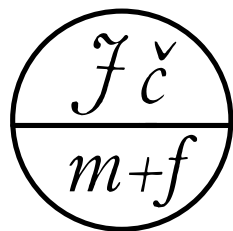




ROBUST 2008

ROBUST 2008

Sborník prací 15. letní školy JČMF ROBUST 2008
uspořádané Jednotou českých matematiků a fyziků
za podpory CQR, ČStS, KPMS MFF UK a ÚM SAV
ve dnech 8. – 12. září 2008 v Pribylině (SK)



Praha 2009

Všechna práva vyhrazena. Tato publikace ani žádná její část nesmí být reprodukována nebo šířena v žádné formě, elektronické nebo mechanické, včetně fotokopí, bez písemného souhlasu vydavatele.

© (eds.) Jaromír Antoch a Gejza Dohnal

© Jednota českých matematiků a fyziků

ISBN 978-80-7015-004-7

JČMF 57551-09

ROBUST 2008 – PÁR SLOV ÚVODEM

Ve dnech 8. – 12. září 2008 se v hotelu Mier v podhůří Roháčů uskutečnila již patnáctá letní škola JČMF ROBUST 2008. Tato akce byla organizována skupinou pro výpočetní statistiku ČMS JČMF za podpory CQR, ČStS, KPMS MFF UK a ÚM SAV. Tak jako v minulosti, i tentokrát byl ROBUST věnován vybraným trendům matematické statistiky, teorie pravděpodobnosti a analýzy dat. Počet účastníků z pěti evropských zemí (České republiky, Slovenska, Holandska, Itálie a Rakouska) přesáhl sto dvacet.

Mezi účastníky bylo k naší velké radosti mnoho mladých tváří¹. Téměř polovinu účastníků totiž tvořili pregraduální a postgraduální studenti či ti, kteří obhájili doktorskou práci v roce 2008. Doufáme, že naše akce přispěje k tomu, že česko-slovenská stochastika nevymře po meči ani po přeslici. Chtěli bychom také touto cestou poděkovat především firmám ČSOB, a. s., Median, s. r. o., a grantu GA ČR 201/05/H007, které podpořily účast mnoha studentů.

Pozvání přednést následující přehledné přednášky přijali:

- Kolektiv pracovníků oddělení nelineárního modelování ÚI AV ČR, *Zkušenosti se statistickým modelováním spotřeby zemního plynu*,
- Prof. Dr. Roberto Caliere, Wolfram Inc., Itálie, *Mathematica and statistics*,
- RNDr. R. Harman, PhD., KAMS FMFI UK, Bratislava, *Algoritmy optimálního navrhování experimentov*,
- Doc. RNDr. D. Hlubinka, PhD., KPMS MFF UK, Praha, *Výprava do hlubin dat*,
- Prof. RNDr. A. Pázman, DrSc., KAMS FMFI UK, Bratislava, *Ako lepšie porozumieť nelineárnej metóde najmenších štvorcov pomocou geometrie*,
- Prof. RNDr. J. Štěpán, DrSc., KPMS MFF UK, Praha, *Stochastické modelování epidemií*.

Vedle toho bylo předneseno 41 delších příspěvků a 47 krátkých sdělení doplněných posterem.

V soutěži o nejlepší práci studentů a doktorandů odborná komise ve složení prof. M. Hušková (MFF UK v Praze), prof. A. Pázman (MFF UK v Bratislavě), Ing. Z. Roth (SZÚ v Praze) a prof. J. Štěpán (MFF UK v Praze) ocenila práce následujících doktorandů (v abecedním pořadí):

- Barbora Arendacká (MFF UK v Bratislavě),
- Martin Branda (MFF UK v Praze),
- Šárka Došlá (MFF UK v Praze),
- Kateřina Helisová (MFF UK v Praze),
- Michal Pešta (MFF UK v Praze),
- Ján Somorčík (MFF UK v Bratislavě).

¹Nejmłodším účastníkům bylo okolo jednoho roku.

Ocenění obdrželi programy a knihy věnované společnostmi Elkan² a Humusoft³.

Mnoho času též bylo věnováno diskusím. Pondělní večer byl zasvěcen historii stereologie, pravděpodobnosti a statistiky vůbec, úterní večer pak otázkám výpočetní statistiky. Vystoupili na něm zástupci firem *Humusoft* a *Median*, kteří předvedli nejnovější verze jimi šířených programů.

Vedle odborných diskusí se též konaly debaty volnější, ať již během středního výletu k Račkovu plesu, nebo při opékání sice možná nezdravých, ale o to lahodnějších buřtů a klobás. Také tanečníci, zpěváci a milovníci místních specialit si tentokrát přišli na své. Wimmerovci, hombres de tango, čardáš i polka, odpadli až k pátečnímu ránu, zatímco tanečníci až s kuropěním. I počasí nám tentokrát vyšlo! Celou dobu bylo překrásné tatranské babí léto, při němž sluníčko statečně ukazovalo svoji sílu. Zvláště ti, kteří bydleli ve vyšších patrech hotelu, nestačili ráno co ráno obdivovat jeho překrásné východy.

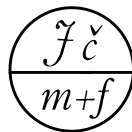
Z proslovených přednášek naleznete převážnou většinu v přiloženém sborníku. Budete se možná divit, že jeden příspěvek je publikován tak říkajíc *per partes*. Editorům totiž přišlo nevhodné, tak jako před dvěma lety, omezovat jedny v „rozmachu pera“ a a na straně dát podstatně více místa druhým. Rozhodli se proto v případě velmi zajímavého článku kolegy Komárka do tištěného sborníku zařadit pouze „krátkou upoutávku“, zatímco rozsáhlejší pokračování čtenář nalezne na CD. Poslední příspěvek dorazil příliš pozdě, takže byl zařazen pouze na CD bez toho, že bychom měnili stránkování.

Z vystavených posterů naleznete téměř všechny na přiloženém CD. Zároveň CD obsahuje elektronickou verzi sborníku a vybrané fotografie. Jinými slovy to znamená, že do sborníku a jeho elektronické přílohy přispěli prakticky všichni přednášející. Chtěli bychom tímto poděkovat nejenom jim, ale též všem těm, kteří články recenzovali.

ROBUST 2008 by se neuskutečnil a jeho publikace by neexistovaly, nebýt nezištné pomoci mnoha lidí. Zvláště bychom chtěli poděkovat občianskemu združení Vendelín, Centru J. Hájka pro aplikovanou a teoretickou statistiku – projekt LC06024, paní Haně Bílkové, pracovníkům tiskárny Vězeňské služby v Praze na Pankráci, panu Zdeňkovi Roubíčkoví ze společnosti Diskus, s.r.o., a grantu GA ČR 201/09/0775. Díky nim vám všem můžeme popřát příjemné čtení.

V Praze 1. května 2009

Jaromír Antoch a Gejza Dohnal



²<http://www.elkan.cz>

³<http://www.humusoft.cz>

Antoch Jaromír	
<i>Použití genetických algoritmů pro analýzu intervalových dat</i>	1
Arendacká Barbora	
<i>Aproximácia zovšeobecneného konfidenčného intervalu pre σ_1^2</i>	9
Boháčová Hana	
<i>Oblasti necitlivosti pro varianční komponenty ve smíšeném</i>	
<i>lineárním regresním modelu</i>	17
Branda Martin	
<i>Recourse reformulation of chance constrained problems</i>	25
Cézová Eliška	
<i>Zónový regulační diagram</i>	33
Cimermanová Katarína	
<i>Klasifikácia zašumených dát</i>	41
Dienstbier Jan	
<i>Problems with extremes</i>	47
Dohnal Gejza	
<i>O optimalizaci regulačních diagramů</i>	55
Dostál Petr	
<i>Robust filtering and portfolio management</i>	65
Fajfrová Lucie	
<i>Generalized zero range process as a traffic model</i>	73
Friesl Michal	
<i>Koziolův-Greenův model se zleva useknutými pozorováními</i>	81
Helisová Kateřina	
<i>Model pro náhodné sjednocení kruhů se vzájemnými interakcemi</i>	89
Hlubinka Daniel	
<i>Výpravy do hlubin dat</i>	97
Hönschová Erika	
<i>Estimation of the scale parameter in Burr type XII distribution</i>	131
Chochola Ondřej	
<i>Kvantily v sekvenční analýze bodu změny</i>	137
Janáček Jiří	
<i>Řešení problémů analýzy obrazu pomocí minima totální variace</i>	145
Janžura Martin	
<i>On the importance of entropy</i>	151
Jarušková Daniela	
<i>Testování normality ve vícerozměrném případě</i>	159
Karlová Andrea	
<i>Lèvy processes and balayage space</i>	171
Klaschka Jan	
<i>Intervalový odhad parametru p binomického rozdělení:</i>	
<i>Co je (relativně) nového?</i>	177
Klášterecký Petr, Kulich Michal	
<i>Logistická regrese v subkohortních studiích</i>	185

Komárek Arnošt	
<i>R balíček pro odhad jedno i vícerozměrných rozdělení pomocí normálních směsí, případně s intervalově cenzorovanými daty</i> . . .	193
Kovářová Milena	
<i>Text mining of articles in life science</i>	201
Král' Pavol, Ďurica Marek, Lašová Lenka	
<i>Základy statistiky s fuzzy datami</i>	205
Kvitkovičová Andrea	
<i>Testing hypotheses about the parameters of the shifted and rescaled Wiener process</i>	215
Lachout Petr, Lstibůrek Milan	
<i>Model přenosu genů</i>	223
Lechnerová Radka, Lechner Tomáš	
<i>Analýza rozmístění elektronických podatelů obcí v České republice</i>	231
Maciak Matúš	
<i>Spline models with change-points</i>	239
Madurkayová Barbora	
<i>Ratio type statistics for detection of changes in linear regression models</i>	247
Marek Jaroslav	
<i>Tajný ostrov templářů</i>	255
Marek Tomáš, Tichavský Petr	
<i>On the estimation of mutual information</i>	263
Marušiaková Miriam	
<i>Multiple changes in coefficients of autoregressive models</i>	271
Neubauer Jiří, Veselý Vítězslav	
<i>Change point detection by BASIS PURSUIT</i>	279
Pázman Andrej	
<i>Nelineární metoda nejmenších štvorců pomocí geometrie</i>	287
Pečánka Jakub	
<i>Change detection in AR processes</i>	305
Pelikán Emil, Brabec Marek, Malý Marek, Konár Ondřej, Vondráček Jiří	
<i>Zkušenosti se statistickým modelováním spotřeby zemního plynu</i>	315
Pešta Michal	
<i>Regression in Sobolev spaces using total least squares</i>	343
Prokešová Michaela	
<i>Moment estimates for inhomogeneous planar cluster processes</i> .	351
Půlpán Zdeněk, Čihák Michal, Bomberová Jaroslava	
<i>Některé míry neurčitosti měření za nejistoty</i>	359
Ryšavá Eva	
<i>Determinanty PZI ve zpracovatelském průmyslu ČR v letech 2000-2006</i>	369

Řezanková Hana, Húsek Dušan <i>Zjišťování optimálního počtu shluků ve statistických programových systémech</i>	377
Somorčík Ján, Rublík František <i>Multi-sample spatial median test not requiring distributions of the same type</i>	385
Staněk Jakub <i>Stochastický model vývoje epidemie s vakcinací</i>	393
Strouhal Jan A. <i>Jak posuzovat spolehlivost softwaru</i>	399
Šedová Michaela, Kulich Michal <i>Maximálně věrohodné odhady a lineární regrese ve výběrových šetřeních</i>	409
Šimečková Marie, Šimeček Petr, Rasch Dieter <i>Testy aditivity</i>	417
Štěpán Josef <i>Mathematical models for epidemics</i>	425
Tuček Pavel, Tuček Jiří, Marek Jaroslav, Procházka Roman <i>Nelineární statistické inference v nanotechnologiích</i>	445
Tvrdlík Josef <i>Adaptivní stochastické algoritmy v nelineární regresi</i>	453
Václavík Vladimír <i>Metoda MLE pro logaritmicko-normální rozdělení s prahem</i>	461
Vajda Igor <i>On efficiencies of decisions about statistical models based on f-divergences of empirical distributions</i>	469
Vávra František, Pavelka Tomáš, Šedivá Blanka, Vokáčová Kateřina, Marek Patrice, Neumanová Martina <i>Ratio statistics</i>	479
Vencálek Ondřej <i>Weighted data depth and its properties</i>	487
Víšek Jan Ámos <i>Consistency of weighted generalized method of moments</i>	495
Wimmer Gejza, Jr. <i>Algoritmus výpočtu přibližných konfidenčních intervalov parametra polohy z digitalizovaných meraní</i>	505
Witkovský Viktor, Wimmer Gejza <i>Konfidenční intervaly založené na digitalizovaných meraniach</i> .	513
Žambochová Marta <i>Shlukování velkých souborů dat pomocí algoritmů k-průměrů</i> ...	533
Timková Jana <i>Testing homogeneity and goodness of fit in survival data</i>	541

Komárek Arnošt

A new R package for Bayesian estimation of multivariate normal mixtures allowing for selection of the number of components and interval-censored data 549

POUŽITÍ GENETICKÝCH ALGORITMŮ PRO ANALÝZU INTERVALOVÝCH DAT

Jaromír Antoch

Klíčová slova: Genetické algoritmy, intervalová data, míry variability.

Abstrakt: Cílem článku je ukázat možnost použití genetických algoritmů pro analýzu intervalových dat.

Abstract: Goal of the paper is to show how genetic algorithms can be used for analysis of interval data.

1 Problém

V reálné praxi přibývá dat intervalového typu. Narazíme na ně zpravidla tehdy, kdy nejsme schopni sledovanou veličinu přesně změřit, nýbrž jsme pouze nuceni konstatovat, že hodnota leží v nějakém intervalu. Jako typický příklad si vezmeme úlohu délku onemocnění pacienta apod. V informatice k zájmu o tuto problematiku přispívá fakt, že i „přesně měřené“ hodnoty jsou v počítači pouze „nepřesně aproximovány“.

Předpokládejme, že k dispozici nejsou přesně naměřené hodnoty, nýbrž pouze intervaly v nichž jednotlivá data mohou ležet, a zajímá nás, v jakém rozmezí mohou ležet vybrané charakteristiky takovýchto dat. V našem případě se jedná o míry variability jako rozptyl apod. Na získané výsledky se ze statistického hlediska můžeme dívat jako na analýzu citlivosti uvažovaných metod v případě, kdy jsou klasická „přesná“ data nahrazena daty intervalovými. Hlavní výhodou použití genetických algoritmů je, vedle poměrně snadné implementovatelnosti, především skutečnost, že jsou schopny nás v řadě případů dovést k řešení mnohem rychleji než úplné prohledávání stavového prostoru.

Mějme intervaly $I_i = [a_i, b_i], a_i \leq b_i, i = 1, \dots, n$, a označme \mathcal{K} jejich kartézský součin, tj. $\mathcal{K} = I_1 \otimes \dots \otimes I_n \subset \mathbb{R}_n$. Naším cílem budiž:

P1. Nalézt mezi všemi body z \mathcal{K} ten, který má největší rozptyl, tj.

$$\mathbf{x}^{\max} = \arg \max_{\mathbf{x} \in \mathcal{K}} \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}_n)^2. \quad (1)$$

P2. Nalézt mezi všemi body z \mathcal{K} ten, který má nejmenší rozptyl, tj.

$$\mathbf{x}^{\min} = \arg \min_{\mathbf{x} \in \mathcal{K}} \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}_n)^2. \quad (2)$$

P3. Nalézt mezi všemi body z \mathcal{K} ten, pro který je maximální střední vzdálenost jeho složek od jejich mediánu, tj.

$$\tilde{\mathbf{x}}^{\max} = \arg \max_{\mathbf{x} \in \mathcal{K}} \frac{1}{n} \sum_{i=1}^n \left| x_i - \operatorname{med}_{1 \leq j \leq n} x_j \right|, \quad (3)$$

P4. Nalézt mezi všemi body z \mathcal{K} ten, pro který je minimální střední vzdálenost jeho složek od jejich mediánu, tj.

$$\tilde{\mathbf{x}}^{\min} = \arg \min_{\mathbf{x} \in \mathcal{K}} \frac{1}{n} \sum_{i=1}^n \left| x_i - \operatorname{med}_{1 \leq j \leq n} x_j \right|. \quad (4)$$

Podobné úlohy dostaneme, nahradíme-li v (3) a (4) medián průměrem nebo použijeme-li L_∞ normu místo L_1 či L_2 normy. Zde je dobré si všimnout, že pokud v (3) a (4) nahradíme $\operatorname{med}_{1 \leq j \leq n} x_j$ průměrem \bar{x}_n , tj. hledáme-li

$$\tilde{\mathbf{x}}^{\max} = \arg \max_{\mathbf{x} \in \mathcal{K}} \frac{1}{n} \sum_{i=1}^n \left| x_i - \bar{x}_n \right| \quad \text{či} \quad \tilde{\mathbf{x}}^{\min} = \arg \min_{\mathbf{x} \in \mathcal{K}} \frac{1}{n} \sum_{i=1}^n \left| x_i - \bar{x}_n \right|, \quad (5)$$

řešení je totéž jako pro úlohu (1) a (2). Máme-li tedy vyřešeny úlohy (1) a (2), máme vyřešenu úlohu (5) a naopak.

Na analogické problémy narazíme při výpočtu charakteristik polohy, závislosti, měr informace, lineární regrese, konfidenčních intervalů, apod. Z hlediska výpočetní složitosti se touto problematikou již delší dobu zabývá skupina V. Kreinoviche, viz [3], [4] či [7]. Statistické aspekty jsou živě diskutovány v komunitě statistiků zabývajících se intervalovými daty, viz např. [2].

2 Řešení využívající optimalizačních metod

Na první pohled je zřejmé, že každá z výše uvedených úloh může mít několik řešení, a že řešení může být i nekonečně mnoho. Ukažme to podrobně. Za tím účelem je vhodné si přepsat vzorec pro rozptyl ve tvaru

$$\operatorname{var} \mathbf{x} = \frac{1}{n} \left[\sum_{i=1}^n x_i^2 - \frac{1}{n-1} \sum_{1 \leq i \neq j \leq n} x_i x_j \right] = \frac{1}{n(n-1)} \sum_{1 \leq i < j \leq n} (x_i - x_j)^2, \quad (6)$$

odkud okamžitě vidíme, že platí následující tvrzení:

Věta 1. Předpokládejme, že $\mathcal{I} = \cap_i I_i \neq \emptyset$. Potom existuje buď jedno nebo nekonečně mnoho řešení úloh P2 a P4. Jedno řešení existuje tehdy, je-li \mathcal{I} jednobodová množina, nekonečně mnoho řešení existuje tehdy, jedná-li se o interval. Jde přitom o ta $\mathbf{x} = (x_1, \dots, x_n)' \in \mathcal{K}$, pro něž

$$a_{\max} \leq x_1 \leq b_{\min} \quad \text{a} \quad x_i = x_j \quad \forall 1 \leq i, j \leq n,$$

$$a_{\min} = \min_{1 \leq i \leq n} a_i, \quad a_{\max} = \max_{1 \leq i \leq n} a_i, \quad b_{\min} = \min_{1 \leq i \leq n} b_i, \quad b_{\max} = \max_{1 \leq i \leq n} b_i.$$

Hodnota rozptylu, resp. absolutní mediánové odchylky, je pro všechna tato řešení rovna nule.

Podíváme-li se na (6) pro potřeby úloh P1 a P2 z geometrického hlediska, vidíme, že nejde o nic jiného než o hledání takových bodů v \mathcal{K} , které

leží nejdále (nejblíže) přímce \mathcal{L} procházející počátkem a bodem $(1, \dots, 1)'$. Charakterizaci řešení shrnuje:

Věta 2. Řešení úloh P1 a P2 leží na hranici \mathcal{K} . Navíc, řešení úlohy P1 se shoduje s jedním nebo více vrcholy \mathcal{K} .

Důkaz: Viz [1]. □

Další výhodou vztahu (6) a Věty 2 je to, že umožňují úlohy P1 a P2 převést na úlohy kvadratického programování. Skutečně, položíme

$$\mathbf{H} = \begin{pmatrix} \frac{n-1}{n} & \dots & -\frac{1}{n} \\ \vdots & \ddots & \vdots \\ -\frac{1}{n} & \dots & \frac{n-1}{n} \end{pmatrix}, \quad \mathbf{G} = 2\mathbf{H}, \quad \mathbf{lb} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \quad \mathbf{ub} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}, \quad (7)$$

kde diagonální prvky matice \mathbf{H} jsou rovny $(n-1)/n$, zatímco všechny ostatní prvky jsou rovny $-1/n$. Potom úloha P1 odpovídá úloze kvadratického programování

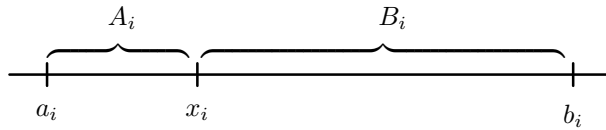
$$\min \left\{ -\frac{1}{2} \mathbf{x}' \mathbf{G} \mathbf{x} \right\} \quad \text{při omezení} \quad \mathbf{lb} \leq \mathbf{x} \leq \mathbf{ub}, \quad (8)$$

zatímco úloha P2 odpovídá úloze kvadratického programování

$$\min \left\{ \frac{1}{2} \mathbf{x}' \mathbf{G} \mathbf{x} \right\} \quad \text{při omezení} \quad \mathbf{lb} \leq \mathbf{x} \leq \mathbf{ub}. \quad (9)$$

Poznamenejme, že pro úlohy tohoto druhu existuje celá řada více či méně dobrých softwarových řešičů. Je však třeba být opatrný a trpělivý, k hledanému řešení nevede v praxi cesta zdaleka tak přímočará jako na papíře.

Podívejme se dále na úlohy P3 a P4. Nechť $x_i \in [a_i, b_i]$, $1 \leq i \leq n$, označuje hodnotu i -té souřadnice hledaného řešení. Pro každý interval I_i zavedme pomocné proměnné A_i a B_i reprezentující vzdálenost mezi a_i a x_i , resp. mezi b_i a x_i tak, jak zobrazuje následující obrázek.



Označme $r_i = a_i + A_i - c_A$, resp. $r_i = b_i - B_i - c_B$, kde c_A , resp. c_B , je neznámá hodnota mezi $1 \leq j \leq n$ x_j , a položme $\mathbf{r} = (r_1, \dots, r_n)'$. Dále zavedme matice

$$\mathbf{f} = \begin{pmatrix} \mathbf{1}_{n,1} \\ \mathbf{1}_{n,1} \\ \mathbf{0}_{n,1} \\ 0 \end{pmatrix}, \quad \mathbf{y}_A = \begin{pmatrix} \mathbf{r}^+ \\ \mathbf{r}^- \\ \mathbf{A} \\ c_A \end{pmatrix}, \quad \mathbf{y}_B = \begin{pmatrix} \mathbf{r}^+ \\ \mathbf{r}^- \\ \mathbf{B} \\ c_B \end{pmatrix}, \quad \mathbf{lb} = \begin{pmatrix} \mathbf{0}_{n,1} \\ \mathbf{0}_{n,1} \\ \mathbf{0}_{n,1} \\ a_{\min} \end{pmatrix}, \quad \mathbf{ub} = \begin{pmatrix} \mathbf{b} - \mathbf{a} \\ \mathbf{b} - \mathbf{a} \\ \mathbf{b} - \mathbf{a} \\ b_{\max} \end{pmatrix},$$

kde $\mathbf{A}eq = (\mathbf{I}_n \dot{\vdash} -\mathbf{I}_n \dot{\vdash} -\mathbf{I}_n \dot{\vdash} \mathbf{1}_{n,1})$, $\mathbf{B}eq = (\mathbf{I}_n \dot{\vdash} -\mathbf{I}_n \dot{\vdash} \mathbf{I}_n \dot{\vdash} \mathbf{1}_{n,1})$, $\mathbf{a}eq = \mathbf{a}$, $\mathbf{b}eq = \mathbf{b}$, \mathbf{r}^+ označuje vektor kladných částí a \mathbf{r}^- vektor záporných částí vektoru \mathbf{r} , $\mathbf{A} = (A_1, \dots, A_n)'$, $\mathbf{B} = (B_1, \dots, B_n)'$, $\mathbf{0}_{n,1}$ vektor nul a $\mathbf{1}_{n,1}$ vektor jedniček řádu $n \times 1$.

Potom úloha P4 odpovídá buď problému lineárního programování

$$\min \mathbf{f}' \cdot \mathbf{y}_A \quad \text{při omezení} \quad \mathbf{A}eq \cdot \mathbf{y}_A = \mathbf{a}eq \quad \text{a} \quad \mathbf{lb} \leq \mathbf{y}_A \leq \mathbf{ub}, \quad (10)$$

nebo

$$\min \mathbf{f}' \cdot \mathbf{y}_B \quad \text{při omezení} \quad \mathbf{B}eq \cdot \mathbf{y}_B = \mathbf{b}eq \quad \text{a} \quad \mathbf{lb} \leq \mathbf{y}_B \leq \mathbf{ub}, \quad (11)$$

který lze řešit například použitím simplexové metody. Hledané řešení dostaneme jako $\tilde{\mathbf{x}}^{\min} = \mathbf{a} + \mathbf{y}_A(2n+1 : 3n)$ nebo jako $\tilde{\mathbf{x}}^{\min} = \mathbf{b} - \mathbf{y}_B(2n+1 : 3n)$. Pro úlohu P3 pouze nahradíme v (10) a (11) minimum maximem.

Poznamenejme, že použití reprezentace s pomocnými proměnnými A_i a B_i umožňuje přeformulovat také úlohy P1 a P2 a převést je (opět) na úlohy kvadratického programování, tentokrát v jiných proměnných. Reprezentaci s použitím A_i a B_i lze užít také v situaci, kdy v (3) a (4) nahradíme medián průměrem, nebo použijeme-li L_∞ normu místo L_1 či L_2 normy.

3 Genetické algoritmy

Metodika genetických algoritmů byla inspirována Darwinovou teorií evoluce, a proto i názvosloví odpovídá biologickým termínům. Vzhledem k tomu, že ne všichni čtenáři jsou s danou problematikou obeznámeni, připomeňme několik nezbytných pojmů. Další podrobnosti čtenář nalezne např. v [5], [8] či [9]. Učiníme tak na značně zjednodušeném problému, který, ne náhodou, úzce připomíná úlohy P1–P4.

Mějme funkci $f : \{0, 1\}^n \rightarrow \mathbb{R}_1$, $n \in \mathbb{N}$, kterou chceme maximalizovat, a zároveň hledíme bod ve kterém je maxima dosaženo, tj.

$$\boldsymbol{\alpha}^{opt} = \arg \max_{\boldsymbol{\alpha} \in \{0, 1\}^n} f(\boldsymbol{\alpha}). \quad (12)$$

Bod $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n)'$ nazveme jedincem, *chromozomem*. Je to nějaký prvek prostoru $\{0, 1\}^n$, na kterém funkci f maximalizujeme. Jednotlivé složky tohoto vektoru budeme v souladu s biologickou terminologií nazývat *geny*. Množinu $\mathcal{P} = \{\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_S\} \subset \{0, 1\}^n$ nazveme populací jedinců rozsahu S .

Funkce f reprezentuje prostředí ve kterém jedinci dané populace žijí, jedinec $\boldsymbol{\alpha}$ reprezentuje genotyp organismu a hodnota $f(\boldsymbol{\alpha})$ jeho fenotyp. Mírou síly chromozomu může být například jeho funkční hodnota. Funkci f chceme maximalizovat, takže čím větší je hodnota $f(\boldsymbol{\alpha})$, tím by měl být jedinec $\boldsymbol{\alpha}$ silnější. To nás vede k zavedení funkce $F : \{0, 1\}^n \rightarrow \mathbb{R}_1$, tzv. *fitness*, jež hodnotí sílu jedince. Tato funkce musí pro úlohu (12) splňovat podmínku

$$\forall \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 \in \mathcal{P} \quad f(\boldsymbol{\alpha}_1) \leq f(\boldsymbol{\alpha}_2) \Rightarrow F(\boldsymbol{\alpha}_1) \leq F(\boldsymbol{\alpha}_2). \quad (13)$$

Normalizovanou fitness rozumíme funkci $F'(\alpha) = F(\alpha) / (\sum_{\alpha' \in \mathcal{P}} F(\alpha'))$. Hodnotu normalizované fitness jedince v rámci dané populace můžeme interpretovat tak, že nám poskytuje pravděpodobnost přežití daného jedince, tj. odpovídá šanci jedince na reprodukci. Možností, jak volit fitness, je celá řada. Pro úlohu (12) přiřadíme chromozomům hodnoty funkce kterou maximalizujeme, tj. $F(\alpha) = f(\alpha)$.

Poznamenejme, že kdybychom v (12) nahradili funkci $\arg \max$ funkcí $\arg \min$, potom je třeba podmínku na funkci fitness změnit na

$$\forall \alpha_1, \alpha_2 \in \mathcal{P} \quad f(\alpha_1) \leq f(\alpha_2) \Rightarrow F(\alpha_1) \geq F(\alpha_2).$$

Pro problém P2 se osvědčila volba $F(\alpha) = 1/f(\alpha)$.

Nejdůležitějším operátorem na prostoru jedinců \mathcal{P} je *křížení chromozomů*, jež napodobuje svou biologickou předlohu. Možností, jak tuto operaci zavést, je celá řada. Základním typem, který se nejvíce osvědčil i pro náš problém, je tzv. jednobodové křížení. Operace spočívá v tom, že náhodně vybereme bod křížení $k \in \{1, \dots, n-1\}$, po bod křížení do nových chromozomů zkopírujeme obsah chromozomů původních, a za tímto bodem jejich obsah prohodíme. Schematicky můžeme jednobodové křížení znázornit tak, že chromozomy

$$\alpha = (\alpha_1, \dots, \alpha_k, \alpha_{k+1}, \dots, \alpha_n) \quad \text{a} \quad \beta = (\beta_1, \dots, \beta_k, \beta_{k+1}, \dots, \beta_n)$$

dají po zkřížení

$$\alpha' = (\alpha_1, \dots, \alpha_k, \beta_{k+1}, \dots, \beta_n) \quad \text{a} \quad \beta' = (\beta_1, \dots, \beta_k, \alpha_{k+1}, \dots, \alpha_n).$$

Křížení samozřejmě můžeme analogicky definovat jako dvou-, tří- i vícebodové, respektive náhodné. Pro náš problém ani vícebodové ani náhodné křížení žádné podstatné vylepšení nepřineslo.

Poslední základní operací je *mutace chromozomů*, jež vnáší, stejně jako v přírodě, do celého procesu evoluce něco nového. Zde slouží zejména k tomu, aby se algoritmus dostal z lokálních extrémů. Mutací jednoho či více genů chromozomu můžeme získat jedince, který se v populaci vůbec nenachází. Jeho následným zkřížením s chromozomy ostatními tak může GA dostat zcela nový směr. Pro úlohu (12) se osvědčil základní, a v praxi také nejrozšířenější, typ mutace, kdy pro pevně zvolenou pravděpodobnost mutace p_M postupně generujeme náhodná čísla R_i z $R[0, 1]$, a každý gen buď změníme nebo ponecháme nezměněný podle toho, zda k němu nagerované náhodné číslo R_i nepřekročí, resp. překročí, p_M . Jinými slovy, pro každý chromozom $\alpha \in \mathcal{P}$ vytvoříme nový chromozom $\alpha^* = (\alpha_1^*, \dots, \alpha_n^*)'$ tak, že

$$\alpha_i^* = \begin{cases} 1 - \alpha_i, & R_i < p_M, \\ \alpha_i, & R_i \geq p_M, \end{cases} \quad i = 1, \dots, n.$$

Nyní jsme již připraveni popsat základní postup genetického algoritmu. Na počátku náhodně vygenerujeme populaci chromozomů z $\{0, 1\}^n$, tedy

$\mathcal{P}_0 = \{\alpha_1, \dots, \alpha_S\} \subset \{0, 1\}^n$, přičemž pod náhodně vygenerovanou populací rozumíme to, že jednotlivé geny pocházejí z alternativního rozdělení $Alt(1/2)$. Všechny chromozomy populace \mathcal{P}_0 ohodnotíme pomocí normalizované fitness. Chromozom z \mathcal{P}_0 je vybírán jako rodičovský chromozom další generace s pravděpodobností rovnající se jeho normalizované fitness. Takto náhodně vybíráme dvojice chromozomů z \mathcal{P}_0 , aplikujeme na ně reprodukční operátor, tj. zkřížíme je mezi sebou, a výsledek následně ještě zmutujeme. Výběry provádíme s vrácením. Takto postupujeme tak dlouho, dokud nemáme m nových chromozomů, kde $m \in \mathbb{N}$, $m \geq S$, je pevně zvolené číslo. Z m nových chromozomů vybereme S s největšími hodnotami účelové funkce, které vytvoří novou populaci \mathcal{P}_1 , chromozomy ohodnotíme normalizovanou fitness, atd. Celý postup opakujeme tak dlouho, dokud není splněna nějaká ukončovací podmínka. Stále si přitom udržujeme informaci o několika doposud nejlepších řešeních dané úlohy.

Zatímco někteří autoři doporučují doposud nejlepšího jedince (několik doposud nejlepších jedinců) v populaci \mathcal{P} neustále udržovat, jiní před touto modifikací spíše varují. V našem případě se ukázalo, že udržet doposud nejlepšího jedince v populaci se jednoznačně vyplácí, neboť k řešení dokongrujeme mnohem rychleji. Zkusili jsme dokonce variantu, před kterou většina „guru“ v oblasti genetických algoritmů varuje, totiž, vždy křížit danou populaci \mathcal{P}_\bullet s doposud nejlepším jedincem. Obavou je, že tato strategie může snadno skončit v některém lokálním minimu. Je třeba říci, že tato strategie vedla v případě úlohy P1 jednoznačně k nejrychlejšímu dosažení řešení, a to i v úlohách s několika tisíci intervaly.

Ukončovací podmínkou může být například maximální počet generací, které projdeme. Další možností je zastavit algoritmus tehdy, jsou-li všichni jedinci populace stejní nebo je-li hodnota maximalizované (minimalizované) funkce po určitý počet generací „prakticky“ táž, apod. Vzhledem k tomu, že slabší jedinci jsou ohodnoceni malou fitness, k reprodukci se dále prakticky neprosadí, neboť jsou vybíráni jako rodiče další generace jen s velmi malou pravděpodobností. Na základě tohoto typu výběru k reprodukci bychom měli postupně napříč generacemi získávat jedince stále „silnější“. Důkaz konvergence genetického algoritmu lze nalézt například v [6].

4 Příklad

Vraťme se nyní k úloze P1 a podívejme se na ni pohledem genetických algoritmů. Reprezentace rozptylu ve tvaru (6) má tu velkou výhodu, že je dovoluje na daný problém velmi efektivně použít, neboť Věta 2 umožňuje se soustředit pouze na vrcholy \mathcal{K} . Ke každému vrcholu totiž můžeme jednoznačně přiřadit některý vektor $\alpha = (\alpha_1, \dots, \alpha_n)' \in \{0, 1\}^n$, přičemž hodnota $\alpha_i = 1$ odpovídá volbě $x_i = a_i$, resp. hodnota $\alpha_i = 0$ odpovídá volbě $x_i = b_i$. K jednotlivým vrcholům tak máme naprosto přirozeně přiřazeny chromozomy, funkce fitness je dána hodnotou rozptylu spočtenou pro daný vrchol \mathbf{x} , takže pro vlastní implementaci stačí zvolit rozsah populace S , bod křížení k ,

pravděpodobnost mutace p_M a ukončovací pravidlo. Uvažujme krychli \mathcal{K} vytvořenou z následujících čtyřiceti intervalů, tj.

a_i	b_i	a_i	b_i	a_i	b_i	a_i	b_i
-47.50	28.75	40.75	95.00	-10.00	07.75	27.75	57.00
47.25	91.75	-47.00	30.50	-81.50	-72.25	49.75	85.50
38.50	81.50	-95.75	-25.25	-94.00	-18.75	12.75	90.75
-53.50	45.00	-34.00	-28.25	-65.50	22.00	18.50	85.50
-46.50	93.50	-1.25	34.50	-3.25	83.25	-93.00	-83.00
-98.25	-40.75	16.50	81.25	-90.25	-77.00	-95.75	-52.00
-34.50	-28.00	-21.75	39.75	-24.75	-01.25	-66.00	-61.25
51.00	64.00	30.25	80.25	42.50	79.75	-77.50	-42.75
-88.50	-71.50	-14.50	-06.00	-11.00	-05.00	-32.50	-13.50
-93.75	-33.00	-22.50	1.25	-19.50	06.75	-42.25	20.00

Nalezení řešení pro P1 je při použití vyčerpávajícího prohledávání na hranici spočitatelnosti v „rozumném čase“, neboť projít všech 2^{40} vrcholů a spočítat pro každý z nich hodnotu rozptylu si vyžádalo na počítači s 16ti vícejádrovými procesory více než čtyři hodiny výpočtů, což je pro praxi samozřejmě neúnosné. Podle [4] je přitom exhaustivní prohledávání jediným způsobem, který se 100%ní jistotou zaručuje nalezení globálního extrému pro „obecné intervaly“ tohoto typu. Naproti tomu genetický algoritmus našel totéž řešení při vhodném nastavení parametrů již po několika desítkách/stovkách generací. Pro zajímavost uvedme, že globálního maxima je dosaženo v bodě s vrcholy

-47.50	91.75	81.50	45.00	93.50	-98.25	-34.50	64.00
95.00	-47.00	-95.75	-34.00	34.50	81.25	39.75	80.25
7.75	-81.50	-94.00	-65.50	83.25	-90.25	-24.75	79.75
57.00	85.50	90.75	85.50	-93.00	-95.75	-66.00	-77.50
-88.50	-93.75	-14.50	-22.50	-11.00	-19.50	-32.50	-42.25

Hodnota rozptylu v tomto bodě je rovna 5 037.948.

Podívejme se na nastavení parametrů použitého genetického algoritmu. Na základě doporučení z literatury a „doladění“ pomocí simulací se jako „optimální“ ukázala následující volba:

- Počáteční generace byla zvolena náhodně, tj. jednotlivé geny byly generovány z alternativního rozdělení $Alt(1/2)$.
- Bod křížení $k \approx 0.6n$.
- Pravděpodobnost mutace $p_M \approx 0.01$.
- Fitness $f(\alpha) = var \mathbf{x}$, kde \mathbf{x} je vrchol \mathcal{K} odpovídající chromozomu α .
- Rozsah populace $card(S) = 100$.
- Počet generací 300.

Dále se podívejme na to, jaký vliv na řešení má změna těchto parametrů. Hlavním závěrem je, že rozsah populace a počet generací podstatně závisí na pravděpodobnosti mutace, mnohem méně již na volbě ostatních parametrů.

- Volba počáteční generace nemá podstatný vliv na rychlost dosažení optimálního řešení.
- Volba bodu křížení nemá podstatný vliv na rychlost dosažení optimálního řešení. Podobně dobré výsledky jako křížení v pevném bodě dává křížení náhodné.
- Zvyšujeme-li pravděpodobnost mutace, je zapotřebí podstatně zvýšit buď počet generací nebo rozsah populace. Tak například volba $p_M = 0.025$, kterou literatura doporučuje pro tento typ úloh ($p_M \approx 1/n$), si vyžádala buď zdvojnásobení rozsahu populace nebo ztrojnásobení počtu generací.

Literatura

- [1] Antoch J., Brzezina M., Miele R. *A note on variability of interval data*. Přijato k publikaci v J. Computational Statistics.
- [2] J. Computational Statistics. Special Issue on Interval Data, **21(2)**, 2006.
- [3] Ferson et al. *Computing variance for interval data is NP-hard*. <http://www.cs.utep.edu/vladik/2002/tr02-13b.pdf>
- [4] Ferson S. et al. (2007). *Experimental uncertainty estimation and statistics for data having interval uncertainty*. Technical report SAND2007-0939, May 2007, Sandia National Laboratories.
- [5] Goldberg D.E. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison–Wesley.
- [6] Hartl R.F. (1990). *A global convergence proof for a class of genetic algorithms*. Technische Universität Wien.
- [7] Kreinovich V. et al. (2006). *Towards combining probabilistic and interval uncertainty in engineering calculations: Algorithms for computing statistics under interval uncertainty, and their computational complexity*. Reliable Computing **12**, 471–501.
- [8] Kvasnička V., Pospíchal J., Tiňo P. (2000). *Evoluční algoritmy*. STU Bratislava.
- [9] Reeves C.R., Rowe J.E. (2002). *Genetic Algorithms: Principles and Perspectives: A Guide to GA Theory*. Kluwer.
- [10] Xiang G., Ceberli M., Kreinovich V. (2007). *Computing population variance and entropy under interval uncertainty: Linear-time algorithm*. Reliable Computing **13(6)**, 467–488.

Poděkování: Tato práce byla podporována granty GAČR 201/09/0755 a MSM 0021620839.

Adresa: J. Antoch, MFF UK, KPMS, Sokolovská 83, CZ–186 75 Praha 8

E-mail: jaromir.antoch@mff.cuni.cz

APROXIMÁCIA ZOVŠEOBECNENÉHO KONFIDENČNÉHO INTERVALU PRE σ_1^2

Barbora Arendacká

Kľúčové slová: Lineárny zmiešaný model, variančné komponenty, zovšeobecnený pivot, približný konfidenčný interval.

Abstrakt: V príspevku sa zaoberáme zovšeobecneným konfidenčným intervalom pre variančný komponent zodpovedajúci náhodnému faktoru v lineárnych zmiešaných modeloch, ktorý je založený na Parkovom-Burdickovom zovšeobecnenom pivate, a približným, El-Bassiouniho-Williamsovým-Tukeyho intervalom pre spomínaný variančný komponent. Ukážeme, že okrem očividných podobností medzi oboma procedúrami, El-Bassiouniho-Williamsov-Tukeyho interval môžeme v istom zmysle považovať za aproximáciu spomenutého zovšeobecneného intervalu. Tento výsledok je inšpirovaný prácou Brossa (Biometrics, 6, 1950).

Abstract: A generalized confidence interval for the variance component corresponding to the random factor in a linear mixed model with two variance components based on the Park-Burdick generalized pivot and an approximate, the El-Bassouni-Williams-Tukey, interval for the variance component are considered. Besides apparent similarities, we show that the El-Bassiouni-Williams-Tukey interval can be regarded as an approximation of the generalized interval. This result is inspired by Bross (Biometrics, 6, 1950).

1 Model a značenie

Uvažujeme lineárny zmiešaný model

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{a} + \boldsymbol{\epsilon} \quad (1)$$

kde \mathbf{Y} ($n \times 1$) je vektor pozorovaní, \mathbf{X} ($n \times p$) je známa matica nezávislých premenných, $\boldsymbol{\beta}$ ($p \times 1$) je vektor neznámych parametrov, $\boldsymbol{\epsilon}$ ($n \times 1$) je vektor náhodných chýb, \mathbf{Z} ($n \times q$) je známa matica a \mathbf{a} ($q \times 1$) je nepozorovateľný náhodný vektor. Predpokladáme, že \mathbf{a} a $\boldsymbol{\epsilon}$ sú navzájom nezávislé a každý z nich má mnohorozmerné normálne rozdelenie s nulovou strednou hodnotou a kovariančnou maticou $\sigma_1^2 \mathbf{I}_q$, resp. $\sigma^2 \mathbf{I}_n$ (kde \mathbf{I}_n je jednotková matica $n \times n$), $\sigma_1^2 \geq 0$, $\sigma^2 > 0$. Ďalej predpokladáme, že $\mathcal{R}(\mathbf{Z}) \not\subseteq \mathcal{R}(\mathbf{X})$, kde $\mathcal{R}(\mathbf{A})$ označuje lineárny priestor generovaný stĺpcami matice \mathbf{A} , a tiež že $n > \text{rank}([\mathbf{X}, \mathbf{Z}])$, čo zaručuje existenciu jednoduchého odhadu pre rušivý parameter σ^2 . Táto posledná podmienka sa dá chápať tak, že máme k dispozícii viac ako len minimálny počet pozorovaní, a teda nie je veľmi obmedzujúca.

Neznáme variančné komponenty σ_1^2 , σ^2 sú invariantné vzhľadom na posunutie v strednej hodnote. Môžeme teda uplatniť princíp invariencie,

skonstruovať maximálny invariant $\mathbf{B}^T \mathbf{Y}$ ($\mathbf{B}^T \mathbf{B} = \mathbf{I}_{n-\text{rank}(\mathbf{X})}$, $\mathbf{B} \mathbf{B}^T = \mathbf{I}_n - \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$) vzhľadom na túto transformáciu a následným uplatnením princípu postačujúcosti obdržíme minimálnu postačujúcu štatistiku $\mathbf{U} = (U_1, \dots, U_r)^T$ pre $(\sigma_1^2, \sigma^2)^T$ v triede rozdelení

$$\mathbf{B}^T \mathbf{Y} \sim N_{n-\text{rank}(\mathbf{X})}(\mathbf{0}, \sigma_1^2 \mathbf{B}^T \mathbf{Z} \mathbf{Z}^T \mathbf{B} + \sigma^2 \mathbf{I}_{n-\text{rank}(\mathbf{X})}).$$

Zložky \mathbf{U} sú navzájom nezávislé náhodné premenné

$$U_i \sim (\lambda_i \sigma_1^2 + \sigma^2) \chi_{\nu_i}^2, i = 1, \dots, r,$$

kde $\lambda_1 > \dots > \lambda_r = 0$ sú vlastné čísla matice $\mathbf{B}^T \mathbf{Z} \mathbf{Z}^T \mathbf{B}$ a $\nu_1, \dots, \nu_r > 0$ sú ich násobnosti (pozri napr. [7], str. 880). Práve vďaka našim predpokladom ($\mathcal{R}(\mathbf{Z}) \not\subseteq \mathcal{R}(\mathbf{X})$, $n > \text{rank}([\mathbf{X}, \mathbf{Z}])$) platí, že $\lambda_r = 0$, t.j. $U_r \sim \sigma^2 \chi_{\nu_r}^2$, a $r \geq 2$. Spomeňme ešte, že v najjednoduchšom prípade modelu (1), t.j. vyváženom modeli jednoduchého triedenia s náhodným efektom $Y_{ij} = \mu + a_i + \epsilon_{ij}$, $i = 1, \dots, I > 1$, $j = 1, \dots, J > 1$, je $r = 2$ a štatistiky U_1, U_2 sú známe sumy štvorcov z tabuľky analýzy rozptylu.

V ďalšom texte používame: $s = \sum_{i=1}^{r-1} \nu_i$, $\alpha_l, \alpha_u \in (0, 1)$ sú zvolené tak, že $\alpha_u - \alpha_l = 1 - \alpha$, čo je konfidenčná úroveň, na ktorej sa intervaly konštruujú. $\chi_{m;b}^2$, $F_{m,n;b}$ označujú b kvantily príslušného χ^2 a F rozdelenia. Konvergenciu podľa distribúcie značíme \xrightarrow{D} .

2 El-Bassiouniho-Williamsov-Tukeyho interval

V [4] El-Bassiouni navrhol približný konfidenčný interval pre σ_1^2 , ktorého dolnú L_{EBWT} a hornú U_{EBWT} hranicu dostaneme ako nezáporné riešenia alebo nuly, ak nezáporné riešenia neexistujú, nasledujúcich rovníc:

$$\sum_{i=1}^{r-1} \frac{U_i}{\lambda_i L_{EBWT} \frac{\chi_{s;\alpha_u}^2}{s F_{s,\nu_r;\alpha_u}} + \frac{U_r}{\nu_r}} = s F_{s,\nu_r;\alpha_u} \quad (2)$$

$$\sum_{i=1}^{r-1} \frac{U_i}{\lambda_i U_{EBWT} \frac{\chi_{s;\alpha_l}^2}{s F_{s,\nu_r;\alpha_l}} + \frac{U_r}{\nu_r}} = s F_{s,\nu_r;\alpha_l}. \quad (3)$$

(Nezápornosť riešenia sa vyžaduje kvôli nezápornosti parametra σ_1^2 . Nedostaneme potom intervaly so zápornými hranicami, ale na druhej strane, výsledkom procedúry môže byť aj degenerovaný interval $[0, 0]$. Kladenie záporných hraníc konfidenčného intervalu rovných 0 samozrejme ovplyvňuje pravdepodobnosť pokrytia skutočnej nulovej hodnoty σ_1^2 , ktorá by potom aj v prípade presného konfidenčného intervalu bola väčšia ako $1 - \alpha$.)

El-Bassiouni odvodil uvedený interval (pod názvom Waldov interval) ako analógiu Williamsovho-Tukeyho intervalu (pozri [12]), ktorý sa považuje za

dobré približné riešenie v situácii, keď $r = 2$. Všimol si totiž, že Williamsov-Tukeyho interval môžeme zapísať v tvare

$$[d_1 \hat{\sigma}^2 \hat{\rho}_l, d_2 \hat{\sigma}^2 \hat{\rho}_u] \quad (4)$$

kde d_1, d_2 sú korekčné konštanty, $\hat{\sigma}^2 = U_2/\nu_2$ je nevychýlený odhad parametra σ^2 a $\hat{\rho}_l, \hat{\rho}_u$ sú hranice presného¹ konfidenčného intervalu pre $\rho = \sigma_1^2/\sigma^2$. El-Bassiouni si uvedomil, že v prípade, keď $r > 2$, vieme tiež skonštruovať presný konfidenčný interval pre ρ (konkrétne uvažoval ten, ktorý navrhol Wald [11]), ako aj nevychýlený odhad pre σ^2 (U_r/ν_r). Tieto kvantily dosadil do (4) a stupne voľnosti vystupujúce v d_1, d_2 nahradil ich náprotivkami, keď $r > 2$, čiže $d_1 = sF_{s,\nu_r;\alpha_u}/\chi_{s;\alpha_u}^2$, $d_2 = sF_{s,\nu_r;\alpha_l}/\chi_{s;\alpha_l}^2$. Po jednoduchých algebraických úpravách dostaneme, že hranice takto vzniknutého intervalu môžeme spočítať z (2), (3). El-Bassiouni odporúča tento interval používať iba na základe výsledkov simulačnej štúdie, v ktorej sa zdalo, že jeho pravdepodobnosť pokrytia nikdy neklesá pod nominálnu úroveň (na rozdiel napr. od známeho Thomasovho-Hultquistovho intervalu [9]). Dá sa však ukázať, že okrem dobrých simulačných výsledkov má El-Bassiouniho interval aj pekné teoretické vlastnosti: existuje dolné ohraničenie pre jeho pravdepodobnosť pokrytia (konkrétne $1 - 2\alpha$) a tiež, keď $\sigma_1^2/\sigma^2 \rightarrow \infty$ alebo keď $\nu_r \rightarrow \infty$ (a σ_1^2 je kladné), sa jeho pravdepodobnosť pokrytia blíži k nominálnej úrovni. Ako presný konfidenčný interval sa El-Bassiouniho interval správajú aj čo sa týka pokrytia skutočnej nulovej hodnoty σ_1^2 . Pozri [1].

Interval daný (2) a (3) nazývame El-Bassiouniho-Williamsov-Tukeyho (EBWT) interval kvôli jeho prepojeniu s Williamsovým-Tukeyho intervalom, na ktorý sa redukuje, keď $r = 2$. EBWT interval nie je len formálnou analógiou Williamsovho-Tukeyho intervalu, ale dá sa odvodiť priamo použitím Williamsovho prístupu v [12] (pozri [1]).

3 Parkov-Burdickov zovšeobecný pivot

Skôr ako uvedieme vzorec pre výpočet Parkovho-Burdickovho (PB) zovšeobecného pivota, stručne priblížime pojem zovšeobecného pivota a spôsob jeho použitia na konštrukciu zovšeobecných konfidenčných intervalov, postupovať budeme podľa [6].

V našom prípade chceme interval pre σ_1^2 založiť na minimálnej postačujúcej štatistike \mathbf{U} , ktorej distribúcia závisí na $\boldsymbol{\sigma} = (\sigma_1^2, \sigma^2)^T$. Označme \mathbf{U}^* nezávislú kópiu \mathbf{U} a nech \mathbf{u} označuje realizáciu \mathbf{U} . Zovšeobecný pivot pre σ_1^2 , ozn. $R(\mathbf{U}, \mathbf{U}^*, \boldsymbol{\sigma})$, je funkcia $\mathbf{U}, \mathbf{U}^*, \boldsymbol{\sigma}$, pre ktorú platí

V1 Podmienené rozdelenie $R(\mathbf{U}, \mathbf{U}^*, \boldsymbol{\sigma})$ pri $\mathbf{U} = \mathbf{u}$ nezávisí na $\boldsymbol{\sigma}$.

V2 Pre každú možnú realizáciu \mathbf{u} závisí $R(\mathbf{u}, \mathbf{u}, \boldsymbol{\sigma})$ na $\boldsymbol{\sigma}$ iba cez σ_1^2 .

¹V skutočnosti je tento interval presný len pre nenulové ρ , skutočnú nulu pokrýva s pravdepodobnosťou väčšou ako $1 - \alpha$, keďže prípadné záporné $\hat{\rho}_l, \hat{\rho}_u$ sa kladú rovné nule.

Hodnota $R(\mathbf{u}, \mathbf{u}, \boldsymbol{\sigma})$ sa nazýva napozorovaná hodnota zovšeobecného pivota R . Konštrukcia zovšeobecnenej oblasti spoľahlivosti pre σ_1^2 pomocou R pripomína postup s obyčajnými pivotmi, akurát sa všetko deje podmienene, pri daných napozorovaných hodnotách. Vďaka V1 vieme pre každú realizáciu \mathbf{u} nájsť podmnožinu $C(\mathbf{u})$ výberového priestoru $R(\mathbf{u}, \mathbf{U}^*, \boldsymbol{\sigma})$, pre ktorú $P(R(\mathbf{u}, \mathbf{U}^*, \boldsymbol{\sigma}) \in C(\mathbf{u})) \geq 1 - \alpha$, pričom rovnosť je možné vždy dosiahnuť pre spojitú náhodnú premennú $R(\mathbf{u}, \mathbf{U}^*, \boldsymbol{\sigma})$. Zovšeobecnú konfidenčnú oblasť potom tvoria všetky hodnoty σ_1^2 , pre ktoré napozorovaná hodnota pivota $R(\mathbf{u}, \mathbf{u}, \boldsymbol{\sigma})$ patrí do $C(\mathbf{u})$. Vďaka V2 závisí táto oblasť len na napozorovaných dátach. Ak chceme garantovať, aby zovšeobecnená konfidenčná oblasť bola intervalom, môžeme bez straty všeobecnosti sprísniť vlastnosť V2 a požadovať, aby pre každé pozorovateľné \mathbf{u} , $R(\mathbf{u}, \mathbf{u}, \boldsymbol{\sigma}) = \sigma_1^2$. Potom budú zovšeobecný konfidenčný interval tvoriť priamo dolný a horný kvantil $q_{\alpha_l}^R(\mathbf{u})$, $q_{\alpha_u}^R(\mathbf{u})$ rozdelenia $R(\mathbf{u}, \mathbf{U}^*, \boldsymbol{\sigma})$. (Obvykle sa volí $\alpha_l = \alpha/2$, $\alpha_u = 1 - \alpha/2$.) To je aj prípad PB zovšeobecného pivota R^{PB} , ktorý je definovaný ako nezáporné riešenie alebo nula, ak nezáporné riešenie neexistuje, rovnice

$$\sum_{i=1}^{r-1} \frac{U_i}{\lambda_i R^{PB} + U_r / Q_r} = \sum_{i=1}^{r-1} Q_i, \quad (5)$$

kde $Q_i = U_i^* / (\lambda_i \sigma_1^2 + \sigma^2)$, $i = 1, \dots, r$. Požadovanie nezápornosti má tie isté dôvody ako v prípade EBWT intervalu. Keďže zovšeobecný konfidenčný interval pre σ_1^2 založený na R^{PB} tvoria vždy dolný a horný kvantil jeho rozdelenia pri daných napozorovaných hodnotách u_i , celú situáciu by sme mohli vnímať aj tak, že sme na základe napozorovaných dát priradili neznámemu parametru σ_1^2 distribúciu, akú má $R^{PB}(\mathbf{u}, \mathbf{U}^*, \boldsymbol{\sigma})$. V takom prípade by totiž intervalový odhad pre σ_1^2 tvorili prirodzene dolný a horný kvantil rozdelenia tohto parametra. Tento pohľad jasne odhaľuje spojitosť zovšeobecnenej inferencie s fiduciálnou metódou zavedenou Fisherom [5], pozri tiež [6]. Nie je preto prekvapujúce, že v prípade $r = 2$ sa interval založený na PB pivote zhoduje s fiduciálnym intervalom pre σ_1^2 , pozri [5], [3].

Všimnime si ešte, že ak by sme vo V2 mali $R(\mathbf{u}, \mathbf{u}^*, \boldsymbol{\sigma})$ pre každé pozorovateľné \mathbf{u} , \mathbf{u}^* , tak $R(\mathbf{u}, \mathbf{U}^*, \boldsymbol{\sigma})$ by bol klasický pivot pre σ_1^2 a $\{\sigma_1^2; R(\mathbf{u}, \mathbf{u}^*, \boldsymbol{\sigma}) \in C(\mathbf{u})\}$ by bola presná konfidenčná oblasť pre σ_1^2 . V skutočnosti je ale \mathbf{U}^* nepozorovateľný a jediná realizovaná hodnota, ktorú máme k dispozícii je \mathbf{u} . Jej použitie namiesto realizácie \mathbf{u}^* spôsobuje, že pravdepodobnosť pokrytia zovšeobecných intervalov nie je teoreticky garantovaná. Často sa preto ich použitie v praxi obhajuje len na základe simulačných výsledkov. O PB zovšeobecnenom pivote sa však dá navyše ukázať, že vedie k intervalom, ktorých pravdepodobnosť pokrytia sa blíži k nominálnej úrovni, keď $\nu_r \rightarrow \infty$ (a $\sigma_1^2 > 0$) alebo keď $\sigma_1^2 / \sigma^2 \rightarrow \infty$. Taktiež, skutočnú nulovú hodnotu σ_1^2 pokrývajú tieto intervaly ako keby boli presné. Pozri [2].

4 Aproximácia

Podľa popisu v predchádzajúcej časti je zrejmé, že pri konštrukcii PB zovšeobecneného intervalu pre σ_1^2 potrebujeme pre konkrétnu realizáciu $\mathbf{u} = (u_1, \dots, u_r)^T$ spočítať dolný a horný kvantil rozdelenia implicitne zadanej náhodnej premennej \tilde{R} , ktorá je daná ako nezáporné riešenie alebo nula, ak nezáporné riešenie neexistuje, rovnice

$$\sum_{i=1}^{r-1} \frac{u_i}{\lambda_i \tilde{R} + u_r/Q_r} = \sum_{i=1}^{r-1} Q_i, \quad (6)$$

pozri (5). Spočítať príslušné kvantily je možné simulačne (náhodným generovaním r -tíc $(Q_1, \dots, Q_r)^T$, spočítaním hodnoty \tilde{R} pre každú r -ticu a odhadnutím kvantilov ich výberovými ekvivalentmi) alebo numericky, keďže pre nenulový kvantil \tilde{q}_α platí

$$\alpha = E_{Q_r} \left(1 - F_{\sum_{i=1}^{r-1} Q_i} \left(\sum_{i=1}^{r-1} \frac{u_i}{\lambda_i \tilde{q}_\alpha + u_r/Q_r} \right) \right)$$

kde $F_{\sum_{i=1}^{r-1} Q_i}$ označuje distribučnú funkciu $\sum_{i=1}^{r-1} Q_i$.

V každom prípade však výpočet týchto kvantilov nie je práve jednoduchý. Preto by sme sa mohli pokúsiť nájsť pre ne nejakú výpočtovo menej náročnú aproximáciu. Pri jej hľadaní sa inšpirujeme prácou Brossa [3], ktorý túto úlohu riešil pre prípad $r = 2$.

Podobne ako Bross, využijeme fakt, že v špeciálnych prípadoch poznáme tvar \tilde{q}_α , t.j. α kvantilu \tilde{R} . Označme $s_i = u_i/\nu_i$, $i = 1, \dots, r$, potom (6) môžeme prepísať ako

$$\sum_{i=1}^{r-1} \frac{\nu_i s_i}{\lambda_i \tilde{R} + \nu_r s_r/Q_r} = \sum_{i=1}^{r-1} Q_i. \quad (7)$$

Keďže

$$P(\tilde{R} = 0) = P \left(\sum_{i=1}^{r-1} \nu_i s_i / s_r \leq \nu_r \sum_{i=1}^{r-1} Q_i / Q_r \right),$$

je zrejmé, že \tilde{q}_α je rovný nule, keď

$$\sum_{i=1}^{r-1} \nu_i s_i / s_r \leq s F_{s, \nu_r; 1-\alpha}. \quad (8)$$

Ďalej sa dá ukázať, že keď $\nu_r \rightarrow \infty$ a s_1, \dots, s_r považujeme za pevné, tak \tilde{R} konverguje v distribúcii k náhodnej premennej T , ktorú dostaneme ako nezáporné riešenie alebo nulu, ak nezáporné riešenie neexistuje, rovnice

$$\sum_{i=1}^{r-1} \frac{\nu_i s_i}{\lambda_i T + s_r} = \sum_{i=1}^{r-1} Q_i. \quad (9)$$

Tento fakt vyplýva zo spojitosti \tilde{R} ako implicitne zadanej funkcie a toho, že $s_r \nu_r / Q_r \xrightarrow{\mathcal{D}} s_r$, keď $\nu_r \rightarrow \infty$. Označme α kvantil T ako $q_{T;\alpha}$. Potom $q_{T;\alpha}$ je kladné riešenie alebo nula, ak kladné riešenie neexistuje, rovnice

$$\sum_{i=1}^{r-1} \frac{\nu_i s_i}{\lambda_i q_{T;\alpha} + s_r} = \chi_{s;1-\alpha}^2, \quad (10)$$

keďže pre nenulové $q_{T;\alpha}$ platí

$$\alpha = P(T \leq q_{T;\alpha}) = P\left(\sum_{i=1}^{r-1} \frac{\nu_i s_i}{\lambda_i q_{T;\alpha} + s_r} \leq \sum_{i=1}^{r-1} Q_i\right).$$

Spojitosť kvantilovej funkcie náhodnej premennej T spolu s faktom $\tilde{R} \xrightarrow{\mathcal{D}} T$ implikujú $\tilde{q}_\alpha \rightarrow q_{T;\alpha}$ pre $\nu_r \rightarrow \infty$.

Nakoniec uvažujme situáciu, keď $s_i/s_r \rightarrow \infty$, $i = 1, \dots, r-1$, tak, že $s_i/\sum_{j=1}^{r-1} s_j \rightarrow a_i > 0$, $i = 1, \dots, r-1$, pre nejaké kladné konštanty a_i . Potom sa dá ukázať, že pre $\tilde{q}_\alpha/\sum_{i=1}^{r-1} s_i$, čo je α kvantil $\tilde{R}/\sum_{i=1}^{r-1} s_i$, platí

$$\frac{\tilde{q}_\alpha}{\sum_{i=1}^{r-1} s_i} \rightarrow \frac{\sum_{i=1}^{r-1} a_i \nu_i / \lambda_i}{\chi_{s;1-\alpha}^2}, \quad (11)$$

keďže v danej situácii $\tilde{R}/\sum_{i=1}^{r-1} s_i \xrightarrow{\mathcal{D}} (\sum_{i=1}^{r-1} a_i \nu_i / \lambda_i) / \sum_{i=1}^{r-1} Q_i$ a kvantilová funkcia limitnej náhodnej premennej je spojitá. V súvislosti s touto konvergenciou si všimnime, že (7) sa dá prepísať ako

$$\sum_{i=1}^{r-1} \frac{\nu_i s_i / \sum_{i=1}^{r-1} s_i}{\lambda_i \tilde{R} / \sum_{i=1}^{r-1} s_i + \nu_r s_r / (Q_r \sum_{i=1}^{r-1} s_i)} = \sum_{i=1}^{r-1} Q_i.$$

Ak to zhrnieme, pre aproximáciu α kvantilu \tilde{R} , t.j. funkciu B_α , ktorá závisí na s_1, \dots, s_r a nejakých tabelovaných kvantiloch, by malo platiť:

1. $B_\alpha = 0$ práve vtedy, keď $\sum_{i=1}^{r-1} \nu_i s_i / s_r \leq s F_{s, \nu_r; 1-\alpha}$;
2. ak $\nu_r \rightarrow \infty$, tak $B_\alpha \rightarrow t$, kde t je nezáporné riešenie alebo nula, ak nezáporné riešenie neexistuje, rovnice

$$\sum_{i=1}^{r-1} \frac{\nu_i s_i}{\lambda_i t + s_r} = \chi_{s;1-\alpha}^2;$$

3. ak $s_i/s_r \rightarrow \infty$, $i = 1, \dots, r-1$, tak, že $s_i/\sum_{i=1}^{r-1} s_i \rightarrow a_i > 0$, $i = 1, \dots, r-1$, tak

$$\frac{B_\alpha}{\sum_{i=1}^{r-1} s_i} \rightarrow \frac{\sum_{i=1}^{r-1} a_i \nu_i / \lambda_i}{\chi_{s;1-\alpha}^2}.$$

Je zrejmé, že aproximácia spĺňajúca uvedené tri podmienky bude nulová práve vtedy, keď aj pôvodný kvantil bude nula, a tiež napríklad, že s narastajúcim ν_r môžeme očakávať zlepšovanie jej presnosti.

Naším pôvodným cieľom bolo nájsť približné vyjadrenie pre hranice PB zovšeobecného intervalu, t.j. pre $q_{\alpha_l}^{R^{PB}}(\mathbf{u})$, $q_{\alpha_u}^{R^{PB}}(\mathbf{u})$, čo sú pre konkrétnu realizáciu \mathbf{u} vlastne kvantily \tilde{q}_{α_l} , \tilde{q}_{α_u} náhodnej premennej \tilde{R} z (6) s vhodnou voľbou u_i , $i = 1, \dots, r$. Od približných vyjadrení, B_{α_l} , B_{α_u} , uvedených kvantilov teda môžeme vyžadovať splnenie vlastností 1-3 vyššie a je ľahké dokázať, že tieto budú splnené, ak B_{α_l} , B_{α_u} zadefinujeme ako nezáporné riešenia alebo nuly, ak nezáporné riešenie neexistujú, rovníc

$$\sum_{i=1}^{r-1} \frac{\nu_i s_i}{\lambda_i B_{\alpha_l} \frac{\chi_{s;1-\alpha_l}^2}{s F_{s,\nu_r;1-\alpha_l}} + s_r} = s F_{s,\nu_r;1-\alpha_l}$$

$$\sum_{i=1}^{r-1} \frac{\nu_i s_i}{\lambda_i B_{\alpha_u} \frac{\chi_{s;1-\alpha_u}^2}{s F_{s,\nu_r;1-\alpha_u}} + s_r} = s F_{s,\nu_r;1-\alpha_u}.$$

Porovnaním posledných rovníc s (2), (3) hneď vidíme, že pre konkrétnu realizáciu \mathbf{u} môžeme hranice PB zovšeobecného intervalu pri použití hladín $\hat{\alpha}_l$, $\hat{\alpha}_u$ v procedúre aproximovať realizáciou dolnej a hornej hranice EBWT intervalu pri daných pozorovaniach, ak v EBWT procedúre zvolíme $\alpha_l = 1 - \hat{\alpha}_u$, $\alpha_u = 1 - \hat{\alpha}_l$.

Vráťme sa ešte k Brossovemu výsledku. Bross [3] uvažoval prípad vyváženého modelu jednoduchého triedenia s náhodným efektom a fiduciálny interval pre σ_1^2 , ktorý sa zhoduje so zovšeobecným intervalom, na ktorý sa v tejto situácii redukuje PB zovšeobecný interval. Bross si všimol, že hľadané kvantily sa dajú ľahšie počítať v niektorých špeciálnych situáciách, čím dostal podmienky, ktoré by mala aproximácia týchto kvantilov spĺňať a z jej rôznych možných podôb si vybral jednu. Nanešťastie, Brossova voľba mala jeden drobný nedostatok, ktorý si všimol Tukey [10], str. 67, navrhol ako ho odstrániť (pri zachovaní splnenia všetkých Brossových podmienok) a dostal tak to, čo dnes poznáme pod názvom Williamsov-Tukeyho interval. Z toho je zrejmý náš výsledok v špeciálnom prípade $r = 2$.

5 Záver

Vzhľadom na Brossove [3] a Tukeyho [10] výsledky, ako aj podobnosť vzorcov (2), (3) a (5), nie je zistenie, že približný EBWT interval možno považovať za aproximáciu PB zovšeobecného intervalu, prekvapujúce. V skúmanom probléme však odhaľuje prepojenie medzi pomerne ľahko konštruovateľnými zovšeobecnými intervalmi a približnými metódami, ktoré sú zas ľahšie spočítateľné. Náš výsledok bol inšpirovaný prácou Brossa [3], ktorý sa zaoberal špeciálne prípadom vyváženého modelu jednoduchého triedenia s náhodným efektom a skúmal zovšeobecný interval, na ktorý sa v tomto prípade

zjednoduší nami skúmaná Parkova-Burdickova procedúra. Kým Bross priamo využil jednoduchosť konštrukcie zovšeobecného konfidenčného intervalu pre σ_1^2 , keď hľadaním aproximácie jeho hraníc navrhol približný konfidenčný interval pre parameter záujmu, v našom prípade šlo o skúmanie súvislostí medzi približnými a zovšeobecnými metódami na konštrukciu konfidenčných intervalov pre σ_1^2 , ktoré sa objavujú v literatúre.

Literatúra

- [1] Arendacká B. (2007). *A modification of the Hartung-Knapp interval on the variance component in two-variance component models*. Kybernetika, **43**, 471–480.
- [2] Arendacká B. (2008). *Approximate and generalized approaches to confidence intervals on a variance component*. Tatra Mount. Math. Publ., **39**, 27–35.
- [3] Bross I. (1950). *Fiducial intervals for variance components*. Biometrics, **6**, 136–144.
- [4] El-Bassiouni M.Y. (1994). *Short confidence intervals for variance components*. Commun. in Stat. Theory and Methods, **23**, 1915–1933.
- [5] Fisher R.A. (1935). *The fiducial argument in statistical inference*. Annals of Eugenics, **6**, 391–398.
- [6] Hannig J., Iyer H., Patterson P. (2006). *Fiducial generalized confidence intervals*. JASA, **101**, 254–269.
- [7] Olsen A., Seely J., Birkes D. (1976). *Invariant quadratic unbiased estimation for two variance components*. The Annals of Statistics, **4**, 878–890.
- [8] Park D.J., Burdick R.K. (2003). *Performance of confidence intervals in regression models with unbalanced one-fold nested error structures*. Commun. in Stat. Simulation and Computation, **32**, 717–732.
- [9] Thomas J.D., Hultquist R.A. (1978). *Interval estimation for the unbalanced case of the one-way random effects model*. The Annals of Statistics, **6**, 582–587.
- [10] Tukey J.W. (1951). *Components in regression*. Biometrics, **7**, 33–69.
- [11] Wald A. (1947). *A note on regression analysis*. The Annals of Math. Stat., **18**, 586–589.
- [12] Williams J.S. (1962). *A confidence interval for variance components*. Biometrika, **49**, 278–281.

Podakovanie: Práca bola podporená grantmi VEGA 1/3016/06 a APVV RPEU-0008-06.

Adresa: B. Arendacká, Ústav merania SAV, Dúbravská cesta 9, 841 04 Bratislava

E-mail: barendacka@gmail.com

OBLASTI NECITLIVOSTI PRO VARIANČNÍ KOMPONENTY VE SMÍŠENÉM LINEÁRNÍM REGRESNÍM MODELU

Hana Boháčová

Klíčová slova: Smíšený lineární regresní model, varianční komponenty, oblast necitlivosti.

Abstrakt: Skutečné hodnoty parametru varianční matice θ ve smíšeném lineárním modelu obvykle bývají neznámé. Při odhadování parametrů střední hodnoty β pak musíme vycházet z odhadnutých hodnot parametrů varianční matice. V příspěvku se používá θ_0 -MINQUE odhad parametrů varianční matice, kde θ_0 je daná počáteční hodnota. Je zde popsán postup odvození oblasti necitlivosti pro lineární funkci parametrů varianční matice, uveden tvar této množiny a je nastíněno, jaké by mohla mít vlastnosti pro speciální případ, kdy parametr θ má dvě složky.

Abstract: The real values of the variance components θ are usually not known in case of the mixed linear model. Thus we need to use their estimates when estimating the fixed effects parameters β . The θ_0 -MINQUE estimator is used in this paper and an insensitivity region for a linear function of θ is described and some properties of such a set are outlined for a special case of two components of the parameter θ .

1 Seznam použitých symbolů

$r(\mathbf{A})$	hodnota matice \mathbf{A}
$\text{Tr}(\mathbf{A})$	stopa matice \mathbf{A} , definuje se pro čtvercové matice jako součet prvků na hlavní diagonále matice
\mathbf{e}_i	i -tý jednotkový vektor; vektor, který má na i -té pozici jedničku, na ostatních pozicích nuly
\mathbf{I}	jednotková matice
\mathbf{A}^-	pseudoinverzní matice k matici \mathbf{A}
\mathbf{A}^+	Mooreova-Penroseova pseudoinverze matice \mathbf{A}
$\mathcal{M}(\mathbf{A})$	sloupcový prostor matice \mathbf{A}
$\mathbf{M}_{\mathbf{A}}$	projekční matice (v euklidovské normě) na ortogonální komplement prostoru $\mathcal{M}(\mathbf{A})$, tzn. $\mathbf{M}_{\mathbf{A}} = \mathbf{I} - \mathbf{A}\mathbf{A}^+$
\mathbf{Y}	n -rozměrný náhodný vektor
$\mathbf{Y} \sim N_n(\mathbf{X}\beta, \Sigma_{\theta})$	náhodný vektor \mathbf{Y} má n -rozměrné normální rozdělení se střední hodnotou $\mathbf{X}\beta$ a varianční maticí Σ_{θ}
$\text{Var}_{\theta_0}(\mathbf{Y})$	varianční matice náhodného vektoru \mathbf{Y} při skutečné hodnotě parametru varianční matice rovné θ_0

$\text{cov}_{\boldsymbol{\theta}_0}(X, Z)$ kovariance veličin X a Z za předpokladu, že skutečná hodnota parametru varianční matice je $\boldsymbol{\theta}_0$

$\left. \frac{\partial f(\boldsymbol{\theta})}{\partial \theta_i} \right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}$ hodnota parciální derivace funkce $f(\boldsymbol{\theta})$ podle θ_i v bodě $\boldsymbol{\theta}_0$

2 Smíšený lineární model

Uvažujme následující lineární regresní model (viz [5], str. 229):

$$\mathbf{Y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}). \quad (1)$$

Nechť \mathbf{X} rozměru $n \times k$ je známá matice plné sloupcové hodnosti $r(\mathbf{X}) = k$, $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_k)'$ je vektor neznámých parametrů střední hodnoty a nechť varianční matice $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ splňuje

$$\boldsymbol{\Sigma}_{\boldsymbol{\theta}} = \sum_{i=1}^r \theta_i \mathbf{V}_i, \quad (2)$$

kde $\theta_1, \theta_2, \dots, \theta_r > 0$ jsou neznámé varianční komponenty - objekt našeho zájmu - a $\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_r$ jsou známé lineárně nezávislé symetrické pozitivně semidefinitní matice, přičemž uvažujeme pouze takové hodnoty variančních komponent, pro které platí:

$$\boldsymbol{\theta} = (\theta_1, \dots, \theta_r)' \in \mathcal{S}, \quad (3)$$

kde \mathcal{S} je množina takových $\boldsymbol{\theta} \in \mathbb{R}^r$, pro která je (při daných $\mathbf{V}_1, \dots, \mathbf{V}_r$) matice $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ pozitivně definitní.

3 Odhady parametrů smíšeného lineárního modelu

Cílem je odhadnout parametry střední hodnoty $\boldsymbol{\beta} = (\beta_1, \dots, \beta_k)'$ a parametry varianční matice - varianční komponenty - $\boldsymbol{\theta} = (\theta_1, \dots, \theta_r)$. Použijeme-li $\boldsymbol{\theta}_0 \in \mathbb{R}^r$ - počáteční hodnotu parametru $\boldsymbol{\theta}$, pak $\boldsymbol{\theta}_0$ -lokálně nejlepším lineárním nestranným odhadem parametru $\boldsymbol{\beta}$ je (viz např. [4, str. 13])

$$\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}_0) = (\mathbf{X}'\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}^{-1}\mathbf{X})^{-1} \mathbf{X}'\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}^{-1}\mathbf{Y}. \quad (4)$$

$\boldsymbol{\theta}_0$ -MINQUE (nestranný kvadratický odhad s minimální normou při $\boldsymbol{\theta} = \boldsymbol{\theta}_0$, viz např. [3], str.97-101) vektoru $\boldsymbol{\theta}$ je

$$\hat{\boldsymbol{\theta}}(\boldsymbol{\theta}_0) = \mathbf{S}_{(\mathbf{M}_{\mathbf{X}}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_{\mathbf{X}})^+}^{-1} \begin{pmatrix} \mathbf{Y}'(\mathbf{M}_{\mathbf{X}}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_{\mathbf{X}})^+ \mathbf{V}_1 (\mathbf{M}_{\mathbf{X}}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_{\mathbf{X}})^+ \mathbf{Y} \\ \vdots \\ \mathbf{Y}'(\mathbf{M}_{\mathbf{X}}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_{\mathbf{X}})^+ \mathbf{V}_r (\mathbf{M}_{\mathbf{X}}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_{\mathbf{X}})^+ \mathbf{Y} \end{pmatrix}, \quad (5)$$

za předpokladu, že matice $\mathbf{S}_{(\mathbf{M}_X \boldsymbol{\Sigma}_{\theta_0} \mathbf{M}_X)^+}$ je regulární. Matice $\mathbf{S}_{(\mathbf{M}_X \boldsymbol{\Sigma}_{\theta_0} \mathbf{M}_X)^+}$ má na pozici (i, j) prvek

$$\left\{ \mathbf{S}_{(\mathbf{M}_X \boldsymbol{\Sigma}_{\theta_0} \mathbf{M}_X)^+} \right\}_{i,j} = \text{Tr} \left[\mathbf{V}_i (\mathbf{M}_X \boldsymbol{\Sigma}_{\theta_0} \mathbf{M}_X)^+ \mathbf{V}_j (\mathbf{M}_X \boldsymbol{\Sigma}_{\theta_0} \mathbf{M}_X)^+ \right].$$

Pro varianční matice odhadů (4) a (5) platí

$$\text{Var}_{\boldsymbol{\theta}_0} \left[\widehat{\boldsymbol{\beta}}(\boldsymbol{\theta}_0) \right] = (\mathbf{X}' \boldsymbol{\Sigma}_{\theta_0}^{-1} \mathbf{X})^{-1}, \quad (6)$$

$$\text{Var}_{\boldsymbol{\theta}_0} \left[\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}_0) \right] = 2 \mathbf{S}_{(\mathbf{M}_X \boldsymbol{\Sigma}_{\theta_0} \mathbf{M}_X)^+}^{-1}. \quad (7)$$

4 Oblast necitlivosti pro lineární funkci parametru $\boldsymbol{\theta}$ ve smíšeném lineárním modelu

Při odvozování oblasti necitlivosti pro lineární funkci vektoru variančních komponent budeme vycházet z $\boldsymbol{\theta}_0$ -MINQUE vektoru $\boldsymbol{\theta}$ (5) a příslušné varianční matice (7). Nechť je dán vektor $\mathbf{g} \in \mathbb{R}^r$. Uvažujme v dalším $\mathbf{g}'\boldsymbol{\theta}$ - lineární kombinaci složek vektoru $\boldsymbol{\theta}$. Její $\boldsymbol{\theta}_0$ -MINQUE je

$$\widehat{\mathbf{g}'\boldsymbol{\theta}}(\boldsymbol{\theta}_0) = \mathbf{g}'\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}_0) \quad (8)$$

a podle (7) a (8) platí

$$\text{Var}_{\boldsymbol{\theta}_0} \left[\widehat{\mathbf{g}'\boldsymbol{\theta}}(\boldsymbol{\theta}_0) \right] = 2 \mathbf{g}' \mathbf{S}_{(\mathbf{M}_X \boldsymbol{\Sigma}_{\theta_0} \mathbf{M}_X)^+}^{-1} \mathbf{g}. \quad (9)$$

Budeme se zabývat otázkou, o jaké $\delta\boldsymbol{\theta} = (\delta\theta_1, \dots, \delta\theta_r)$ můžeme změnit počáteční hodnotu $\boldsymbol{\theta}_0$, aby směrodatná odchylka odhadu $\widehat{\mathbf{g}'\boldsymbol{\theta}}(\boldsymbol{\theta}_0 + \delta\boldsymbol{\theta})$ nevrstla ve srovnání se směrodatnou odchylkou odhadu $\widehat{\mathbf{g}'\boldsymbol{\theta}}(\boldsymbol{\theta}_0)$ o víc než o ε -násobek. Budeme tedy hledat množinu $\mathcal{N}_{\mathbf{g}'\boldsymbol{\theta}, \boldsymbol{\theta}_0}$, pro kterou platí, že je-li $(\boldsymbol{\theta}_0 + \delta\boldsymbol{\theta}) \in \mathcal{N}_{\mathbf{g}'\boldsymbol{\theta}, \boldsymbol{\theta}_0}$, pak

$$\sqrt{\text{Var}_{\boldsymbol{\theta}_0} \left[\widehat{\mathbf{g}'\boldsymbol{\theta}}(\boldsymbol{\theta}_0 + \delta\boldsymbol{\theta}) \right]} \leq (1 + \varepsilon) \sqrt{\text{Var}_{\boldsymbol{\theta}_0} \left[\widehat{\mathbf{g}'\boldsymbol{\theta}}(\boldsymbol{\theta}_0) \right]} \quad (10)$$

pro dané $\varepsilon > 0$ dostatečně malé. Tuto množinu budeme nazývat oblastí necitlivosti pro lineární funkci $\mathbf{g}'\boldsymbol{\theta}$ parametrů varianční matice.

Může vzniknout otázka, čím je oblast necitlivosti pro statistika tak zajímavá, když varianční komponenty určíme iterací a tedy samotný odhad na počáteční hodnotě $\boldsymbol{\theta}_0$ nezávisí. Motivace k určení uvedené oblasti necitlivosti je následující. Odhad $\widehat{\boldsymbol{\theta}}$ vektoru variančních komponent je charakterizován určitou mírou nejistoty, která v uvažovaném modelu je různá pro různé hodnoty vektoru $\boldsymbol{\theta}$. Jestliže v daném bodě $\boldsymbol{\theta}$ je v oblasti necitlivosti obsažena oblast spolehlivosti s dostatečně vysokou pravděpodobností, je

zaručeno (s praktickou jistotou), že použitý odhad, v němž v roli θ_0 vystupuje hodnota odhadu $\hat{\theta}$ z poslední iterace, má skutečnou disperzi jen nevýznamně odchýlenou od hodnoty získané numericky. V případě, že oblast necitlivosti je příliš malá vzhledem k oblasti spolehlivosti, pak skutečná dosperze se může od vypočítané odchylovat významně.

Poznámka 4.1. Budeme-li volit za vektor \mathbf{g} určující lineární kombinaci složek vektoru θ i -tý jednotkový vektor $\mathbf{e}_i \in \mathbb{R}^r$, udává pro pevně zvolenou počáteční hodnotu θ_0 oblast necitlivosti $\mathcal{N}_{\mathbf{g}'\theta, \theta_0}$ množinu všech přípustných počátečních hodnot variančních komponent, které nevedou ke zvýšení směrodatné odchylky odhadu i -té varianční komponenty θ_i o více než o ε -násobek ve srovnání se směrodatnou odchylkou odhadu θ_i založeného na počáteční hodnotě θ_0 . Vezmeme-li takto postupně $\mathbf{g} = \mathbf{e}_i$ pro $i = 1, 2, \dots, r$ a určíme-li průnik příslušných oblastí necitlivosti

$$\bigcap_{i=1}^r \mathcal{N}_{\mathbf{e}_i'\theta, \theta_0},$$

dostaneme množinu vstupních hodnot variančních komponent, jejichž použití nepůsobí větší než 100 ε -procentní zvětšení směrodatné odchylky žádné varianční komponenty v porovnání s odhadem založenými na počáteční hodnotě θ_0 .

$\hat{\theta}(\theta_0 + \delta\theta)$ můžeme pomocí Taylorova polynomu prvního stupně aproximovat takto

$$\hat{\theta}(\theta_0 + \delta\theta) \approx \hat{\theta}(\theta_0) + \sum_{i=1}^r \left[\frac{\partial \hat{\theta}(\theta)}{\partial \theta_i} \bigg|_{\theta=\theta_0} \delta\theta_i \right]. \quad (11)$$

Pro odpovídající lineární funkci $\mathbf{g}'\hat{\theta}(\theta_0 + \delta\theta)$ tedy platí

$$\mathbf{g}'\hat{\theta}(\theta_0 + \delta\theta) \approx \mathbf{g}'\hat{\theta}(\theta_0) + \sum_{i=1}^r \left[\frac{\partial \mathbf{g}'\hat{\theta}(\theta)}{\partial \theta_i} \bigg|_{\theta=\theta_0} \delta\theta_i \right]. \quad (12)$$

Označme $\mathbf{C}_{\mathbf{A}, \mathbf{B}}$ matici, která má na pozici (i, j) prvek $\text{Tr}(\mathbf{V}_i \mathbf{A} \mathbf{V}_j \mathbf{B})$, $S_{\mathbf{A}}$ matici s $\text{Tr}(\mathbf{V}_i \mathbf{A} \mathbf{V}_j \mathbf{A})$ na pozici (i, j) a

$$\gamma = \begin{pmatrix} \mathbf{Y}'(\mathbf{M}_{\mathbf{X}} \Sigma_{\theta} \mathbf{M}_{\mathbf{X}})^+ \mathbf{V}_1 (\mathbf{M}_{\mathbf{X}} \Sigma_{\theta} \mathbf{M}_{\mathbf{X}})^+ \mathbf{Y} \\ \vdots \\ \mathbf{Y}'(\mathbf{M}_{\mathbf{X}} \Sigma_{\theta} \mathbf{M}_{\mathbf{X}})^+ \mathbf{V}_r (\mathbf{M}_{\mathbf{X}} \Sigma_{\theta} \mathbf{M}_{\mathbf{X}})^+ \mathbf{Y} \end{pmatrix}.$$

Pak

$$\begin{aligned} \frac{\partial \hat{\theta}(\theta)}{\partial \theta_k} \bigg|_{\theta=\theta_0} &= -2\mathbf{S}_{(\mathbf{M}_{\mathbf{X}} \Sigma_{\theta_0} \mathbf{M}_{\mathbf{X}})^+}^{-1} \\ &\times \mathbf{C}_{(\mathbf{M}_{\mathbf{X}} \Sigma_{\theta_0} \mathbf{M}_{\mathbf{X}})^+ \mathbf{V}_k (\mathbf{M}_{\mathbf{X}} \Sigma_{\theta_0} \mathbf{M}_{\mathbf{X}})^+, (\mathbf{M}_{\mathbf{X}} \Sigma_{\theta_0} \mathbf{M}_{\mathbf{X}})^+} \mathbf{S}_{(\mathbf{M}_{\mathbf{X}} \Sigma_{\theta_0} \mathbf{M}_{\mathbf{X}})^+}^{-1} \gamma \end{aligned}$$

$$\begin{aligned}
 & + \mathbf{S}_{(\mathbf{M}_X \boldsymbol{\Sigma}_{\theta_0} \mathbf{M}_X)^+}^{-1} \\
 & \left(\begin{array}{c} 2\mathbf{Y}'(\mathbf{M}_X \boldsymbol{\Sigma}_{\theta_0} \mathbf{M}_X)^+ \mathbf{V}_k(\mathbf{M}_X \boldsymbol{\Sigma}_{\theta_0} \mathbf{M}_X)^+ \mathbf{V}_1(\mathbf{M}_X \boldsymbol{\Sigma}_{\theta_0} \mathbf{M}_X)^+ \mathbf{Y} \\ \vdots \\ 2\mathbf{Y}'(\mathbf{M}_X \boldsymbol{\Sigma}_{\theta_0} \mathbf{M}_X)^+ \mathbf{V}_k(\mathbf{M}_X \boldsymbol{\Sigma}_{\theta_0} \mathbf{M}_X)^+ \mathbf{V}_r(\mathbf{M}_X \boldsymbol{\Sigma}_{\theta_0} \mathbf{M}_X)^+ \mathbf{Y} \end{array} \right). \quad (13)
 \end{aligned}$$

K odvození $\text{Var}_{\boldsymbol{\theta}_0} [\mathbf{g}'\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}_0 + \delta\boldsymbol{\theta})]$ ve vztahu (10) použijeme vztahů (12) a (13). Toto odvození je poměrně zdlouhavé a kvůli stručnosti zde nebude uvedeno. Je možné je najít např. v [1]. Výsledná oblast necitlivosti pak je množina

$$\begin{aligned}
 \mathcal{N}_{\mathbf{g}'\boldsymbol{\theta}, \boldsymbol{\theta}_0} &= \left\{ \boldsymbol{\theta}_0 + \delta\boldsymbol{\theta} : \delta\boldsymbol{\theta}' \mathbf{W}_{\mathbf{g}} \delta\boldsymbol{\theta} \leq (2\varepsilon + \varepsilon^2) \text{Var} [\mathbf{g}'\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}_0)] \right\} \\
 &= \left\{ \boldsymbol{\theta}_0 + \delta\boldsymbol{\theta} : \delta\boldsymbol{\theta}' \mathbf{W}_{\mathbf{g}} \delta\boldsymbol{\theta} \leq 2(2\varepsilon + \varepsilon^2) \mathbf{g}' \mathbf{S}_{(\mathbf{M}_X \boldsymbol{\Sigma}_{\theta_0} \mathbf{M}_X)^+}^{-1} \mathbf{g} \right\}, \quad (14)
 \end{aligned}$$

kde

$$\begin{aligned}
 \mathbf{W}_{\mathbf{g}} &= \text{Var}_{\boldsymbol{\theta}_0} \left[\frac{\partial \mathbf{g}'\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right] \\
 &= 8 \begin{pmatrix} \mathbf{a}'_1 \\ \vdots \\ \mathbf{a}'_r \end{pmatrix} \mathbf{S}_{(\mathbf{M}_X \boldsymbol{\Sigma}_{\theta_0} \mathbf{M}_X)^+} (\mathbf{a}_1, \dots, \mathbf{a}_r) + 8 \mathbf{S}_{(\mathbf{M}_X \boldsymbol{\Sigma}_{\theta_0} \mathbf{M}_X)^+} \mathbf{V}_{\mathbf{g}} (\mathbf{M}_X \boldsymbol{\Sigma}_{\theta_0} \mathbf{M}_X)^+ \\
 &\quad - 8 \begin{pmatrix} \mathbf{a}'_1 \\ \vdots \\ \mathbf{a}'_r \end{pmatrix} \mathbf{C}_{(\mathbf{M}_X \boldsymbol{\Sigma}_{\theta_0} \mathbf{M}_X)^+, (\mathbf{M}_X \boldsymbol{\Sigma}_{\theta_0} \mathbf{M}_X)^+} \mathbf{V}_{\mathbf{g}} (\mathbf{M}_X \boldsymbol{\Sigma}_{\theta_0} \mathbf{M}_X)^+ \\
 &\quad - 8 \mathbf{C}'_{(\mathbf{M}_X \boldsymbol{\Sigma}_{\theta_0} \mathbf{M}_X)^+, (\mathbf{M}_X \boldsymbol{\Sigma}_{\theta_0} \mathbf{M}_X)^+} \mathbf{V}_{\mathbf{g}} (\mathbf{M}_X \boldsymbol{\Sigma}_{\theta_0} \mathbf{M}_X)^+ (\mathbf{a}_1, \dots, \mathbf{a}_r). \quad (15)
 \end{aligned}$$

V zápisu matice $\mathbf{W}_{\mathbf{g}}$ bylo použito následující značení:

$$\begin{aligned}
 \mathbf{a}'_i &= \mathbf{g}' \mathbf{S}_{(\mathbf{M}_X \boldsymbol{\Sigma}_{\theta_0} \mathbf{M}_X)^+}^{-1} \mathbf{C}_{(\mathbf{M}_X \boldsymbol{\Sigma}_{\theta_0} \mathbf{M}_X)^+} \mathbf{v}_i (\mathbf{M}_X \boldsymbol{\Sigma}_{\theta_0} \mathbf{M}_X)^+, (\mathbf{M}_X \boldsymbol{\Sigma}_{\theta_0} \mathbf{M}_X)^+ \\
 &\quad \cdot \mathbf{S}_{(\mathbf{M}_X \boldsymbol{\Sigma}_{\theta_0} \mathbf{M}_X)^+}^{-1}, \quad i = 1, \dots, r
 \end{aligned}$$

$$\mathbf{V}_{\mathbf{g}} = \sum_{i=1}^r \mathbf{b}_{\mathbf{g}_i} \mathbf{V}_i,$$

(kde $\mathbf{b}_{\mathbf{g}_i}$ je i -tá složka vektoru $\mathbf{b}_{\mathbf{g}} = \mathbf{S}_{(\mathbf{M}_X \boldsymbol{\Sigma}_{\theta_0} \mathbf{M}_X)^+}^{-1} \mathbf{g}$).

5 Vlastnosti oblasti necitlivosti pro lineární funkci parametrů varianční matice

Podívejme se nejprve na okamžik na oblast necitlivosti pro lineární funkci $\mathbf{h}'\boldsymbol{\beta}$ parametru střední hodnoty určenou vektorem $\mathbf{h} \in \mathbb{R}^k$. Podobným postupem, jaký byl popsán v předešlé části, lze pro dané $\varepsilon > 0$, daný vektor \mathbf{h} a daný počáteční bod $\boldsymbol{\theta}_0$ získat oblast necitlivosti v tomto tvaru (podrobnější odvození je uvedeno v článku [2], který vychází z šesté kapitoly [4])

$$\mathcal{N}_{\mathbf{h}'\boldsymbol{\beta}, \boldsymbol{\theta}_0} = \{\boldsymbol{\theta}_0 + \delta\boldsymbol{\theta} : \delta\boldsymbol{\theta}'\mathbf{W}_{\mathbf{h}}\delta\boldsymbol{\theta} \leq 2\varepsilon \text{Var}_{\boldsymbol{\theta}_0}[\mathbf{h}'\widehat{\boldsymbol{\beta}}(\boldsymbol{\theta}_0)]\}. \quad (16)$$

Matice $\mathbf{W}_{\mathbf{h}}$ je rozměru $r \times r$ a má na pozici (i, j) prvek

$$\begin{aligned} & \{\mathbf{W}_{\mathbf{h}}\}_{i,j} \\ &= \text{cov}_{\boldsymbol{\theta}_0} \left\{ \mathbf{L}_{\mathbf{h}}' \mathbf{V}_i \boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}^{-1} [\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}}(\boldsymbol{\theta}_0)], \mathbf{L}_{\mathbf{h}}' \mathbf{V}_j \boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}^{-1} [\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}}(\boldsymbol{\theta}_0)] \right\} \\ &= \mathbf{L}_{\mathbf{h}}' \mathbf{V}_i (\mathbf{M}_{\mathbf{X}} \boldsymbol{\Sigma}_{\boldsymbol{\theta}_0} \mathbf{M}_{\mathbf{X}})^+ \mathbf{V}_j \mathbf{L}_{\mathbf{h}}, \end{aligned}$$

kde

$$\mathbf{L}_{\mathbf{h}}' = \mathbf{h}' (\mathbf{X}' \boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}^{-1} \mathbf{X})^{-1} \mathbf{X}' \boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}^{-1}.$$

Lze ukázat, že platí:

Lemma 5.1. (viz [3])

Pro matici $\mathbf{W}_{\mathbf{h}} = \begin{pmatrix} \mathbf{L}_{\mathbf{h}}' \mathbf{V}_1 \\ \vdots \\ \mathbf{L}_{\mathbf{h}}' \mathbf{V}_r \end{pmatrix} [\mathbf{M}_{\mathbf{X}} \boldsymbol{\Sigma}_{\boldsymbol{\theta}_0} \mathbf{M}_{\mathbf{X}}]^+ (\mathbf{V}_1 \mathbf{L}_{\mathbf{h}}, \dots, \mathbf{V}_r \mathbf{L}_{\mathbf{h}})$, která je přiřazena funkci $h(\boldsymbol{\beta}) = \mathbf{h}'\boldsymbol{\beta}$, $\boldsymbol{\beta} \in \mathbb{R}^k$ platí

$$\mathcal{M}(\mathbf{W}_{\mathbf{h}}) \perp \boldsymbol{\theta}_0. \quad (17)$$

Z lemmatu vyplývá, že matice $\mathbf{W}_{\mathbf{h}}$ je singulární. Podle (17) je vektor $\boldsymbol{\theta}_0$ kolmý na podprostor prostoru \mathbb{R}^r generovaný sloupci matice $\mathbf{W}_{\mathbf{h}}$. Pak tedy pro speciální případ $r = 2$ můžeme množinu (16) graficky znázornit jako pás vymezený dvojicí přímkou rovnoběžných s orientovanou úsečkou spojující počátek soustavy souřadnic s bodem $\boldsymbol{\theta}_0$. Otázka je, zda podobné tvrzení lze vyslovit i o matici $\mathbf{W}_{\mathbf{g}}$ z množiny (14). Vzhledem k jejímu poměrně komplikovanému tvaru je ovšem problém zabývat se jejími vlastnostmi teoreticky. Otázkou tedy je, jak vypadá prostor $\mathcal{M}(\mathbf{W}_{\mathbf{g}})$ generovaný sloupci této matice a zda i v tomto případě platí, že vektor počátečních hodnot $\boldsymbol{\theta}_0$ je na prostor $\mathcal{M}(\mathbf{W}_{\mathbf{g}})$ kolmý. Ve všech provedených numerických studiích byla matice $\mathbf{W}_{\mathbf{g}}$ singulární a použitý vektor $\boldsymbol{\theta}_0$ byl na prostor $\mathcal{M}(\mathbf{W}_{\mathbf{g}})$ kolmý, pro $r = 2$ byla tedy vždy množina $\mathcal{N}_{\mathbf{g}'\boldsymbol{\theta}, \boldsymbol{\theta}_0}$ množinou všech bodů mezi dvojicí rovnoběžných přímkou se směrovým vektorem $\boldsymbol{\theta}_0$ (včetně těchto hraničních přímkou). Vzhledem ke komplikovanému tvaru matice $\mathbf{W}_{\mathbf{g}}$ se však tyto její vlastnosti nepodařilo dokázat obecně.

Literatura

- [1] Boháčová H. (2008). *Insensitivity regions for variance components in general linear model*. In: Acta Universitatis Palackianae Olomucensis, Facultas Rerum Naturalium, Mathematica, 7–22.
- [2] Boháčová H., Heckenbergerová, J. (2007). *Insensitivity regions for fixed effects parameters in mixed linear model with type I constraints and calculation problems linked with them* (in Czech). In: Forum Statisticum Slovacum 6/2007, 31–35.
- [3] Kubáček L., Kubáčková L. (2000). *Statistika a metrologie*, Univerzita Palackého v Olomouci.
- [4] Kubáček L., Kubáčková L., Volaufová J. (1995). *Statistical models with linear structures*. Veda, 1995.
- [5] Rao C.R., Kleffe J. (1988). *Estimation of variance components and application*. North-Holland, Amsterdam-New York-Oxford-Tokyo.
- [6] Rényi A. (1970). *Probability theory*, Akadémiai Kiadó, Budapest.

Poděkování: Autorka děkuje panu profesorovi L. Kubáčkovi za všestrannou pomoc.

Adresa: H. Boháčová, Ústav matematiky, Fakulta ekonomicko správní, Univerzita Pardubice, Studentská 84, 532 10, Pardubice

E-mail: hana.bohacova@upce.cz

RECOURSE REFORMULATION OF CHANCE CONSTRAINED PROBLEMS

Martin Branda

Keywords: Stochastic programming, chance constrained problems, penalty function problems, asymptotic equivalence.

Abstract: We explore reformulation of stochastic programs with chance constraints by stochastic programs with suitably chosen penalty-type objectives. We show that the two problems are asymptotically equivalent. The obtained problems with penalties and with a fixed set of feasible solutions are much simpler to solve and analyze than the chance constrained programs.

Abstrakt: Článek se zabývá možností přeformulovat úlohu s pravděpodobnostními omezeními pomocí úlohy s penalizací v účelové funkci. Je ukázána asymptotická ekvivalence obou úloh. Získané úlohy s penalizací jsou poté snadněji řešitelné a i jejich struktura umožňuje snažší studium vlastností.

1 Introduction

Stochastic programming treats problems where optimization and uncertainty appears together. In general, we consider the following program with a random factor

$$\min \{g_0(x, \omega) : x \in X, g_i(x, \omega) \leq 0, i = 1, \dots, k\}, \quad (1)$$

where $g_i, i = 0, \dots, k$, are real functions on $\mathbb{R}^n \times \Omega$, $X \subseteq \mathbb{R}^n$ and ω is a realization of a random vector with support $\Omega \subseteq \mathbb{R}^s$. However, ω is unknown for us and we would like to hedge against a few worst realizations of ω .

Expectation is usually incorporated into objective, i.e. $f(x) := \mathbb{E}[g_0(x, \omega)]$ if it exists is minimized over stochastic constraints. A question is how to deal with them. In [6], three suggestions how to deal with stochastic constraints of the special form $g_i(x, \omega) = \omega_i - g_i(x) \leq 0$, $i = 1, \dots, k$, where ω_i are random bounds with distribution functions F_i , are introduced. First, the constraints can be incorporated into the objective function of the optimization problems as penalty function

$$\sum_{i=1}^k q_i \int_{g_i(x)}^{\infty} [\omega_i - g_i(x)] F_i(d\omega).$$

with $q_i > 0$ being constant. Next, reliability type model with chance or probabilistic constraints is considered

$$P(g_i(x) \geq \omega_i, i = 1, \dots, k) \geq 1 - \varepsilon.$$

for some level $\varepsilon \in (0, 1)$. Finally, constraints involving conditional expectation

$$\mathbb{E}[\omega_i - g_i(x) | \omega_i - g_i(x) > 0] \leq l_i, \quad i = 1, \dots, k$$

for some small levels $l_i > 0$.

In this paper, we will study relation between general chance constrained and penalty function problems. We will show that the model with chance constraints and penalty type model are asymptotically equivalent under quite mild assumptions.

The paper is organized as follows. In section 2, we will define the problems and show that they are asymptotically equivalent. In section 3, equivalence for generalized chance constrained problem is proven. In section 4, we will summarize our results and suggest topics for future research.

2 Reformulation

The multiple chance constrained problem can be formulated as follows:

$$\begin{aligned} \psi_\varepsilon &= \min_{x \in X} f(x), \\ \text{s.t.} \quad & P(g_{11}(x, \omega) \leq 0, \dots, g_{1k_1}(x, \omega) \leq 0) \geq 1 - \varepsilon_1, \\ & \vdots \\ & P(g_{m1}(x, \omega) \leq 0, \dots, g_{mk_m}(x, \omega) \leq 0) \geq 1 - \varepsilon_m, \end{aligned} \quad (2)$$

with optimal solution x_ε , where $\varepsilon = (\varepsilon_1, \dots, \varepsilon_m)$, with the levels $\varepsilon_j \in (0, 1)$.

At the time of writing this paper we do not know any solver which could be used for solving previous multiple jointly chance constrained problems. On the other hand, stochastic programs with penalties and fixed set of feasible solution can be solved much simpler, see [5]. Thus, the recourse reformulation which is stated below may be very useful.

In [4], asymptotic equivalence between problem with one joint chance constraint and problem with simple recourse penalty function is shown. The approach by [4] can be extended to a whole class of penalty functions with desirable properties which was done in [3]. We propose further extension to multiple jointly chance constrained problems (2).

Below, we will consider penalty functions $\vartheta_j : \mathbb{R}^{k_j} \rightarrow \mathbb{R}_+$, $j = 1, \dots, m$, which are continuous nondecreasing in their components, equal to 0 on $\mathbb{R}_-^{k_j}$ and positive otherwise. We denote

$$g_j(x, \omega) = (g_{j1}(x, \omega), \dots, g_{jk_j}(x, \omega)) : \mathbb{R}^n \times \Omega \rightarrow \mathbb{R}^{k_j}$$

and set $p_j(x, \omega) = \vartheta_j(g_j(x, \omega))$. Our choice is appropriate, because it holds

$$P(g_{ji}(x, \omega) \leq 0, \quad i = 1, \dots, k_j) \geq 1 - \varepsilon_j \iff P(p_j(x, \omega) > 0) \leq \varepsilon_j.$$

The corresponding penalty function problem can be formulated as follows:

$$\varphi_N = \min_{x \in X} \left[f(x) + N \cdot \sum_{j=1}^m \mathbb{E}[p_j(x, \omega)] \right] \quad (3)$$

with N a positive parameter. We denote x_N an optimal solution of (3).

A rigorous proof of the relationship between optimal values of (2) and those of (3) for a special additive penalty function and one chance constraint was given by [4]. The following main theorem states the asymptotic equivalence of the models in generalized settings.

Theorem 1. Consider the two problems (2) and (3) and assume: $X \neq \emptyset$ compact, $f(x) = \mathbb{E}[g_0(x, \omega)]$ a finite continuous function of x , $\vartheta_j : \mathbb{R}^{k_j} \rightarrow \mathbb{R}_+$, $j = 1, \dots, m$, are continuous functions, nondecreasing in their components, which are equal to 0 on $\mathbb{R}_-^{k_j}$ and positive otherwise, denote $p_j(x, \omega) = \vartheta_j(g_j(x, \omega))$, $j = 1, \dots, m$, and assume

- (i) $g_{ji}(\cdot, \omega)$, $i = 1, \dots, k_j$, $j = 1, \dots, m$, are almost surely continuous;
- (ii) there exists a nonnegative random variable $C(\omega)$ with $\mathbb{E}[C^{1+\kappa}(\omega)] < \infty$ for some $\kappa > 0$, such that $|p_j(x, \omega)| \leq C(\omega)$, $j = 1, \dots, m$, for all $x \in X$;
- (iii) $\mathbb{E}[p_j(x', \omega)] = 0$, $j = 1, \dots, m$, for some $x' \in X$;
- (iv) $P(g_{ji}(x, \omega) = 0) = 0$, $i = 1, \dots, k_j$, $j = 1, \dots, m$, for all $x \in X$.

Denote $\gamma = \kappa/(2(1 + \kappa))$, and for arbitrary $N > 0$ and $\epsilon \in (0, 1)^m$ put

$$\begin{aligned} \varepsilon_j(x) &= P(p_j(x, \omega) > 0), \quad j = 1, \dots, m, \\ \alpha_N(x) &= N \cdot \sum_{j=1}^m \mathbb{E}[p_j(x, \omega)], \\ \beta_\epsilon(x) &= \varepsilon_{max}^{-\gamma} \sum_{j=1}^m \mathbb{E}[p_j(x, \omega)], \end{aligned}$$

where ε_{max} denotes maximum of the vector $\epsilon = (\varepsilon_1, \dots, \varepsilon_m)$ and $[1/N^{1/\gamma}] = (1/N^{1/\gamma}, \dots, 1/N^{1/\gamma})$ is the vector of length m .

THEN for any prescribed $\epsilon \in (0, 1)^m$ there always exists N large enough so that minimization (3) generates optimal solutions x_N which also satisfy the chance constraints (2) with the given ϵ .

Moreover, bounds on the optimal value ψ_ϵ of (2) based on the optimal value φ_N of (3) and vice versa can be constructed:

$$\begin{aligned} \varphi_{1/\varepsilon_{\max}(x_N)} - \beta_{\epsilon(x_N)}(x_{\epsilon(x_N)}) &\leq \psi_{\epsilon(x_N)} \leq \varphi_N - \alpha_N(x_N), \\ \psi_{\epsilon(x_N)} + \alpha_N(x_N) &\leq \varphi_N \leq \psi_{[1/N^{1/\gamma}]} + \beta_{[1/N^{1/\gamma}]}(x_{[1/N^{1/\gamma}]}) \end{aligned} \quad (4)$$

with

$$\lim_{N \rightarrow +\infty} \alpha_N(x_N) = \lim_{N \rightarrow +\infty} \varepsilon_j(x_N) = \lim_{\varepsilon_{max} \rightarrow 0_+} \beta_\epsilon(x_\epsilon) = 0$$

for any sequences of optimal solutions x_N and x_ϵ .

Proof. We denote

$$\delta_N = \sum_{j=1}^m \mathbb{E}[p_j(x_N, \omega)]$$

for some sequence x_N of optimal solutions of the problem (3). Our assumptions and general properties of the penalty function method, see [2, Theorem 9.2.2], ensure that for any sequence x_N of optimal solutions $\delta_N \rightarrow 0_+$ and also $\alpha_N(x_N) = N\delta_N \rightarrow 0$ as $N \rightarrow \infty$. Then by Chebysev inequality

$$\begin{aligned} & P(p_j(x_N, \omega) > 0) = \\ &= P(0 < p_j(x_N, \omega) \leq \sqrt{\delta_N}) + P(p_j(x_N, \omega) > \sqrt{\delta_N}) \\ &\leq G_j(x_N, \sqrt{\delta_N}) - G_j(x_N, 0) + \frac{1}{\sqrt{\delta_N}} \mathbb{E}[p_j(x_N, \omega)] \\ &\leq G_j(x_N, \sqrt{\delta_N}) - G_j(x_N, 0) + \sqrt{\delta_N} \rightarrow 0, \text{ as } N \rightarrow \infty, j = 1, \dots, m. \end{aligned}$$

Here for a fixed x , $G_j(x, \cdot)$ denotes the distribution function of $p_j(x, \omega)$ defined by

$$G_j(x, y) = P(p_j(x, \omega) \leq y), \quad j = 1, \dots, m.$$

Assumption (iii) implies that for every vector $\epsilon > 0$ (with small components) there exists some $x_\epsilon \in X$ such that

$$P(g_{ji}(x_\epsilon, \omega) \leq 0, i = 1, \dots, k_j) \geq 1 - \epsilon_j, \quad j = 1, \dots, m.$$

Then for any $\epsilon > 0$ the following relations hold

$$\begin{aligned} & \sum_{j=1}^m \mathbb{E}[p_j(x_\epsilon, \omega)] = \\ &= \sum_{j=1}^m \int_{\Omega} |p_j(x_\epsilon, \omega)| I_{(p_j(x_\epsilon, \omega) > 0)} P(d\omega) \\ &\leq \sum_{j=1}^m \int_{\Omega} C(\omega) I_{(p_j(x_\epsilon, \omega) > 0)} P(d\omega) \\ &\leq \left(\int_{\Omega} C^{1+\kappa}(\omega) P(d\omega) \right)^{1/(1+\kappa)} \cdot \sum_{j=1}^m \left(\int_{\Omega} I_{(p_j(x_\epsilon, \omega) > 0)} P(d\omega) \right)^{\kappa/(1+\kappa)} \\ &\leq c \cdot \sum_{j=1}^m P(p_j(x_\epsilon, \omega) > 0)^{\kappa/(1+\kappa)} \\ &\leq c \cdot m \cdot \varepsilon_{max}^{\kappa/(1+\kappa)}, \end{aligned}$$

where $c := \left(\int_{\Omega} C^{1+\kappa}(\omega) P(d\omega) \right)^{1/(1+\kappa)}$, which is finite due to the assumption (ii). Accordingly, for $\varepsilon_{max} \rightarrow 0_+$

$$0 \leq \sum_{j=1}^m \mathbb{E}[p_j(x_\varepsilon, \omega)] \leq c \cdot m \cdot \varepsilon_{max}^{\kappa/(1+\kappa)} \rightarrow 0,$$

and also $\beta_\varepsilon(x_\varepsilon) \rightarrow 0$. If we set

$$\varepsilon_j(x_N) = P(p_j(x_N, \omega) > 0), \quad j = 1, \dots, m,$$

then the optimal solution x_N of the expected value problem is feasible for the chance constrained program with $\varepsilon(x_N) = (\varepsilon_1(x_N), \dots, \varepsilon_m(x_N))$, because the following relations hold

$$\begin{aligned} P(g_{ji}(x_N, \omega) \leq 0, \quad i = 1, \dots, k_j) &\geq 1 - \varepsilon_j(x_N) \\ \iff P(p_j(x_N, \omega) > 0) &\leq \varepsilon_j(x_N). \end{aligned}$$

Hence, we get

$$\begin{aligned} \varphi_N &= f(x_N) + N \cdot \sum_{j=1}^m \mathbb{E}[p_j(x_N, \omega)] \\ &\geq f(x_{\varepsilon(x_N)}) + N \cdot \sum_{j=1}^m \mathbb{E}[p_j(x_N, \omega)] \\ &= \psi_{\varepsilon(x_N)} + \alpha_N(x_N). \end{aligned}$$

Finally,

$$\begin{aligned} \psi_\varepsilon &= \left(\psi_\varepsilon + \varepsilon_{max}^{-\gamma} \sum_{j=1}^m \mathbb{E}[p_j(x_\varepsilon, \omega)] \right) - \varepsilon_{max}^{-\gamma} \sum_{j=1}^m \mathbb{E}[p_j(x_\varepsilon, \omega)] \\ &\geq \varphi_{\varepsilon_{max}^{-\gamma}} - \varepsilon_{max}^{-\gamma} \sum_{j=1}^m \mathbb{E}[p_j(x_\varepsilon, \omega)] \\ &= \varphi_{\varepsilon_{max}^{-\gamma}} - \beta_\varepsilon(x_\varepsilon). \end{aligned}$$

This completes the proof.

The bounds (4) and the terms $\alpha_N(x)$, $\varepsilon(x)$ and $\beta_\varepsilon(x)$ depend on the choice of the penalty function ϑ . Two special penalty functions are readily available: $\vartheta^1(u) = \sum_{i=1}^k [u_i]^+$ applied in [4] and $\vartheta^2(u) = \max_{1 \leq i \leq k} [u_i]^+$. However, only ϑ^1 preserves convexity, whereas ϑ^2 may work better for joint chance constraints.

Notice, however, that when we want to evaluate one of the bounds in (4), we must be prepared to face some problems. We are able to compute $\alpha_N(x_N)$, $\varepsilon(x_N)$, hence the upper bound for the optimal value $\psi_{\varepsilon(x_N)}$ of the chance constrained program (2) with probability levels $\varepsilon(x_N)$. But we are not able to compute $\beta_{\varepsilon(x_N)}(x_{\varepsilon(x_N)})$ without having the solution $x_{\varepsilon(x_N)}$ which we do not want to find or even may not be able to find.

3 Generalization

Let H_j be set-valued mappings which map \mathbb{R}^n into measurable subsets of $\Omega \subset \mathbb{R}^s$ with closed graphs. According to [7, Theorem 5.7] mappings H_j are outer semicontinuous. Let the support Ω be compact and λ be a Borel probability measure on Ω . We consider the generalized chance constrained problem

$$\tilde{\psi}_\epsilon = \min\{f(x) : x \in X, \lambda(H_j(x)) \geq 1 - \epsilon_j, j = 1, \dots, m\}, \quad (5)$$

with optimal solution \tilde{x}_ϵ . According to [8, Proposition 3.1] the functions $h_j(x) = \lambda(H_j(x))$ are upper semicontinuous. Hence, $h'_j(x) = 1 - h_j(x) = \lambda(\Omega \setminus H_j(x))$ are lower semicontinuous. From [7, Theorem 5.7] and our assumptions it follows that the inverse mappings defined by

$$H_j^{-1}(\omega) = \{x \in \mathbb{R}^n : \omega \in H_j(x)\}, j = 1, \dots, m,$$

are outer semicontinuous. We define scalarization functions as

$$\tilde{p}_j(x, \omega) = \min_{z \in H_j^{-1}(\omega)} \rho(z, x), j = 1, \dots, m,$$

where $\rho(\cdot, \cdot)$ denotes a continuous metric on $\mathbb{R}^n \times \mathbb{R}^n$. The following equivalences hold and may help us to penalize possible infeasibilities:

$$\begin{aligned} \omega \in H_j(x) &\iff x \in H_j^{-1}(\omega) \iff \min_{z \in H_j^{-1}(\omega)} \rho(z, x) = 0, \\ \omega \in \Omega \setminus H_j(x) &\iff x \notin H_j^{-1}(\omega) \iff \min_{z \in H_j^{-1}(\omega)} \rho(z, x) > 0. \end{aligned}$$

Moreover, $\tilde{p}_j(x, \omega)$, $j = 1, \dots, m$, are continuous in x for all $\omega \in \Omega$ and lower semicontinuous in ω for all $x \in X$, see [1, Theorems 4.2.2, 4.2.3].

We show that the following penalty function problem

$$\tilde{\varphi}_N = \min_{x \in X} f(x) + N \sum_{j=1}^m \mathbb{E} \left[\min_{z \in H_j^{-1}(\omega)} \rho(z, x) \right]. \quad (6)$$

is asymptotically equivalent to the generalized chance constrained problem (5). We denote \tilde{x}_N the optimal solution of (6). The following generalization of Theorem 1 can be proved.

Theorem 2. Consider the two problems (5) and (6) and assume: $X \neq \emptyset$ compact, $f(x) = \mathbb{E}[g_0(x, \omega)]$ is a finite continuous function of x ,

Let H_j , $j = 1, \dots, m$ be set-valued mappings which map \mathbb{R}^n into measurable subsets of Ω with closed graphs and denote

$$\tilde{p}_j(x, \omega) = \min_{z \in H_j^{-1}(\omega)} \rho(z, x), j = 1, \dots, m,$$

where $\rho(\cdot, \cdot)$ denotes a continuous metric on $\mathbb{R}^n \times \mathbb{R}^n$.

Assume that

- (i) there exists a nonnegative random variable $C(\omega)$ with $\mathbb{E}[C^{1+\tilde{\kappa}}(\omega)] < \infty$ for some $\tilde{\kappa} > 0$, such that $|\tilde{p}_j(x, \omega)| \leq C(\omega)$, $j = 1, \dots, m$, for all $x \in X$;
- (ii) $\mathbb{E}[\tilde{p}_j(x', \omega)] = 0$, $j = 1, \dots, m$, for some $x' \in X$;
- (iii) There exist $x' \in X$ such that $\lambda(H_j(x')) = 1$, $j = 1, \dots, m$.

Denote $\gamma = \tilde{\kappa}/(2(1 + \tilde{\kappa}))$, and for arbitrary $N > 0$ and $\varepsilon \in (0, 1)$ put

$$\begin{aligned}\tilde{\varepsilon}_j(x) &= \lambda(\tilde{p}_j(x, \omega) > 0), \quad j = 1, \dots, m, \\ \tilde{\alpha}_N(x) &= N \cdot \sum_{j=1}^m \mathbb{E}[\tilde{p}_j(x, \omega)], \\ \tilde{\beta}_\varepsilon(x) &= \varepsilon_{max}^{-\gamma} \sum_{j=1}^m \mathbb{E}[\tilde{p}_j(x, \omega)],\end{aligned}$$

where ε_{max} denotes maximum of the vector $\varepsilon = (\varepsilon_1, \dots, \varepsilon_m)$ and $[1/N^{1/\gamma}] = (1/N^{1/\gamma}, \dots, 1/N^{1/\gamma})$ is the vector of length m .

THEN for any prescribed $\varepsilon \in (0, 1)^m$ there always exists N large enough so that minimization (6) generates optimal solutions x_N which also satisfy the chance constraints (5) with the given ε .

Moreover, bounds on the optimal value $\tilde{\psi}_\varepsilon$ of (5) based on the optimal value $\tilde{\varphi}_N$ of (6) and vice versa can be constructed:

$$\begin{aligned}\tilde{\varphi}_{1/\varepsilon_{max}^\gamma(\tilde{x}_N)} - \tilde{\beta}_{\varepsilon(\tilde{x}_N)}(\tilde{x}_{\varepsilon(\tilde{x}_N)}) &\leq \tilde{\psi}_{\varepsilon(\tilde{x}_N)} \leq \tilde{\varphi}_N - \tilde{\alpha}_N(\tilde{x}_N), \\ \tilde{\psi}_{\varepsilon(\tilde{x}_N)} + \tilde{\alpha}_N(\tilde{x}_N) &\leq \tilde{\varphi}_N \leq \tilde{\psi}_{[1/N^{1/\gamma}]} + \tilde{\beta}_{[1/N^{1/\gamma}]}(\tilde{x}_{[1/N^{1/\gamma}]})\end{aligned}\tag{7}$$

with

$$\lim_{N \rightarrow +\infty} \tilde{\alpha}_N(\tilde{x}_N) = \lim_{N \rightarrow +\infty} \tilde{\varepsilon}_j(\tilde{x}_N) = \lim_{\varepsilon_{max} \rightarrow 0_+} \tilde{\beta}_\varepsilon(\tilde{x}_\varepsilon) = 0$$

for any sequences of optimal solutions \tilde{x}_N and \tilde{x}_ε .

Proof. The proof is similar to those of Theorem 1 using properties of scalarization functions \tilde{p}_j , $j = 1, \dots, m$ and assumptions (i)-(iii).

4 Conclusion

Reformulation of chance constrained programs by incorporating a suitably chosen penalty function into the objective helps to arrive at problems with expectation in objective and a fixed set of feasible solutions. The obtained problems are much simpler to solve and analyze than the chance constrained programs. The recommended form of the penalty function follows the basic ideas of penalty methods and its suitable properties follow by generalization of results of [4]. The questions for future research are how to choose the parameter N so that the probability levels ε are ensured and to find bounds which can be evaluated.

References

- [1] Bank B., Guddat J., Klatte D., Kummer B., and Tammer K. (1982). *Non-linear parametric optimization*. Akademie-Verlag, Berlin.
- [2] Bazara M.S., Sherali H.D., Shetty C.M. (1993). *Nonlinear programming: theory and algorithms*. Wiley, Singapore.
- [3] Branda M., Dupačová J. (2008). *Approximations and contamination bounds for probabilistic programs*. SPEPS-13, 2008.
- [4] Ermoliev Y.M., Ermolieva T.Y., Macdonald G.J., Norkin V.I. (2000). *Stochastic optimization of insurance portfolios for managing exposure to catastrophic risks*. Annals of Operations Research **99**, 207–225.
- [5] Kall P., Mayer J. (2005). *Stochastic linear programming: models, theory, and computation*. International Series in Operations Research & Management Science, Kluwer Academic Publishers, New York.
- [6] Prékopa A. (1973). *Contributions to the theory of stochastic programming*. Mathematical programming **4**, 202–221.
- [7] Rockafellar R.T., Wets R. (2004). *Variational analysis*. 2nd edition, Springer-Verlag, Berlin.
- [8] Römisch W., Schultz R. (1991). *Stability analysis for stochastic programs*. Annals of Operations Research **30**, 241–266.

Acknowledgement: I would like to thank to my supervisor Doc. RNDr. Petr Lachout, CSc. and an anonymous reviewer for their valuable comments. This work was supported by the grant "Statistical Dynamic Models and their Applications in Economic, Natural and Technical Science" (201-05-H007).

Address: M. Branda, MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8 – Karlín

E-mail: branda@karlin.mff.cuni.cz

ZÓNOVÝ REGULAČNÍ DIAGRAM

Eliška Cézová

Klíčová slova: Regulační diagram, zónový regulační diagram, stabilita procesu, ARL, markovský model, optimalizace.

Abstrakt: Zónový regulační diagram je speciálním typem regulačního diagramu pro \bar{X} a R . V příspěvku je uveden nový návrh flexibilního zónového diagramu pro \bar{X} , který odstraňuje nevýhody klasického návrhu a svými vlastnostmi se blíží ke složitějšímu CUSUM diagramu. Zároveň je prezentována metoda pro výpočet průměrné doby běhu (ARL) tohoto diagramu pomocí Markovského řetězce. V tabulce jsou srovnány různé varianty zónového diagramu pomocí ARL při různých hodnotách posunu střední hodnoty procesu oproti předpokládané (stabilní) nulové hodnotě. Ve spojení s ekonomicko-statistickou optimalizací parametrů může tento diagram představovat velmi dobrou alternativu pro automatickou regulaci procesu.

Abstract: Zone control chart is a special type of control chart \bar{X} and R chart. In this article a new flexible zone control chart for \bar{X} chart is suggested, which eliminate disadvantage standard suggestion and his property near complicated CUSUM chart. Simultaneously is presented method for calculation of average run length (ARL) this chart whereby Markov chain. There are compared various variation zone control chart whereby ARL by various values of shift mean process over against notional (stable) zero value in the table. This chart can introduce very good alternative for automatic adjustment process in connection with economic-statistic optimalization parametr.

1 Úvod

V dnešní době se stále více výroba soustřeďuje na automatizaci výroby. Každý z nás zná mnoho případů automatických linek, bez kterých bychom si nedokázali představit dnešní výrobu. Díky tomuto pokroku, můžeme používat i pro automatizaci Shewhartovi regulační diagramy, pomocí nichž se snažíme odhalit chybu ve výrobě, ať už je to vlivem opotřebení stroje nebo jeho špatným nastavením.

V tomto příspěvku se zaměříme na Zónový regulační diagram, který je specifickým typem Shewhartova regulačního diagramu. Tento diagram je přitom pro obsluhu mnohem pochopitelnější než Shewhartův regulační diagram či např. CUSUM diagram.

Základní Shewhartův diagram vznikl od roku 1931. V roce 1956 byla doporučována běžná pravidla Western Electric Company. V roce 1959 Barnard navrhl CUSUM diagramy, které berou v úvahu jak minulost procesu tak jeho přítomnost. V roce 1971 Reynolds navrhl jednodušší diagram pro obsluhu. V roce 1987 Jaehn vyvinul zónový regulační diagram, který měl

bohužel velmi vysoký poměr falešných poplachů. V tom samém roce Champ a Woodall vypočítali běžná pravidla užitím Markovského řetězce.

V roce 1989 Jaehn modifikoval zónový regulační diagram, kde namísto ve skóre 1 původního zónového regulačního diagramu kolem centrální přímký dosadil 0. Tímto způsobem se výkon zónového regulačního diagramu podstatně vylepšil. V roce 1989 Hendrix přiblížil průměrnou délku přeběhu zónového regulačního diagramu pomocí simulací technik.

Zónový regulační diagram je doporučován v praxi díky jeho výkonu, jednoduchosti, efektivitě, snadnosti užití a snadnému porozumění.

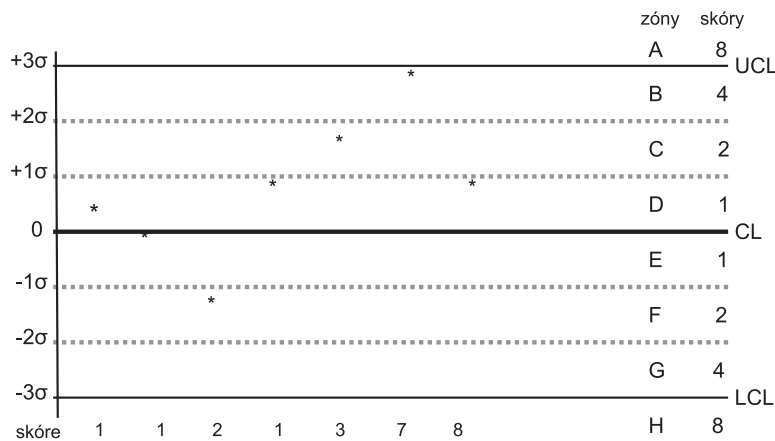
2 Zónový regulační diagram

Původní zónový regulační diagram obsahuje osm zón, čtyři na každé straně centrální přímký (CL). Zóny jsou pásy mezi přímkami rovnoběžnými s centrální přímkou, ve vzdálenosti $\pm 1\sigma$, $\pm 2\sigma$ a $\pm 3\sigma$. Na Obr. 1 jsou označeny písmeny A, B, C, D (nad CL) a E, F, G, H (pod CL). Zónám jsou přiřazeny skóry 8, 4, 2, 1, 1, 2, 4, 8 (na obr. 1 jsou uvedeny napravo od označení zón). Regulace potom probíhá obvyklým způsobem: v pravidelných intervalech jsou prováděna měření sledované veličiny, z nich je vypočítána odpovídající charakteristika (průměr, rozpětí, rozptyl) a tato je zakreslena jako jeden bod do regulačního diagramu. Těmto bodům jsou přiřazovány skóry podle toho, do které zóny padnou a tyto skóry se postupně sčítají. Součet se počítá vždy pro nepřerušovanou posloupnost po sobě jdoucích bodů na jedné straně CL. V okamžiku, kdy následující bod je na opačné straně CL než bod předchozí, součet je „vynulován“ a načítá se znovu. Signál o porušení stability je vyslán ve chvíli, kdy součet dosáhne hodnoty 8.

Dodatečná pravidla, vydaná Western Electric Company, určují kdy má být vyslán signál o porušení stability procesu i v případě, že nebyla překročena horní nebo dolní regulační mez. Lze je popsat následujícím způsobem:

- **Pravidlo 1** – jestliže nějaký bod padne mimo meze tři sigma (nad zónu B nebo pod zónu G dle obr.1.), získá skóre 8 bodů a je vyslán signál.
- **Pravidlo 2** – jestliže dva ze tří po sobě jdoucích bodů padnou na stejnou stranu CL do vzdálenosti větší než 2σ a ne větší než 3σ , je vyslán signál.
- **Pravidlo 3** – jestliže čtyři z pěti po sobě jdoucích bodů padnou na stejnou stranu CL do vzdálenosti mezi σ a 2σ , je vyslán signál.
- **Pravidlo 4** – signál je vyslán, jestliže osm po sobě jdoucích bodů padne na stejnou stranu CL do vzdálenosti 1σ .

Uvědomíme-li si, jak funguje zónový regulační diagram, dostáváme podobná pravidla:

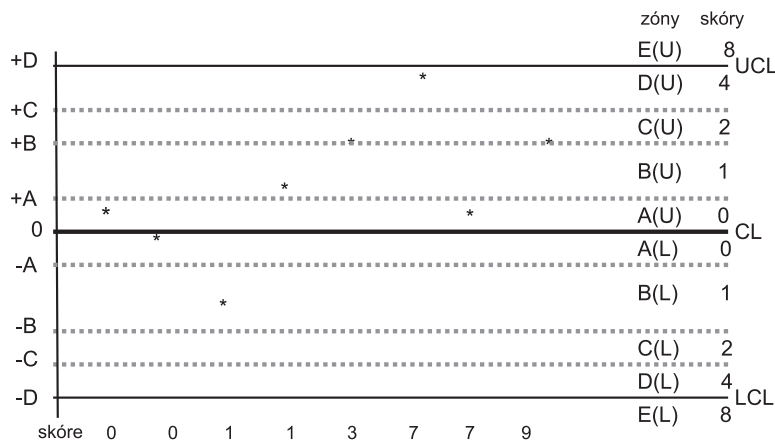


Obrázek 1: Původní zónový regulační diagram.

- **Pravidlo 1'** – jestliže nějaký bod padne mimo meze tři sigma (nad zónu B nebo pod zónu G dle obr.1.), získá skóre 8 bodů a je vyslán signál.
- **Pravidlo 2'** – jestliže tři po sobě jdoucí body padnou na stejnou stranu CL, nenastane situace z Pravidla 1 a nejméně dva z těchto bodů padnou do zóny B nebo G (dle obr. 1.), převyšší skóre 8 bodů a opět je vyslán signál.
- **Pravidlo 3'** – jestliže pět po sobě jdoucích bodů padne na stejnou stranu CL a z nich čtyři do zóny C nebo F, přičemž pátý nepadne od CL dále než 1σ , velikost skóre opět překročí hranici 8 bodů a je vyslán signál.
- **Pravidlo 4'** – signál je vyslán, jestliže osm po sobě jdoucích bodů padne do zóny D nebo E (dle obr. 1).

Cílem bylo umožnit **automatickou signalizaci** následných indikátorů stavu mimo kontrolu v Shewhartově diagramu. Kritici tohoto přístupu tvrdí, že v době výkonné výpočetní techniky není tento argument na místě, ale praxe ukazuje, že jednoduchost zónového diagramu spolu s jeho snadnou algoritmizací jej stále staví na přední místo v oblibě vedle klasického Shewhartova diagramu, který ovšem k aplikaci dodatečných pravidel vyžaduje paměť na předchozí naměřené hodnoty sledované charakteristiky.

Studium výkonnosti diagramu ukazuje, že je pouze mírně lepší než interpretovaná pravidla Shewhartových regulačních diagramů. Velmi rychle odhalí posun. Na druhou stranu, množství falešných poplachů (chyba 1. druhu) je velmi vysoká.



Obrázek 2: Nový zónový regulační diagram se skóre (0, 1, 2, 4, 8).

3 Návrh nového zónového regulačního diagramu

Navržená modifikace spočívá v přidání jedné zóny s nulovým skóre kolem centrální přímkou. Na Obr. 2 je označena jako A(U) resp. A(L). Ukazuje se, že výkonnost diagramu se touto úpravou zlepšuje.

Nový zónový regulační diagram je konstruován následujícím způsobem:

- Parametry diagramu jsou čtyři hodnoty A, B, C a D.
- V diagramu sestrojíme celkem osm přímk, rovnoběžných s CL ve vzdálenostech A, B, C a D od CL v oblasti nad a pod ní.
- Tyto přímk rozdělují diagram celkem do deseti zón, označených A(U), B(U), C(U), D(U), E(U) (nad CL) a A(L), B(L), C(L), D(L), E(L) (pod CL). Krajní zóny E(U) a E(L) zahrnují všechny body, které jsou od CL dále než 3σ (viz Obr. 2).
- Zónám jsou přiřazena skóre postupně 0, 1, 2, 4, 8 bodů (nejblíže CL je skóre 0 pro zóny A(U) a A(L), nejvyšší skóre je přiřazeno nejvzdálenějším zónám E(U) a E(L)).

Meze A, B, C a D jsou dány jako výsledky optimalizačního algoritmu, který minimalizuje pravděpodobnosti chyby druhého druhu pro hodnoty posunu δ_u (nerozlišující posuny) zachovává dané pravděpodobnosti chyby prvního druhu (falešných poplachů). V programu MATLAB byla aplikována optimalizační metoda Nelder-Mead simplexe, pomocí něhož byly vypočítány optimalizační meze, se kterými se dále pracuje.

4 Průměrná délka běhu (ARL)

Ke statistickému hodnocení a nastavení regulačních diagramů se běžně používá průměrná délka běhu ARL (average run length). Délka běhu je očekávaný (střední) počet inspekcí, které proběhnou do okamžiku, kdy regulační diagram vyše signál o vzniku zjištělné příčiny. Očekávaný (střední) počet inspekcí označme $ARL(0)$, kdy je proces pod kontrolou. $ARL(\delta)$ je očekávaný počet inspekcí, které proběhnou než bude indikována chyba, kdy střední hodnota procesu je posunuta o δ -násobek směrodatné odchylky sledované veličiny. Optimální regulační diagram je takový, který má co nejdelší $ARL(0)$ a pro δ , které nás zajímá, co nejkratší $ARL(\delta)$.

Podle hodnot ARL při různých posuvech střední hodnoty sledované charakteristiky můžeme vzájemně porovnávat různé typy regulačních diagramů. K výpočtu ARL se často používají Markovské řetězce. Přechodné stavy takového řetězce odpovídají žádanému chodu procesu pod statistickou kontrolou. Pokud výsledek měření způsobí vyslání signálu, tj. proces se dostane mimo kontrolu, odpovídající stav je absorpčním. Doba do absorpce je potom shodná s dobou do vyslání signálu, tedy přímo ARL. Pravděpodobnosti přechodu mezi jednotlivými stavy a do stavu absorpce závisí přímo na předpokládaném posuvu δ .

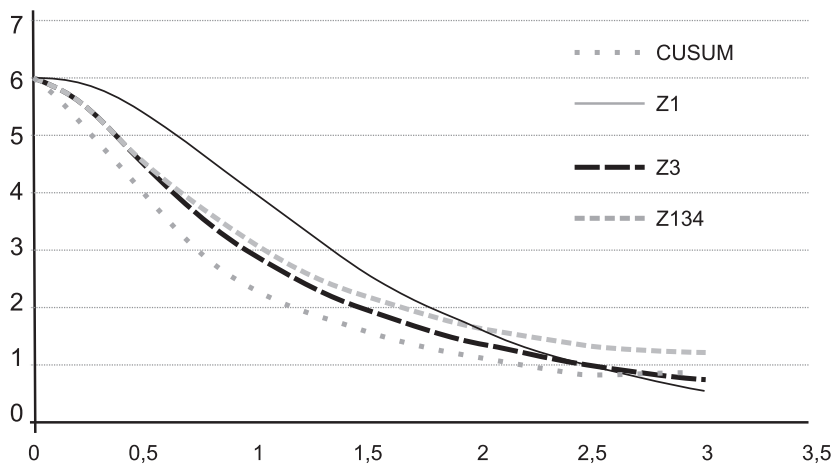
5 Markovský model

Předpokládejme markovský řetězec s těmito stavy $(0, 1, 2, 3, 4, 5, 6, 7, 8+)$. Všechny tyto stavy jsou přechodné kromě stavu $(8+)$, který je stavem absorpčním. Rozdělení doby do absorpce (pohlčení) X může být popsáno užitím PH rozdělení s parametry $(\vec{\pi}, P)$, kde $\vec{\pi}$ je počáteční rozdělení pravděpodobnosti a P je matice intenzit pravděpodobností přechodů mezi přechodnými stavy, která bude mít rozměr (8×8) . Předpokládáme začátek procesu ve stavu 0, tedy $\pi = (1, 0, 0, 0, 0, 0, 0, 0)$.

Střední hodnota X je dána vztahem $E(X) = \vec{\pi}(I - P)^{-1}\vec{e}$, kde I je jednotková matice a \vec{e} je vektor samých jedniček.

Nechť Z značí sledovanou charakteristiku. Pro libovolný posuv s jsme schopni ohodnotit následující pravděpodobnosti (užitím kumulativní distribuční funkce $F(z)$ ze Z). Matice P může být ohodnocena střední hodnotou těchto pravděpodobností. Matice P má potom následující podobu:

$$\begin{aligned} p_{00} &= p_0^2; & p_{01} &= p_0 p_A; & p_{02} &= p_0 p_B; & p_{03} &= p_{05} = p_{06} = p_{07} = 0; & p_{04} &= p_0 p_C; \\ p_{10} &= p_A^+ p_0^- + p_A^- p_0^+; & p_{11} &= p_0 p_A + 2p_A^+ p_{A-}; \\ p_{12} &= p_0 p_B + p_A^-(p_A^- + p_B^+) + p_A^+(p_A^+ + p_B^-); \\ p_{76} &= 0; & p_{77} &= (p_0^- + p_A^- + p_B^- + p_C^-)p_0^- + (p_0^+ + p_A^+ + p_B^+ + p_C^+)p_0^+; \end{aligned}$$



Obrázek 3: Srovnání diagramů: CUSUM, Shewhart, Z_3 a navržený zónový regulační diagram.

kde

$$\begin{array}{lll}
 p_0^- = P(-A < Z + s < 0) & p_0^+ = P(0 < Z + s < A) & p_0 = p_0^- + p_0^+ \\
 p_A^- = P(-B < Z + s < -A) & p_A^+ = P(A < Z + s < B) & p_A = p_A^- + p_A^+ \\
 p_B^- = P(-C < Z + s < -B) & p_B^+ = P(B < Z + s < C) & p_B = p_B^- + p_B^+ \\
 p_C^- = P(-D < Z + s < -C) & p_C^+ = P(C < Z + s < D) & p_C = p_C^- + p_C^+ \\
 p_D^- = P(Z + s < -A) & p_D^+ = P(D < Z + s) & p_D = p_D^- + p_D^+
 \end{array}$$

6 Příklad

Následující výpočty jsou provedeny v MATLABu a jejich výsledky jsou shrnuty v tabulce a dány do grafu. Odpovídající regulační diagramy jsou porovnávány pomocí ARL:

- CUSUM regulační diagram
- Z_1 - regulační diagram Shewhartova typu (s UCL a LCL)
- Z_3 - zónový regulační diagram, který signalizuje zjistitelnou příčinu po 4 následujících pozorováních Z , které padnou nad mez B
- Z_{1234} - navržený zónový regulační diagram

ARL (posun) byl vypočítán dle následujících podmínek:

- $ARL(0) = 400$ pro rozsah výběru $n = 1$ (při každé inspekci je provedeno pouze jediné měření sledované charakteristiky).

Stanovené meze po optimalizaci:

- CUSUM($k = 0,5; h = 5$);
- $Z_1(C = 3,03)$;
- $Z_3(B = 1,06)$;
- $Z_{1234} (A = 0,34; B = 1,24; C = 2,24; D = 3,37)$.

shift	CUSUM	Z_1	Z_3	Z_{1234}
0	400,96	400,00	400,75	400,44
0,3	92,87	272,04	168,40	164,36
0,6	24,09	127,64	51,77	50,40
0,9	11,19	59,15	21,67	20,02
1,2	7,10	29,29	11,30	10,08
1,5	5,16	15,66	7,46	6,45
1,8	4,09	9,04	5,50	4,53
2,1	3,41	5,62	4,58	3,49
2,4	2,93	3,75	4,14	2,79
2,7	2,58	2,68	3,91	2,31
3,0	2,32	2,04	3,79	1,92

7 Závěr

Navržený regulační diagram má lepší vlastnosti než původní zónový regulační diagram a než diagram Shewhartova typu. Při stejném $ARL(0)$, charakterizujícím střední frekvenci výskytu chyby prvního druhu (takzvaných „falešných poplachů“ kdy je vyslán signál, zatímco proces je ve stavu pod statistickou kontrolou) vykazuje rychlejší reakci na posun střední hodnoty sledované charakteristiky ($ARL(s)$). Navíc je jeho použití jednodušší než Shewhartův regulační diagram s dodatečnými pravidly.

Literatura

- [1] Davis R. B., Homer A., Woodall W. H. (1990). *Performance of the zone control chart*. Commun. Statist.-Theory Meth. **19** (5), 1581–1587.
- [2] Dohnal G. (2008). *Regulační diagram, ale jaký?* Sborník z konference REQUEST'08, v tisku
- [3] Zhang S., Wu Z. (2005). *Designs of control charts with supplementary runs rules*. Computers & Industrial Engineering, **49**, 76–97.
- [4] Cézová E. (2008). *A new type of zone control charts*. Current Trends in Statistics in V6 region, in print.

Poděkování: Tato práce byla řešena v rámci projektu MŠMT 1M06047 Centrum pro jakost a spolehlivost výroby.

Adresa: E. Cézová, ČVUT FSI, Ústav technické matematiky, Karlovo náměstí 13, 121 35 Praha 2

E-mail: eliskacqr@email.cz

KLASIFIKÁCIA ZAŠUMENÝCH DÁT

Katarína Cimermanová

Kľúčové slová: Robustná klasifikačná metóda, zašumené dáta, analýza dychu, fajčiarsky návyk.

Abstrakt: Klasifikácia viacrozmerých pozorovaní do dvoch tried je dôležitý problém. Existuje niekoľko klasifikačných metód na zatriedenie pozorovaných vektorov do jednej z dvoch tried, avšak v reálnom živote sú vektory pozorovaní zašumené. Riešením klasifikácie zašumených dát je robustná formulácia vychádzajúca z metódy oporných bodov. Formulácia je konvexný optimalizačný problém, ktorý je súčasťou problematiky kónického programovania druhého rádu. V robustnej formulácii sa predpokladá elipsoidálny model šumu. Nie je nutný predpoklad typu rozdelenia pozorovaných dát, predpokladá sa len konečnosť momentov druhého rádu.

Robustnú klasifikačnú metódu aplikujeme na analýzu vydychovaných plynov, kde sa budeme venovať klasifikácii dobrovoľníkov do skupiny fajčiarov a nefajčiarov.

Abstract: Classification of multidimensional data into one of two classes is an important issue. There are some classification methods which classify data to one of two classes, but in real live situation the observation vectors are noisy. Solution to this problem is a robust formulation that stems from the Support Vector Machine method. The formulation is a convex optimization problem; in particular, it is an instance of the Second Order Cone Programming problem. An ellipsoidal uncertainty model is assumed in the robust formulation. It is derived from a worst case consideration and assumes only the existence of the second order moments.

The robust classification method is applied to breath gas analysis, where we classify volunteers to group of smokers and non-smokers.

1 Úvod

Analýza dychu (*Breath analysis*, BA) ako neinvazívna diagnostická metóda sa stáva v poslednom období atraktívnou témou. Ukazuje sa, ako potenciálna metóda na skorú detekciu rakoviny pľúc a pažeráka (viac detailov napr. v projekte 6. rámcového programu EU pod skratkou BAMOD - *Breath-gas analysis for molecular-oriented detection of minimal diseases*). Rakovina pľúc je choroba, ktorá je s fajčením spájaná najčastejšie. Preto sa zameriavame na analýzu vydychovaných plynov fajčiarov a nefajčiarov zdravej populácie.

Ľudský dych obsahuje veľké množstvo rôznych prvkov, presnejšie prchavých organických zložiek (*volatile organic compounds*, VOCs), ktoré sú produkované endogénne v ľudskom tele. Na meranie koncentrácií prchavých organických zložiek vydychovaného plynu sa využívajú nové metódy, bližšie popísané

napr. v [1]. Jednou z metód je hmotnostná spektrometria s protónovou prenosovou reakciou (*proton-transfer-reaction mass spectrometry*, PTR-MS). Molekulové hmotnosti detekované PTR-MS sú v rozmedzí od m/z 21 do m/z 230. Namerané množstvá (*counts*) sú prepočítané na základe znalostí z kinetickej chémie na koncentrácie v jednotke časť na miliardu (*particles-per-billion*, ppb), viď bližšie [4]. Zdokonaľovanie procesu merania a analýzy dát pomocou PTR-MS (kvalita a spoľahlivosť získaných meraní) je predmetom ďalšieho výskumu v tejto oblasti. Preto predpokladáme, že namerané koncentrácie prchavých organických zložiek sú zašumené.

2 Robustná klasifikačná metóda

Klasifikácia subjektov do dvoch tried je dôležitým problémom. Na riešenie daného problému existuje niekoľko metód. Jednou z nich je aj metóda oporných bodov (*support vector machine*, SVM). Jedná sa o klasifikátor, ktorý sa snaží lineárne oddeliť dve skupiny tak, aby vzdialenosť najbližších bodov (*support vectors*) od deliacej nadroviny bola čo najväčšia prihliadajúc k čo najmenej chybovosti, bližšie napr. v [5]. Na rovnakom princípe je založená aj klasifikačná metóda, ktorá predpokladá zašumené vstupné dáta, viac v [2] a [6].

Predpokladajme, že naše namerané dáta $\mathbf{X} \in R^N$ sú zašumené a skutočná hodnota je nejaký bod v špecifikovanom elipse, teda predpokladáme elipsoidálny model zašumenia. Nech

$$B(\bar{\mathbf{X}}, \Sigma, \gamma) = \{\mathbf{X} : (\mathbf{X} - \bar{\mathbf{X}})' \Sigma^{-1} (\mathbf{X} - \bar{\mathbf{X}}) \leq \gamma^2\}$$

je elipsoid, Σ je pozitívne semidefinitná matica a $\gamma \geq 0$ je skalár. Pre vstupné dáta nie je nutný predpoklad typu rozdelenia, predpokladáme len konečnosť momentov druhého rádu.

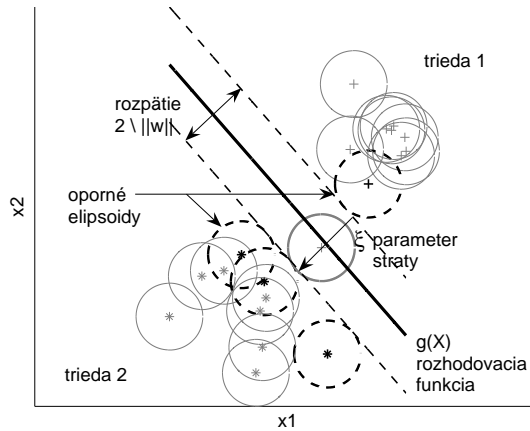
Riešením klasifikácie zašumených dát $\mathbf{X} \in B(\bar{\mathbf{X}}, \Sigma, \gamma)$ je nájdenie optimálnych parametrov \mathbf{w} , \mathbf{b} rozhodovacej funkcie

$$g(\mathbf{X}) = \text{sign}(\langle \mathbf{w}, \mathbf{X} \rangle + \mathbf{b}),$$

tak, aby rozpätie (*margin*) dvoch paralelných nadrovín ku $g(\mathbf{X})$ oddeľujúcich dáta bolo čo najväčšie, teda rovné $2/\|\mathbf{w}\|$. V prípade lineárne neseparovateľných dát ide o maximalizáciu rozpätia tak, aby bol čo najmenší počet zle klasifikovaných pozorovaní, teda minimálna strata ďalej charakterizovaná voľnými (*slack*) parametrami straty $\xi \geq 0$. Elipsoidy, pre ktoré platí

$$y(\langle \mathbf{w}, \mathbf{X} \rangle + b) \geq 1,$$

pre $\forall \mathbf{X} \in B$ a rovnosť paltí len v jednom z bodov, teda sa dotýkajú jednej z paralelných nadrovín a kde $y \in \{1, -1\}$ je kategorizácia pozorovania do jednej z dvoch tried, hrajú rolu tzv. oporných bodov teda ich budeme nazývať oporné elipsoidy (*support ellipsoids*). Tieto elipsoidy sú postačujúce



Obrázok 1: Schéma riešenia metódy na klasifikáciu zašumených dát do dvoch tried, ktorej riešením je nájdenie parametrov rozhodovacej funkcie tak, aby rozpätie medzi dvoma paralelnými nadrovinami bolo čo najväčšie a v prípade lineárne neseparovateľných dát bola strata čo najmenšia.

pri popise rozhodovacej funkcie $g(\mathbf{X})$, predstavujú len malý zlomok všetkých dát, takže efektívny počet bodov definujúcich rozhodovaciu funkciu $g(\mathbf{X})$ je omnoho menší ako počet subjektov v trénovacej množine. Na obrázku 1 je znázornený princíp klasifikačnej metódy.

Riešením je optimalizačná úloha kvadratického programovania

$$\min_{\mathbf{w}, \mathbf{b}, \xi} \quad \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \xi_i$$

$$\text{s podm.} \quad y_i (\langle \mathbf{w}, \mathbf{X}_i \rangle + \mathbf{b}) \geq 1 - \xi_i$$

$$\xi_i \geq 0,$$

pre $\forall X \in B(\bar{X}_i, \Sigma_i, \gamma_i)$ a $i = 1, \dots, n$, kde n je rozsah trsénovacej vzorky a parameter C je regularizačná konštanta, ktorá rieši kompromis medzi maximalizáciou rozpätia a stratou.

Optimalizačná podmienka sa využitím Karush-Kuhn-Tuckerových podmienok dá prepísať na tvar $y_i \langle \mathbf{w}, \mathbf{X}_i \rangle = y_i \langle \mathbf{w}, \bar{X}_i \rangle - \gamma_i \|\Sigma_i^{1/2} \mathbf{w}\|$, [2]. Potom nasledovná robustná formulácia je ekvivalentná s predchádzajúcou optimalizačnou úlohou

$$\min_{\mathbf{w}, \mathbf{b}, \xi} \quad \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \xi_i$$

$$\begin{aligned} \text{s podm.} \quad y_i(\langle \mathbf{w}, \bar{\mathbf{X}}_i \rangle + \mathbf{b}) &\geq 1 - \xi_i + \gamma_i \|\Sigma_i^{\frac{1}{2}} \mathbf{w}\| \\ \xi_i &\geq 0, \end{aligned}$$

pre $i = 1, \dots, n$, kde robustnou ju robí nelineárny člen $\|\Sigma_i^{\frac{1}{2}} \mathbf{w}\|$ nachádzajúci sa v obmedzujúcich podmienkach. Optimalizačná úloha sa rieši ako úloha kónického programovania druhého rádu (*second order cone programming*, SOCP)

$$\min_{\mathbf{w}, \mathbf{b}, \xi} \sum_{i=1}^n \xi_i$$

$$\begin{aligned} \text{s podm.} \quad y_i(\langle \mathbf{w}, \bar{\mathbf{X}}_i \rangle + \mathbf{b}) &\geq 1 - \xi_i + \gamma_i \|\Sigma_i^{\frac{1}{2}} \mathbf{w}\| \\ \|\mathbf{w}\| &\leq W \\ \xi_i &\geq 0, \end{aligned}$$

pre $i = 1, \dots, n$, kde člen $\|\mathbf{w}\|$ je presunutý do podmienky a ohraničený zhora konštantou W , ekvivalentnou s konštantou C . Ide o konvexný optimalizačný problém, v ktorom lineárna funkcia je minimalizovaná na prieniku lineárnej variety (*linear manifold*) s produktom Lorentzových kužeľov, ktoré predstavujú jednotlivé podmienky.

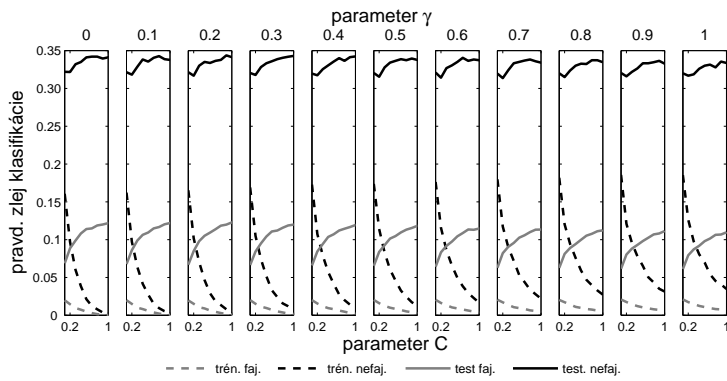
3 Analýza dychu

Pilotná štúdia zostavená na Lekárskej Univerzite v Innsbrucku v rokoch 2006 a 2007 obsahuje namerané koncentrácie prchavých organických zložiek pre 54 fajčiarov a 178 nefajčiarov, u ktorých nebola potvrdená diagnóza rakoviny pľúc. Každému subjektu v databáze sa priradila jedna hodnota, ktorá predstavuje medián z mediánov jednotlivých vzoriek vydýchnutého vzduchu, ktorý bol pre presnosť meraný trikrát. Na spracované dáta aplikujeme klasifikačnú metódu na základe zatriedenia podľa fajčiarskeho návyku.

Klasifikačnú metódu sme otestovali pre naše dáta v dvoch simulačných štúdiách. Výsledky sú priemery pravdepodobností chýb zo 100-krát náhodne rozdelenej databázy fajčiarov a nefajčiarov na tréningovú a testovaciu množinu v pomere 3:2. Predpokladáme $\Sigma_i = \sigma^2 I$, kde σ^2 predstavuje rozdiel medzi najvyššou a najnižšou nameranou koncentráciou v testovacej množine.

V prvej simulačnej metóde sledujeme závislosť empirickej pravdepodobnosti zle klasifikovaných subjektov skupiny fajčiarov a nefajčiarov pri tréningu a testovaní klasifikačnej metódy pre kombinácie parametrov $C = 0.1 : 0.1 : 1$ a $\gamma = 0 : 0.1 : 1$. Výsledky simulačnej metódy sú znázornené na obrázku 2.

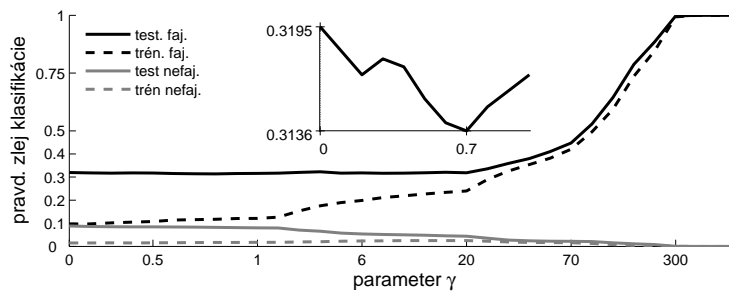
Z grafu vidieť, že klasifikačná metóda prednostne zatriedí subjekty z početnejšej skupiny. Pre rastúci parameter C chybovosť tréningu klesá a zároveň rastie pri testovaní. Pri vyšších hodnotách C dochádza k pretrénovaniu (*overfitting*) klasifikátora. Pre naše dáta sme za optimálny zvolili parameter $C = 0.2$, nakoľko pravdepodobnosť zle zatriedených fajčiarov z tréningovej



Obrázok 2: Empirická pravdepodobnosť zle klasifikovaných subjektov fajčiarov a nefajčiarov pri testovaní a tréňovaní metódy na klasifikáciu zašumených dát pre kombinácie parametrov $\gamma = 0 : 0.1 : 1$ a $C = 0.1 : 0.1 : 1$.

množiny pre všetky parametre γ je najnižšia. Z grafu ďalej vidieť, že pravdepodobnosť zlého zatriedenia subjektov z testovacej množiny pri zvyšujúcom sa šume (parameter γ) pomaly klesá. Tým sme ukázali, že predpoklad, že aplikované dáta sú zašumené znižuje pravdepodobnosť zlého zatriedenia pri testovaní.

Pri veľkej hodnote parametra γ môže dochádzať k situácii, že dáta nie sú lineárne oddeliteľné. Preto sme pre parameter $C = 0.2$ zostrojili ďalšiu simulačnú metódu, kde pravdepodobnosť zlého zatriedenia bola odhadnutá na základe 100-krát náhodne rozdelenej skupiny fajčiarov a nefajčiarov na tréňovaciu a testovaciu v pomere 3:2 pre parameter $\gamma = 0 : 1000$. Výsledky sú znázornené na obrázku 3.



Obrázok 3: Empirická pravdepodobnosť zle klasifikovaných subjektov fajčiarov a nefajčiarov pri testovaní a tréňovaní metódy na klasifikáciu zašumených dát pre kombinácie parametrov $C = 0.2$ a $\gamma = 0 : 1000$.

Najlepšie výsledky dosahovala metóda pre parameter $\gamma = 0.7$ ($P_{testF} = 0.31$, $P_{testN} = 0.08$). Pre $\gamma \geq 20$ klasifikačná metóda nebola schopná oddeliť jednotlivé skupiny, pre $\gamma \geq 300$ zatrieduje všetky subjekty do početnejšej skupiny.

4 Záver

V práci bola prezentovaná klasifikačná metóda na klasifikáciu zašumených dát. Z výsledkov vyplýva, že pri predpoklade zašumenia dát sa znižuje pravdepodobnosť zlej klasifikácie subjektov pri testovaní. Ďalším zlepšením výsledkov by mohlo byť implementovanie nelineárneho mapovania premenných zašumených dát do priestoru s vyššou dimenziou, čím by klasifikačná metóda mala dosiahnuť ešte lepšie výsledky zatriedenia testovaných subjektov.

Literatúra

- [1] Amann A., Smith D. (2005). *Breath analysis for clinical diagnosis and therapeutic monitoring*. World Scientific, Singapore.
- [2] Bhattacharyya Ch. (2004). *Robust classification of noisy data using second order cone programming approach*. In Proceedings International Conference on Intelligent Sensing and Information Processing, 433–438.
- [3] Bi J., Zhang T. (2004). *Support vector classification with input data uncertainty*. Advances in Neural Information Processing systems (NIPS'4), **17**, 161–168.
- [4] Kushch I., et al. (2008). *Compounds enhanced in a mass spectrometric profile of smokers' exhaled breath versus non-smokers as determined in a pilot study using PTR-MS*. Journal of Breath Research, **2**, 1–26.
- [5] Moguerza J.M., Muñoz A. (2006). *Support vector machines with applications*. Statistical Science, **21**, 3, 322–336.
- [6] Shivaswamy P.K., Bhattacharyya Ch., Smola A.J. (2006). *Second order cone programming approach for handling missing and uncertain data*. Journal of Machine Learning Research, **7**, 1283–1314.
- [7] Sturm J.F. (1995). *Using SEDUMI 1.02, a Matlab*toolbox for optimization over symmetric cones*. (Updated for Version 1.05), Optimization Methods and Software, **11**, 625–653.

Podakovanie: Práca bola podporovaná EU projektom BAMOD: LSHC-CT-2005-019031 STREP, Agentúrou na podporu výskumu a vývoja (APVV), grant RPEU-0008-06, Vedeckou grantovou agentúrou Ministerstva školstva SR a Slovenskej akadémie vied (VEGA), grant 1/3016/06 a 2/7087/27.

Adresa: K. Cimermanová, ÚM SAV, Dúbravská cesta 9, 841 04 Bratislava, Slovenská republika

E-mail: katarina.cimermanova@gmail.com

PROBLEMS WITH EXTREMES

Jan Dienstbier

Keywords: Extreme value index, Hill estimator, quantile regression.

Abstract: Univariate extreme value theory (EVT) along with all of its tools, is bound to the validity of the basic assumption, the maximum domain of attraction (MDA) condition. The aim of this paper is to briefly suggest what effect a violation of the condition has on inference methods of EVT. Apart from describing some previously suggested tests, which check the MDA condition, we show how common estimators of the extreme value index γ behave if the MDA doesn't hold. We also pay attention to the inference of extremes in linear models.

Abstrakt: Cílem práce je zkoumat vliv porušení podmínek za nichž platí jednorozměrná teorie extrémálních hodnot.

1 Introduction – Basic assumptions

Extreme value theory (EVT) is a branch of statistics devoted to studying phenomena related to extreme, i.e. high or low, quantiles of distributions. The purpose is to estimate or test hypothesis dealing with extreme quantiles, probabilities, and return periods of events pertaining to the underlying distribution F of a given random sample. Such task can be accomplished only under certain assumptions imposed on the distribution F . More precisely we suppose that the distribution function (d.f.) belongs to the maximum domain of attraction (MDA) of an extreme value distribution G_γ with some $\gamma \in \mathbb{R}$, i.e there exist constants $a_n > 0$, $b_n \in \mathbb{R}$ such that

$$\lim_{n \rightarrow \infty} F^n(a_n x + b_n) = G_\gamma(x) := \exp(-(1 + \gamma x)^{-1/\gamma}), \quad (1)$$

for all $1 + \gamma x > 0$. In case of $\gamma = 0$, $G_0(x)$ is interpreted as $\exp(-e^{-x})$. The previous relation to the domain of attraction is commonly denoted as $F \in \text{MDA}(G_\gamma)$. The parameter γ is called the extreme value index (EVI). From the asymptotic point of view, an adequate approximation of the extreme quantiles and probabilities can be obtained by the respective distribution G_γ , see de Haan and Ferreira [7]. However, we usually do not know the correct EVI γ and hence the importance of its adequate estimation is evident.

During the fast development of the EVT in the last thirty years a lot of various estimates of γ appeared, each based on a different strategy, see e.g. Beirlant et al. [1] or de Haan and Ferreira [7]. Nevertheless, all these estimates share some common properties. Particularly, as γ depends on the properties of the tails only, the estimates depend on the choice of k_n (or k), the number of the highest observations used for the estimation. Roughly speaking, each estimate of the EVI is a function of k . It holds that with growing k the

asymptotic variance of the estimate decreases whereas its asymptotic bias grows. Various techniques have been created determining “the best” k or, equivalently, an optimal portion of the data. Alternatively, other techniques try to reduce the bias of the estimate. Nevertheless, all methods draw their raison d’être from another assumption often imposed in the context of EVT – the second order condition. The second order condition extends the approximation of the tail of F beyond that of (1) up to the second order. More precisely, suppose there exists some positive function a and some positive or negative function A with $\lim_{t \rightarrow \infty} A(t) = 0$ such that for all x with $1 + \gamma x > 0$,

$$\lim_{t \uparrow x_*} \frac{\frac{1-F(t+xa(1/(1-F(t))))}{1-F(t)} - (1+\gamma x)^{-\frac{1}{\gamma}}}{A(1/(1-F(t)))} = (1+\gamma x)^{-\frac{1+\gamma}{\gamma}} H_{\gamma,\rho} \left(-\frac{1}{\gamma}(1-x^{-\gamma}) \right),$$

where $H_{\gamma,\rho}$ is some function that is not a multiple of the function $-\frac{1}{\gamma}(1-x^{-\gamma})$. In particular $H_{\gamma,\rho}$ should not be identically zero. The second order condition implies the MDA condition. The function $H_{\gamma,\rho}$ can be written as

$$H_{\gamma,\rho}(x) = \frac{1}{\rho} \left(\frac{x^{\gamma+\rho} - 1}{\gamma + \rho} - \frac{x^\gamma - 1}{\gamma} \right), \quad (2)$$

which for the cases $\gamma = 0$ or $\rho = 0$ is understood to be equal to the limit of (2) as $\gamma \rightarrow 0$ or $\rho \rightarrow 0$, respectively. As the MDA, or say the first order condition, describes to which extreme value distribution the sample appertains, the second order parameter ρ shows in which “direction” the maxima of the underlying distribution deviate from the respective G_γ . It follows that the asymptotic bias of an arbitrary estimator of γ is a function of the second order parameter ρ and A .

These are the fundamental assumptions used by EVT. In the next section we provide some answers to a natural question what happens if they do not hold.

2 When assumptions fails

We are accustomed to say that all “basic” distributions belong to some domain of attraction. Indeed, normal, exponential, and Laplace distributions represent the examples of the Gumbel domain of attraction ($\gamma = 0$), uniform and Beta belong to the Weibull domain of attraction ($\gamma < 0$) with restricted tails, and Cauchy, Pareto, and Student’s t-distribution are common examples of the distributions belonging to the heavy-tailed Frechet domain ($\gamma > 0$). On the other hand, we know distributions which do not lie in any of these domains. Consider e.g. classical von Mises example

$$F(x) = 1 - e^{-x - \sin(x)}, \quad x \geq 0. \quad (3)$$

The distribution function (3) is infinitely differentiable for $x \geq 0$, yet it does not belong in any domain of attraction due to the oscillating character of its

tail, see de Haan and Ferreira [7, pp. 36]. One can imagine a lot of distributions similar to (3). Also some discrete distributions behave similarly, e.g. Poisson distribution. Another example is truncated exponential distribution $F \sim X = \exp[E]$, where $[E]$ denotes an integer part of the exponentially distributed random variable.

There also exists a slightly different class of distributions with arbitrarily smooth and strictly increasing distribution function without any oscillation as in (3) which, nevertheless, do not belong to any domain of attraction. By the EVT asymptotic expansions, any distribution belonging to the Frechet domain of attraction $\gamma > 0$ can be approximated in tail by generalized Pareto distribution. Hence the distribution function $F(x) = 1 - (\log x)^{-\beta}$, $x \geq e$ and $\beta > 0$, which cannot be approximated by any G_γ , does not belong to any domain of attraction. Roughly speaking, such distributions are “too heavy” – their tails are heavier than of any distribution in the Frechet domain of attraction, c.f. Falk et al. [6, pp. 64].

Suppose we are faced with the following practical task. On the basis of the random sample generated from an unknown distribution we ought to perform a test or an estimate related to the extreme events, e.g. calculate high conditional quantiles. An usual approach is to calculate these characteristics using EVT – estimate γ and then approximate the tail of the unknown distribution of the sample by G_γ . But what happens, if we are not given a sample from a distribution fulfilling the domain of attraction condition? Can we discover solely from the estimates of γ itself that something is wrong and the given task cannot be accomplished using EVT?

We provide a partial answer by the accompanying figures. On Figure 1, which is presented for the comparison, one can see the plots of different estimates of γ against the growing k , the number of the highest observations used

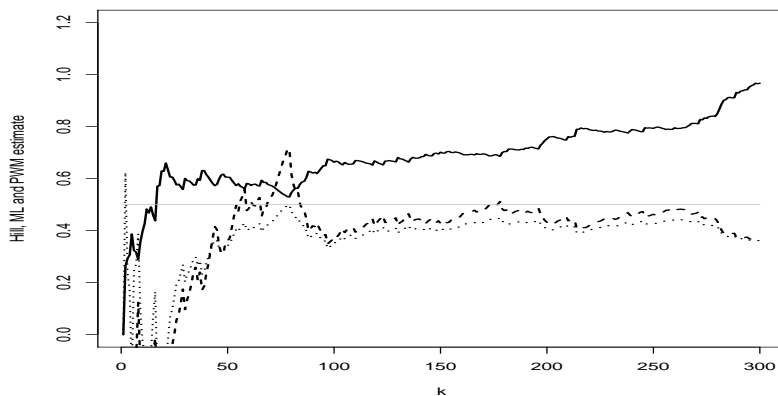


Figure 1: Student’s t-distribution with two degrees of freedom.

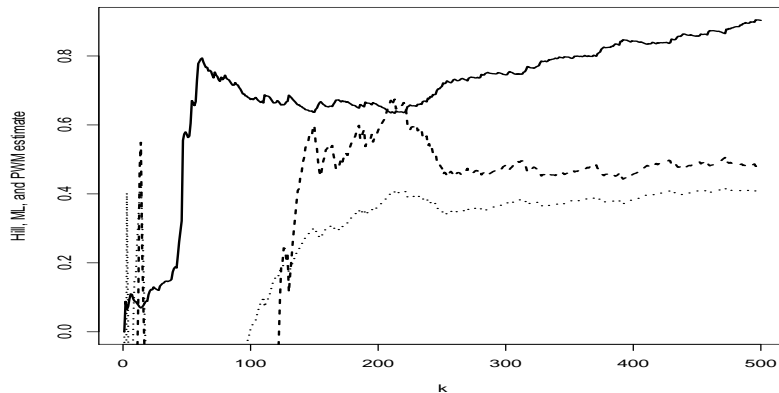


Figure 2: von Mises example $F(x) = 1 - e^{-x - \sin x}$.

for the estimation. The data were generated from the Student’s t-distribution with two degrees of freedom – this distribution belongs to the Fréchet domain of attraction with $\gamma = \frac{1}{2}$. There are plots of the three estimates of γ , the Hill estimate (solid line), the maximum likelihood (ML) estimator (dashed), and the probability weighted moments (PWM) estimator (dotted). Although each of the estimators is based on a different idea, c.f. de Haan and Ferreira [7, pp. 65–126], all curves representing estimated values of γ are roughly the same. Note that all the plots demonstrate the mentioned properties of the estimators of γ – with growing k the asymptotic variance decreases while the asymptotic bias increases. The dilemma which k one should choose can be solved by an adaptive selection of the optimal k on the basis of the sample, c.f. Beirlant et al. [1, pp. 123–129]. The choice is often based simply on an expert decision. One tries to find where the “region of stability” lies, i.e. the first area where the estimate $\hat{\gamma}_n(k_n)$ remains roughly constant. Such method works surprisingly well in practice, although some exceptions exist.

Confront the previous with Figure 2. The situation is different. While one can (with some difficulty) detect a “region of stability” in each graph, the estimator plots sharply differs. In this example the underlying sample distribution have been simulated from the von Mises distribution (3). Therefore, given a sample, calculating the results for more than just one estimator of γ can be a helpful tool for detecting that the whole EVT setting is inappropriate. This is even more important conclusion, since any estimator plot separately may represent a “typical” example how the estimator plot of the EVI “looks like”.

Finally, consider Figure 3. We calculated the estimators from the distribution $F(x) = 1 - 1/\log^{1/1000}(x)$. As noted before, the distribution doesn’t belong to any domain of attraction. One can see that this fact is not reflected neither on the ML nor on the PWM estimate, while results of the Hill esti-

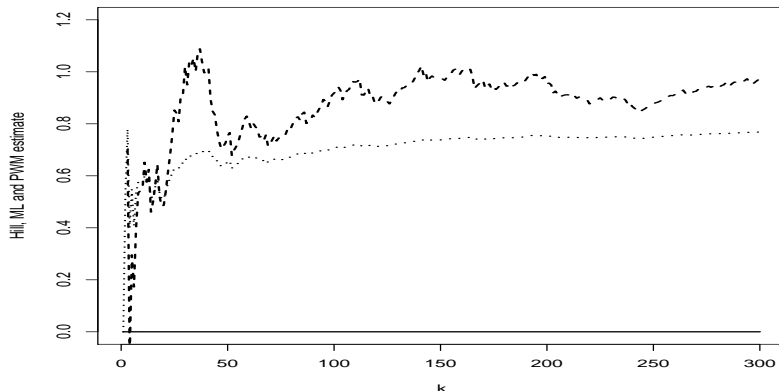


Figure 3: $F(x) = 1 - \frac{1}{\log^{1/1000}(x)}$.

mate (the solid line near the axis) look at least suspicious. Again, the results of the different estimators pointed out that something is wrong.

Clearly, the use of more than one estimator is highly profitable. As each estimator is based on a different strategy of estimating the parameter γ of G_γ , each reacts differently on a violation of the necessary condition. Hence if the estimators show sharply different results, we should ask about the validity of our assumptions. To validate the suspicion one can further use some tests checking the validity of the MDA condition.

3 Tests checking the MDA condition

There had not been paid much attention to testing of the MDA condition yet, nevertheless some results has been published in the past decade. The article Dietrich et al. [3] can be considered as a pioneering work. The topic have been further studied in Drees et al. [5] and Hüsler and Li [8]. In our article, we concentrate on these methods only. Aside from these contributions, one can, being strongly convinced that the sample is heavy tailed, use some goodness-of-fit tests, see Neves and Fraga Alves [9] for references.

Consider the test statistic introduced by Dietrich et al. [3]

$$E_n(k) := k \int_0^1 \left(\frac{\log X_{n-[kt]:n} - \log X_{n-k:n}}{\hat{\gamma}_+} - \frac{t^{-\hat{\gamma}_-} - 1}{\hat{\gamma}_-} (1 - \hat{\gamma}_-) \right)^2 t^\eta dt, \quad (4)$$

where $\hat{\gamma}_+$ and $\hat{\gamma}_-$ are the positive and the negative part of a suitable estimate $\hat{\gamma}$ and $\eta > 0$ is some chosen constant. Under usual condition upon the growth of k , the limit distribution of $E_n(k)$ has been established using an asymptotic expansion of tails due to Drees [4]. For critical point see the numerical study in Hüsler and Li [8]. The limit distribution of $E_n(k)$ has been

described thoroughly only under the null hypothesis, but in Dietrich et al. [3] we can find also some examples indicating that the statistic $E_n(k)$ goes to infinity at least under some alternatives. Nevertheless, the behaviour of $E_n(k)$ under whole spectrum of alternatives has not been studied. Note also that the growth of k can be controlled only through the parameters introduced with the second order condition and hence it is necessary to assume that the second order condition also holds. That can be seen as a little bit of paradox – while testing whether the MDA condition holds, one needs to assume that even the stronger condition holds – remark that the second order condition is sufficient but absolutely not necessary for the MDA. However, the exhaustive study provided in Hüsler and Li [8] shows that the test itself is not very sensitive to the violation of the second order condition, at least in majority of the cases.

Slightly different statistics introduced Drees et al. [5] to test the validity of $F \in \text{MDA}(G_\gamma)$, $\gamma > -1/2$. For any $\eta > 0$ they came with

$$T_n(k) := k \int_0^1 \left(\frac{n}{k} \bar{F}_n \left(\hat{a} \left(\frac{n}{k} \right) \frac{x^{-\hat{\gamma}} - 1}{\hat{\gamma}} + \hat{b} \left(\frac{n}{k} \right) \right) - x \right)^2 x^{\eta-2} dx, \quad (5)$$

for some $\eta > 0$, where $\bar{F}_n = 1 - F_n$ and $\hat{\gamma}$, \hat{a} , and \hat{b} are suitable consistent estimators – it is recommended to use the ML estimators for $\hat{\gamma}$ and \hat{a} , and $\hat{b}(n/k) := X_{n-k:n}$. Again the consistency of the estimator depends on the suitably growing k and the second order condition. One can find the critical values of (5) and a broad numerical comparison between the tests based on (5) and (4) in Hüsler and Li [8].

Summarizing the foregoing, both testing methods share common properties: Both depend on k , the number of the observations used for the calculation. There exist slightly different tests depending on the choice of the estimate $\hat{\gamma}$, yet the difference is negligible under the null hypothesis. We also obtain different tests for different choices of parameter η . Moreover, not much is known about what happens under the alternative. This holds even more for (4) as the statistic (5) has the meaning of the distance between the excesses over $\hat{b}(\frac{n}{k})$ and the approximating Pareto distribution. Hence also the plot of (5) against k can be used as a tool to determine where the Pareto approximation is inaccurate.

For these reasons it is hard to recommend any of these tests as a preliminary filter checking the validity of the EVT model. Such approach could lead to confusing results. Rejecting the EVT model we restrain from the general, asymptotically correct, and simple solution of tail modelling. We should really be persuaded that the concept of EVT does not fit the model before trying even more awkward solutions. Hence the lesson from this section is: Although we often cannot evade estimating of γ we should be aware of the fact that such estimation is not necessarily sufficient in some cases. And, what is even more important, we have the tools to check, whether our assumptions are right.

4 Tests in linear models

The situation is even more complicated if we want to estimate γ in linear regression models, i.e. estimate γ associated with the distribution of “errors” in the model

$$\mathbf{Y}_{n \times 1} = \mathbf{X}_{n \times p} \boldsymbol{\beta}_{p \times 1} + \mathbf{E}_{n \times 1},$$

where $\mathbf{Y}_n = \mathbf{Y} = (Y_1, \dots, Y_n)^\top$ is a vector of observations, $\mathbf{X}_n = \mathbf{X}$ is a known design matrix, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top \in \mathbb{R}^p$ is an unknown parameter, and $\mathbf{E}_n = \mathbf{E} = (E_1, \dots, E_n)^\top$ is the vector of *i.i.d.* errors with a (generally unknown) distribution function F . As one can expect, the choice of some k “largest” observations is not an obvious task in such setting.

In Berlaint et al. [1] authors suggested to find a covariate dependent threshold determining the “largest” observations by a regression quantile $\hat{\beta}(\alpha)$ for α near one. Assuming that β_1 is the intercept, i.e. the first column of \mathbf{X}_n is $(1, \dots, 1)^\top$, we define the empirical regression quantiles $\hat{\beta}_n$ as the solutions of the minimization problem

$$\hat{\beta}_n(\alpha | Y, \mathbf{x}) := \arg \min_{b \in \mathbb{R}^p} \sum_{i=1}^n \rho_\alpha(Y_i - b^\top \mathbf{x}_i), \quad (6)$$

where ρ_α denotes the loss function $\rho_\alpha(u) := u \cdot (\alpha - 1_{\{u < 0\}})$.

In Dienstbier [2] we treated the estimation of γ on the basis of regression quantiles using same techniques as Drees [4] used for i.i.d. random variables. Suppose we have $k = k_n$ highest regression quantiles of the intercept, where k is growing suitably slowly to infinity. In Dienstbier [2] we showed that location and scale invariant smooth functionals of the standardized intercepts of the highest order regression quantiles have the same asymptotic distribution as the same functionals based on the empirical tail quantile function of the underlying sample of errors. It follows that in the situation of linear model same as (4) we can calculate all estimates of γ from suitably chosen regression quantiles instead of unobservable highest order statistics of errors \mathbf{E} . This also proves that the method suggested in Berlaint et al. [1], i.e. to base the estimation of γ on the exceedances over the high quantile regression threshold, leads to consistent estimates.

However, what can be done if different estimates of γ based on the regression quantiles show suspicious results similar to those in Section 2? Clearly, a test checking the validity of the MDA in linear models is needed. In Dienstbier [2] we established strong approximations of the standardized intercepts of regression quantiles to the tail quantile function of the sample of errors. Hence, following the steps of Drees [4] and Drees [5], one can show that also the tests statistic

$$E_n(k) := k \int_0^1 \left(\frac{\log \hat{\beta}_1(1 - \frac{kt}{n}) - \log \hat{\beta}_1(1 - \frac{k}{n})}{\hat{\gamma}_+^Q} - \frac{t^{-\hat{\gamma}_-^Q} - 1}{\hat{\gamma}_-^Q} (1 - \hat{\gamma}_-^Q) \right)^2 t^n dt, \quad (7)$$

has, under null hypothesis $\gamma \in \text{MDA}(G_\gamma)$, $\gamma \in \mathbb{R}$, the same distribution as (4) in the i.i.d. case. Here $\hat{\gamma}^Q$ denote a suitable estimator for the γ based on the regression quantiles as in Dienstbier [2] and $\hat{\beta}_1(\cdot)$ denotes the empirical regression quantile estimating the intercept of the model. To sketch the proof we refer to Theorem 1 of Dietrich et al. [3], which describes asymptotic properties of (4). These properties follows from Theorem 2.1. of Drees [4], an analogue to Theorem 2.1. of Dienstbier [2]. The result finally follows from the continuity of the integral in (7). Details will be published in the forthcoming PhD thesis of the author.

Conclusion is similar as in the previous section. Again we strongly recommend to calculate more different estimates of γ instead of just one. If the estimates sharply differ, consider whether the basic assumptions hold, e.g. make a test checking validity of the MDA condition.

References

- [1] Beirlant J. et al. (2004). *Statistics of extremes, Theory and application*. Wiley, Chichester.
- [2] Dienstbier J. (2008). *Estimators of the extreme value index based on quantile regression*. Revstat, submitted.
- [3] Dietrich D., de Haan L., Hüsler J. (2002). *Testing extreme value conditions*. Extremes **5**, 71–85.
- [4] Drees H. (1998). *On smooth statistical tail functionals*. Scandinavian Journal of Statistics **25**, 187–210.
- [5] Drees H., de Haan L., Li D. (2006). *Approximations to the tail empirical distribution function with application to testing extreme value conditions*, J. Statist. Plann. Inference **136**, 3498–3538.
- [6] Falk M., Hüsler J., Reiss R.D. (2004). *Laws of small numbers: Extremes and rare events*. Birkhäuser, Switzerland.
- [7] de Haan L., Ferreira A. (2006). *Extreme value theory: An introduction*, Springer Verlag, New York.
- [8] Hüsler J., Li D. (2006). *On testing extreme value conditions*. Extremes **9**, 69–86.
- [9] Neves C., Fraga Alves M.I. (2008). *Testing extreme value conditions – overview and recent approaches*. Revstat, Vol. 6, No. 1, 83–100.

Acknowledgement: This work was supported by grant GAČR 201/05/H007.

Address: J. Dienstbier, MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8 – Karlín

E-mail: dienstbi@karlin.mff.cuni.cz

O OPTIMALIZACI REGULAČNÍCH DIAGRAMŮ

Gejza Dohnal

Klíčová slova: Statistická regulace procesu, regulační diagram, ekonomicko-statistický návrh, stabilita procesu, ARL, markovský model, optimalizace.

Abstrakt: Při aplikaci statistické regulace výrobního procesu prostřednictvím regulačních diagramů je třeba postupovat obezřetně, na základě promyšleného návrhu. Příspěvek má za cíl ukázat hlavní kroky při návrhu regulačního diagramu, mezi něž patří vedle důkladné analýzy výrobního procesu a jeho pravděpodobnostních charakteristik také výběr vhodného typu regulačního diagramu a optimalizace jeho parametrů. Ekonomicko-statistický návrh bere do úvahy jak ekonomické parametry (náklady na měření, identifikaci vymezitelné poruchy, zásah v případě falešného poplachu a v případě oprávněné signalizace, ztráty z provozu mimo kontrolu a další), tak i statistické vlastnosti (pravděpodobnost vzniku poruchy, předpokládanou dobu, po kterou proces bude pod kontrolou, předpokládanou dobu běhu mimo kontrolu, pravděpodobnost falešného poplachu a další).

Abstract: The design of control chart is the important part of application of the statistical process control methodology into practice. This paper shows that not only economic-statistical design is essential but the comprehensive analysis of the process is needed. The main target of statistical process control as a part of the manufacturing process management is to keep the process in controllable state which gives the optimal conditions for quality production. For this purpose we use two key tools: control charts for monitoring and maintenance management for improving the manufacturing process.

1 Úvod

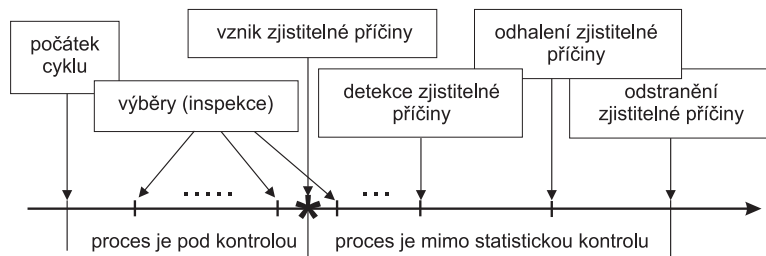
Regulační diagramy jsou velice rozšířeným nástrojem především v oblasti řízení jakosti. Jsou velice oblíbené pro svoji jednoduchost a snadnou aplikovatelnost. Je až s podivem, že se od třicátých let minulého století, kdy byly poprvé uvedeny Shewhartem, v podstatě nezměnily. Nicméně, jejich vývoj se nezastavil a je stále aktuální [5].

Hlavním cílem regulačního diagramu je sledování posloupnosti náhodných veličin (pozorování náhodného procesu v diskrétním čase) a detekce změny jeho parametrů. Základní typy regulačních diagramů byly vytvořeny pro sledování parametru polohy (střední hodnoty, mediánu) nebo rozptýlenosti (rozptyl, rozpětí) posloupnosti nezávislých náhodných veličin s normálním nebo binomickým rozdělením. Zpravidla se jedná o sledování nějakého jakostního znaku při výrobě. Slovo „proces“ je zde používáno ve dvojitým smyslu: jako výrobní proces (z hlediska provozního) a jako náhodný proces (z hlediska

matematického) tvořený realizací sledovaného znaku v čase. Pokud je výrobní proces stabilní (nepůsobí v něm jiné než náhodné vlivy), je sledovaný náhodný proces slabě stacionární. Jeho střední hodnota a rozptyl jsou konstantní v čase. Variabilita sledovaného znaku je způsobena tzv. „běžnou příčinou“ (common cause) o které se předpokládá, že je inherentní vlastností procesu a nelze ji změnit, aniž bychom změnili proces samotný. Kromě toho se v náhodném čase objeví tzv. „zjistitelná příčina“ (assignable cause), která je způsobena nějakou změnou podmínek, zpravidla poruchou. Porucha způsobí trvalou změnu charakteristiky procesu - buď posun střední hodnoty nebo změnu variability, případně obojí. Tato příčina je vymezitelná, to znamená, že předpokládáme, že ji lze identifikovat a po jejím odstranění proces vrátit do „původního stavu“. Jednou z funkcí regulačních diagramů je rozlišit mezi těmito dvěma příčinami, což není tak úplně snadné. Obě příčiny jsou často vzájemně propojeny a jsou odlišitelné pouze v kontextu. To, co je běžnou příčinou dnes, může být zjistitelnou příčinou zítra.

Stacionarita sledovaného náhodného procesu je považována za znak stability výrobního procesu. Po dobu, kdy se náhodný proces chová jako stacionární, říkáme že výrobní proces je „pod statistickou kontrolou“. Po vzniku zjistitelné příčiny (poruchy) se charakteristiky náhodného procesu změní a výrobní proces dostane „mimo statistickou kontrolu“. Nalezením a odstraněním zjistitelné příčiny opět navrátíme výrobní proces pod statistickou kontrolu. Za předpokladu, že poruchy vznikají nezávisle na sobě v čase a doby mezi nimi mají stejné rozdělení pravděpodobnosti, dále že doby do nalezení a odstranění zjistitelné poruchy jsou nezávislé a stejně rozdělené, tvoří okamžiky „navrácení“ procesu pod statistickou kontrolu proces obnovy. Cyklus obnovy začíná v okamžiku navrácení výrobního procesu pod statistickou kontrolu, zahrnuje fázi kdy je proces pod kontrolou až do okamžiku vzniku zjistitelné příčiny (fáze 1), dále fázi mimo statistickou kontrolu, kdy probíhá identifikace a odstranění této příčiny (fáze 2) a končí v okamžiku, kdy je zjistitelná příčina odstraněna. Délka tohoto cyklu závisí na řadě parametrů procesu samotného a parametrů regulačního diagramu.

Proces regulace pomocí regulačního diagramu probíhá tak, že v diskrétních časových okamžicích provádíme několik opakovaných měření sledované



Obrázek 1: Regulační cyklus výrobního procesu.

veličiny na různých vzorcích – takzvaný výběr. Z výsledků tohoto výběru spočteme regulační charakteristiku (průměr, rozpětí, směrodatnou odchylku či jinou statistiku sledované veličiny) a tuto zakreslíme do regulačního diagramu. Regulační cyklus má potom podobu která je znázorněna na obr. 1 [4]:

2 Statistické vlastnosti regulačního diagramu

Cílem regulace procesu je jeho udržení ve stabilním stavu s minimální variabilitou. To v důsledku znamená udržení co nejlepší kvality výroby. Ve skutečnosti však samotná regulace nepřispívá ke zvýšení stability či snižování variability procesu. Proces je ve stabilním stavu pouze „přechodně“. Předpokládáme, že dříve či později vždy dojde k posunutí střední hodnoty, ke zvýšení variability či k jinému porušení stability procesu. Úkolem regulačního diagramu je co nejdříve toto porušení stability indikovat a vyslat příslušný signál. Za určitých podmínek můžeme regulaci chápat jako posloupnost sekvencí testů pro indikaci bodu změny.

Od „dobrého“ regulačního diagramu se očekává, že

- vyšle signál o detekci zjistitelné příčiny co nejdříve po jejím vzniku. To znamená, že proces bude co nejkratší dobu ve stavu mimo kontrolu a bude možno rychle odstranit zjistitelnou příčinu a tím proces vrátit do stavu pod kontrolou.
- nebude vysílat falešné signály v době, kdy je proces pod statistickou kontrolou.

Budeme-li chápat situaci, kdy regulační diagram neodhalí zjistitelnou příčinu přestože už nastala, jako chybu II. druhu, potom první vlastnost souvisí s minimalizací pravděpodobnosti této chyby. Naopak, situaci kdy regulační diagram detekuje zjistitelnou příčinu, která ještě nenastala, můžeme v jistém smyslu mluvit o chybě I. druhu. Potom druhá vlastnost je úzce spojena s pravděpodobností chyby I. druhu. Při výběru regulačního diagramu, případně při jeho optimalizaci, bychom měli sledovat především minimalizaci pravděpodobností obou těchto chyb.

Všeobecně uznávanou mírou pro velikost pravděpodobnosti chyb I. a II. druhu je střední doba běhu (Average Run Length, ARL). Ta vyjadřuje střední počet výběrů při daných podmínkách do vyslání signálu o detekci zjistitelné příčiny. Tak například v případě regulačního diagramu pro střední hodnotu (\bar{X} diagram) budeme symbolem $ARL(0)$ označovat očekávaný (střední) počet výběrů (měření) sledované charakteristiky vycentrováného procesu (jehož střední hodnota je stále rovna požadované hodnotě), než dojde k falešnému signálu, tedy k chybě I. druhu. Naproti tomu, $ARL(d)$ je očekávaný počet výběrů, které proběhnou než bude indikována chyba, za předpokladu, že proces má střední hodnotu posunutou o d -násobek směrodatné odchylky sledované veličiny. Ideálně by tato hodnota měla být rovna 1, což znamená okamžitou detekci v prvním výběru z takového procesu.

Je zřejmé, že optimální regulační diagram by měl mít co nejdelší $ARL(0)$, zatímco $ARL(d)$ by měla být co nejkratší. Pro některé typy provozů však je žádoucí, aby $ARL(d)$ klesala pomalu a malé hodnoty nabývala až od nějakého d_u počínaje dál. Důvodem je to, že malé vychýlení například střední hodnoty může být v takovém provozu nevýznamné a je často chápáno jako „falešný poplach“. Ukazuje se, že pro srovnání různých typů regulačních diagramů stačí porovnat jejich ARL při různých posuvech střední hodnoty sledované charakteristiky.

Optimalizace regulačního diagramu z hlediska $ARL(\cdot)$ se obvykle provádí nastavením parametrů rozhodovací funkce, která na základě výsledku měření (výběru) rozhodne o vyslání varovného signálu. Výpočet $ARL(\cdot)$ může být poměrně komplikovaná procedura. V literatuře se používá několik metod, z nichž nejlepší výsledky dává metoda výpočtu ARL pomocí Markovského řetězce. Při této metodě se možné výsledky vyhodnocení výsledku měření v každém výběru, které nevedou k vyslání signálu o detekci zjistitelné příčiny, chápou jako přechodné stavy nějakého Markovského řetězce. Těchto stavů je obvykle konečný počet. Výsledek měření, který způsobí vyslání signálu, je v tomto modelu chápán jako stav absorpce. Potom doba do absorpce takto zkonstruovaného Markovského řetězce je doba do vyslání signálu, tedy přímo ARL . Hlavní problém při použití tohoto modelu je výpočet pravděpodobností přechodu mezi jednotlivými stavy a do stavu absorpce. Tyto pravděpodobnosti v sobě zahrnují i informaci o případném posunu sledované charakteristiky.

Příklad: Uvažujme zónový diagram pro střední hodnotu, obsahující čtyři limitní hladiny ve vzdálenostech $\pm A\sigma$ a $\pm B\sigma$ od centrální přímký nahoru i dolů. To znamená, že diagram je rozdělen do šesti zón. Při rozhodování se používají následující pravidla:

Signál je vyslán, pokud je sledovaná charakteristika

- $4x$ v zóně (A, B) , aniž by mezi tím nabyla hodnoty na opačné straně centrální přímký (CL)
- $2x$ nad limitou B , aniž by mezi tím nabyla hodnoty na opačné straně CL
- $2x$ v zóně (A, B) a $1x$ nad limitou B , aniž by mezi tím nabyla hodnoty na opačné straně CL
- $4x$ v zóně $(-A, -B)$, aniž by mezi tím nabyla hodnoty na opačné straně CL
- $2x$ pod limitou $-B$, aniž by mezi tím nabyla hodnoty na opačné straně CL
- $2x$ v zóně $(-A, -B)$ a $1x$ pod limitou $-B$, aniž by mezi tím nabyla hodnoty na opačné straně CL

Těmto pravidlům odpovídají následující skóry:

- 0 pro zónu $(-A, A)$
- 2 pro zónu $(-B, -A)$ nebo (A, B) ,
- 4 pro překročení limity $\pm B$.

Skóry se postupně načítají, v případě překročení CL se tento součet počítá znovu od nuly. Pokud součet skórů dosáhne hodnoty 8, regulační diagram vysílá signál [1].

Pro výpočet *ARL* budeme uvažovat Markovský řetězec s pěti stavy, odpovídajícími počtu skórů, tedy stavy $\{0, 2, 4, 6, 8\}$. První čtyři stavy jsou přechodné, pátý stav lze považovat za absorpční.

Ke stanovení pravděpodobností přechodů je třeba navíc předpokládat určité vlastnosti měřené charakteristiky – v tomto případě rozdělení pravděpodobnosti aritmetického průměru měření při jednom výběru \bar{X} a nezávislost výběrů. Budeme tedy předpokládat nezávislá, stejně rozdělená pozorování s normálním rozdělením $N(d, \sigma^2)$. Matice pravděpodobností přechodu bude mít tedy rozměr 5×5 a dobu do absorpce (*ARL*) můžeme popsat rozdělením fázového typu s reprezentací (π, A) , kde π je vektor počátečního rozdělení $(1, 0, 0, 0)$ a A je matice pravděpodobností přechodu mezi prvními čtyřmi stavy [3].

Optimální nastavení limit dostaneme minimalizací funkce $arlZ23(d_u, k)$ vzhledem k proměnné k při podmínce $arlZ23(0, k) = a_0$. Pro tuto minimalizaci lze použít například simplexovu metodu Nelder-Mead.

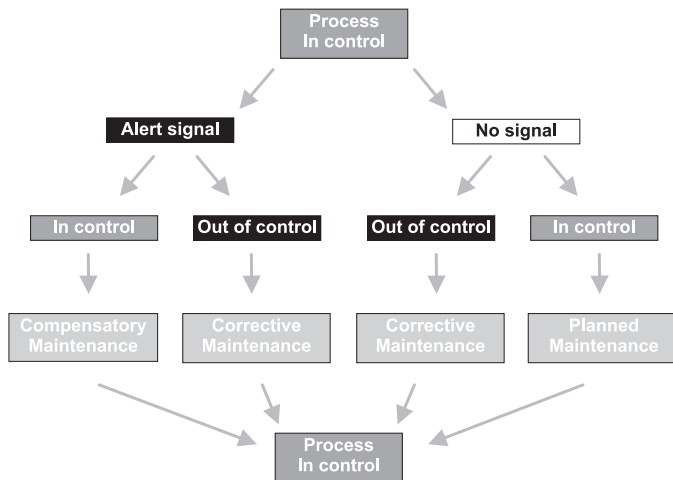
3 Ekonomické aspekty regulace

Udržení maximální kvality je ekonomicky náročné a často není ani nutné. V řadě případů je postačující „přijatelná“ kvalita za „rozumnou“ cenu. Toho lze dosáhnout vhodnou volbou parametrů regulačního diagramu – především doby mezi pozorováním (měřením), počtu pozorování (měření) pro výpočet sledované charakteristiky, nastavením parametrů rozhodovací funkce, na jejímž základě je vyslán signál o tom, že proces je mimo statistickou kontrolu (typicky se jedná o nastavení regulačních mezí). To je předmětem ekonomicko-statistického návrhu regulačního diagramu. V současné praxi se obvykle parametry výběru (frekvence a počet měření) stanoví expertním odhadem a kritické meze jsou nastaveny na trojnásobek směrodatné odchylky sledované charakteristiky. Ekonomicko-statistický návrh bere do úvahy jak ekonomické parametry (náklady na měření, identifikaci vymezené poruchy, zásah v případě falešného poplachu a v případě oprávněné signalizace, ztráty z provozu mimo kontrolu a další), tak i statistické vlastnosti (pravděpodobnost vzniku poruchy, předpokládanou dobu, po kterou proces bude pod kontrolou, předpokládanou dobu běhu mimo kontrolu, pravděpodobnost falešného poplachu a další). Tato dvě hlediska – ekonomické a statistické – nelze od sebe oddělovat. V literatuře jsou publikovány příklady, kdy ekonomická optimalizace sice přislíbí nejnižší náklady, nicméně za cenu příliš nízkého *ARL*(0) nebo příliš malého rozsahu výběru a tím i příliš širokých konfidenčních intervalů pro sledovanou charakteristiku. Naopak, pouze statistická optimalizace nemůže zaručit dosažení co nejnižších nákladů na regulaci. Bohužel, v praxi se velmi často provádí regulace klasickým Shewhartovým regulačním diagramem (s UCL a LCL ve vzdálenosti $\pm 3\sigma$ od centrální přímkou) a intervaly mezi výběry, stejně jako rozsah těchto výběrů se určuje „expertním“ odhadem.

V případě, kdy regulační proces můžeme chápat jako proces obnovy, je sledována střední cena (náklad) na hodinu provozu, kterou spočteme jako podíl středního nákladu na cyklus a střední délky cyklu. Optimalizace se provádí vzhledem ke ztrátové funkci, která vychází z Taguchiho ztrátové funkce. Minimalizaci této ztrátové funkce provádíme za předem daných omezení, například na velikost $ARL(0)$ nebo rozsahu výběrů. Existuje řada numerických metod, jak hledat podmíněné minimum vícerozměrné funkce. Nám se osvědčila simplexová metoda Nelder-Mead.

Nejznámějším modelem pro ekonomickou optimalizaci je model Lorenzen-Vance (Lorenzen T.J. a Vance L.C., 1986), který vychází z výše uvedeného rozdělení regulačního cyklu a vyjadřuje délku trvání a náklady pro každou jeho část. V literatuře lze nalézt řadu modifikací pro různé regulační modely a různé předpoklady o chování sledovaného stochastického procesu [2],[4].

Nedílnou součástí procesu regulace jsou i následná opatření po vyslání signálu o výskytu zjištěné příčiny. Proto je třeba do výpočtu zahrnout i opravy a údržbu výrobního zařízení, [6]. Optimální nastavení údržbové strategie totiž ovlivní předpokládanou dobu do poruchy a tedy do vzniku zjištěné příčiny. Zahrnutí údržby do modelu řízení procesu rozšiřuje klasický Lorenzen-Vance model o několik dalších variant. Na obrázku 2 jsou zobrazeny jako čtyři různé cesty grafem.



Obrázek 2: Model regulace procesu s údržbou.

Pro klasický Shewhartův regulační diagram nám optimalizační procedura vrátí tři parametry: délku mezi výběry h , rozsah výběru n a hodnotu UCL. Pokud bychom požadovali předem danou hodnotu $ARL(0)$, bude hodnota UCL dána tímto požadavkem. Pro požadavek $ARL(0) = 400$, který byl uveden v příkladu, vychází $UCL=3,23$.

4 Návrh regulačního diagramu

Pro praktickou aplikaci je třeba provést návrh regulačního diagramu. Pokud bude regulační diagram chybně navržen (v praxi se tento návrh velice často opomíjí zcela), potom jeho aplikace nemůže přinést optimální výsledky. To je paradoxně i častým důvodem pro odmítání regulačních diagramů jako neefektivního nástroje, který pouze zvyšuje náklady [5], [3].

Návrh regulačního diagramu by měl projít následujícími šesti kroky:

1. analýza výrobního procesu,
2. stochastická analýza,
3. výběr regulačního diagramu,
4. výběr optimální strategie údržby,
5. ekonomicko-statistický návrh,
6. nastavení pravidel pro aplikaci.

4.1 Analýza výrobního procesu

Tato analýza by nám měla odpovědět na následující otázky:

- kterou veličinu/veličiny budeme sledovat?
- které vlastnosti výrobního procesu tato veličina ovlivňuje a naopak, jak je ovlivňována tímto procesem?
- co přesně znamená z provozního hlediska, že je „proces pod kontrolou“?
- jaká jsou rizika v případě, kdy je proces „mimo kontrolu“?
- může proces pokračovat v průběhu identifikace vymezené příčiny (ověřování signálu)? Jaké je riziko?
- může proces pokračovat v průběhu údržbových operací? Jaké je riziko?
- jaké jsou nákladové položky (ztráty) a jejich vyčíslení?

Teprve po zodpovězení těchto otázek můžeme pokračovat v návrhu.

4.2 Stochastická analýza

Vzhledem k tomu, že regulační diagram pracuje s „vnořeným“ stochastickým procesem, nelze provést návrh regulačního diagramu bez znalosti nebo alespoň odhadů všech pravdě-podobnostních charakteristik a analýzy závislosti. Minimálně je třeba odpovědět na tyto otázky:

- jaké je pravděpodobnostní rozdělení sledovaných charakteristik?
- jaké je pravděpodobnostní rozdělení doby do poruchy?
- (auto)korelační analýza sledovaných charakteristik v čase
- jaké je pravděpodobnostní rozdělení trvání opravy nebo údržby

4.3 Výběr regulačního diagramu

Nabídka různých variant regulačních diagramů je velmi široká (a přesto se v drtivé většině používají klasické Shewhartovy diagramy pro \bar{X} a R). Pro výběr té nevhodnější je třeba vzít do úvahy

- charakter sledované charakteristiky/charakteristik
- zda je třeba sledovat jednu či více charakteristik současně
- závislostní struktura sledovaného (stochastického) procesu
- požadavek na citlivost regulačního diagramu
- požadavek na jednoduchost aplikace za daných podmínek

4.4 Výběr optimální strategie údržby

Dobře načasovaná preventivní údržba může významně snížit náklady na výrobu a zvýšit její kvalitu. Proto je třeba určit optimální intervaly a rozsah plánované údržby. Pro ekonomicko-statistický návrh regulačního diagramu je třeba stanovit i vhodný typ údržby:

- renovace (replacement) – uvede systém do stavu „jako nový“,
- minimální oprava – uvede systém do stavu jako před poruchou,
- neúplná oprava – mezi renovací a minimální opravou.

4.5 Ekonomicko-statistický návrh

Výsledkem této části návrhu regulačního diagramu jsou především parametry

- výběru (rozsah výběru, délka intervalu mezi výběry v hodinách),
- rozhodovací funkce (regulační meze, skóry).

4.6 Pravidla pro aplikaci

Pravidla pro aplikaci jsou velmi důležitá: sebelepší návrh regulačního diagramu nebude funkční, pokud nebude aplikován správným způsobem. Proto je třeba tato pravidla stanovit už v okamžiku návrhu regulačního diagramu a pokud možno zajistit jejich dodržování. Mezi pravidla pro aplikaci patří především:

- organizační opatření (podpora ze strany vedení, vyčlenění a pravomoce pracovníků),
- podmínky měření (zabezpečení odběru vzorků a jejich změření),
- odpovědnost pracovníků (a jejich kvalifikační předpoklady),
- technické podmínky (vyhrazený prostor pro odběr vzorků, zastavení provozu při signálu, při údržbě, ...),
- programové a výpočetní zabezpečení (evidence a vyhodnocení výsledků měření).

5 Závěr

Tento příspěvek je především metodologický. Popisuje postup, který by měl provázet každé zavedení regulačních diagramů do praktického provozu. Bez uvedených postupů nelze zaručit, že aplikace statistické regulace bude plnit požadované funkce efektivně. Řada z těchto zásad se v praxi zanedbává a výsledkem je častá skepse k regulačním nástrojům.

Literatura

- [1] Cézová E. (2008). *A new type of zone control charts*. Sborník konference ČStS Current Trends in Statistics in V6 region (v tisku).
- [2] Davis R.B., Homer A., Woodall W.H. (1990). *Performance of the zone control chart*. Commun. Statist.-Theory Meth. **19** (5), 1581–1587.
- [3] Dohnal G. (2008). *Cost-Effective Design of Control Charts: Optimal Design of the Zone Control Chart*. Applied Stochastic Models in Business and Industry (AMSB-08-144, v tisku).
- [4] Prahbu S.S., Montgomery D.C., Runger G. (1997). *Economic-statistical design of an adaptive X chart*. Int. J. Production Economics **49**, 1–15.
- [5] Woodall W.H. (2000). *Controversies and contradictions in statistical process control*. J. of Quality Technology **32**, 341–350.
- [6] Zhou W.H., Zhu G.L. (2007). *Economic design of integrated model of control chart and maintenance management*. Mathematical and Computer Modelling, doi:10.1016/j.mcm.2007.09.08 (in press).

Poděkování: Tato práce byla řešena v rámci projektu MŠMT 1M06047 Centrum pro jakost a spolehlivost výroby.

Adresa: G. Dohnal, ČVUT FSI, Ústav technické matematiky, 121 35 Praha 2, Karlovo náměstí 13

E-mail: dohnal@nipax.cz

ROBUST FILTERING AND PORTFOLIO MANAGEMENT

Petr Dostál

Keywords: Small transaction costs, HARA utility function, non-constant coefficients.

Abstract: We consider a non-consuming agent investing in a stock and a money market interested in the portfolio market price far in the future. We derive an almost log-optimal strategy for small proportional transaction costs in a model with stochastic coefficients. If the rate of return cannot be observed directly, we use robust filtering in order to obtain a model with observable stochastic coefficients. We also modify the obtained strategy for an investor who is more risk averse so that the strategy almost maximizes the asymptotics of expected HARA utility in case of deterministic coefficients.

Abstrakt: Uvažujme investora, který může investovat do akcie a ukládat finanční prostředky na účet. Pro tohoto investora odvodíme téměř optimální strategii vzhledem k logaritmické uživatelské funkci pro malé transakční náklady v modelu akcie se stochastickými koeficienty. Pro případ, že míra návratnosti akcie není přímo pozorovatelný proces, používáme techniku robustního filtrování založenou na Bayesově přístupu tak, abychom obdrželi model s pozorovatelnými koeficienty. V případě, že je investor více averzní vůči riziku, modifikujeme odvozenou strategii tak, že v případě konstantních deterministických koeficientů odpovídá snaze maximalizovat asymptotiku očekávaného užítka mocninných uživatelských funkcí.

1 Introduction

We consider an investment problem with proportional transaction costs without consumption. We seek for a strategy that maximizes the long run growth rate in case of logarithmic utility function and we modify the obtained method for non-logarithmic HARA utility functions. In case of constant coefficients our approach agrees with the aim to maximize the growth rate of the certainty equivalent of the investor's wealth similarly as in [4, 5] in contrast to [1, 2, 3, 6, 7, 8, 9, 10, 12], where the agent maximizes the expected value of the Laplace transform of the utility of consumption at a point chosen according to his/her time preferences. It is shown in [2] that the investment problem is a limiting case of the investment-consumption problem in case of logarithmic utility as the parameter of the Laplace transform goes to zero, i.e. when the consumption is postponed to the future as much as it is possible. In case of non-constant or rather non-deterministic coefficients and non-logarithmic HARA utility, our approach is generally different from the one based on maximizing of the asymptotics of expected utility of the portfolio market price,

but it is stable under reasonable time-change in model and the strategy up to a certain time depends only on the form of the model up to this time. We are searching for a strategy that is in comparison with a wide class of strategies and for small transaction taxes not worse than about some minimized error and therefore we can talk about almost optimal strategy in this sense. In case where the rate of return process cannot be observed directly, we use robust filtering in order to obtain a model, where the coefficients are known, but stochastic.

2 Basic notation and model set-up

Let \mathcal{F}_t be an augmented filtration representing the information flow of our investor on a certain complete probability space (Ω, \mathcal{A}, P) . Let us assume that the investor may invest in a stock market with the stock market price denoted by \mathcal{S}_t and in a money market with the interest rate $r_t \geq 0$. Then the value of money is given by $\mathcal{B}_t = \exp\{\int_0^t r_s ds\}$. Let \mathcal{S}_t be an \mathcal{F}_t -semimartingale with

$$d\mathcal{S}_t = \mathcal{S}_t[\alpha_t dt + \sigma_t dW_t], \quad \mathcal{S}_0 = s_0 > 0, \quad (1)$$

where W_t is a standard \mathcal{F}_t -Brownian motion and where $\theta_t := \sigma_t^{-2}(\alpha_t - r_t)$, $\ln \sigma_t$ are bounded continuous \mathcal{F}_t -Itô processes with bounded coefficients. The process σ_t is referred as the volatility of \mathcal{S} at time $t \geq 0$ and θ_t as the log-optimal proportion and it is the optimal proportion of investor's wealth invested in the stock in case of zero transaction costs and logarithmic utility, see section 3.

We denote by \mathcal{W}_t the portfolio market price, by \mathcal{N}_t the number of shares of the stocks in the portfolio and by π_t the corresponding investor's position in the market given by the equation

$$\mathcal{W}_t \pi_t = \mathcal{N}_t \mathcal{S}_t. \quad (2)$$

Hence, the position π_t is the proportion of investor's wealth invested in the stock. We assume that the deposit part is not discounted, see remark 1, and that we pay $(1+b)$ -multiple of the stock market price in order to obtain the stock and we obtain $(1-c)$ -multiple of the stock market price when we sell it. Then when we buy or sell the stock, the following value remains the same,

$$\mathcal{W}_t(1 + b\pi_t) = \mathcal{W}_t + b\mathcal{N}_t\mathcal{S}_t \quad \text{or} \quad \mathcal{W}_t(1 - c\pi_t) = \mathcal{W}_t - c\mathcal{N}_t\mathcal{S}_t,$$

respectively. In particular, the investor is able to withdraw from the market with positive portfolio market price after withdrawing if and only if the position $\pi_t \in (-1/b, 1/c)$. Further, we introduce the value $\lambda := \ln[(1+b)/(1-c)] \in (0, \infty)$, which describes the real level of transaction costs.

The initial problem was to maximize the asymptotics of expectation of $\mathcal{U}_\gamma(\mathcal{W}_t) = e_\gamma(\ln \mathcal{W}_t)$

$$\max \text{ asymptotics}_{t \rightarrow \infty} E e_\gamma(\ln \mathcal{W}_t), \quad (3)$$

where \mathcal{U}_γ are utility functions with hyperbolic absolute risk aversion (HARA)

$$\mathcal{U}_\gamma(y) = \frac{y^\gamma}{\gamma}, \quad e_\gamma(x) = \frac{1}{\gamma} \exp\{\gamma x\} \quad \text{if } \gamma < 0 \quad (4)$$

$$= \ln y, \quad e_\gamma(x) = x \quad \text{if } \gamma = 0. \quad (5)$$

In case where $\gamma = 0$, we are able to find a strategy that almost maximizes (3) for small transaction costs, and the same holds for $\gamma < 0$, provided that we replace t in (3) by certain stopping times $\tau(t) \rightarrow \infty$ as $t \rightarrow \infty$.

Our approach is based on our ability to find a complementary \mathbb{R}^k -valued process C_t summarizing the essential information, where $k \in \mathbb{N}$, a piecewise smooth function $f \in C^1(-1/b, 1/c) \times \mathbb{R}^k$ that balances the different infinitesimal increments of conditional expected utility of the portfolio market price corresponding to different values of the position π_t and the values of the process C_t , an \mathcal{F}_t -progressive rate process ν_t and a deterministic function δ_λ , which is small enough for $\lambda > 0$ small, such that

$$e_\gamma(\ln \mathcal{W}_t - f(\pi_t, C_t) - \int_0^t \nu_s \, ds - \delta_\lambda t)$$

is an \mathcal{F}_t -supermartingale when applying any of wide class of strategies and such that

$$e_\gamma(\ln \mathcal{W}_t - f(\pi_t, C_t) - \int_0^t \nu_s \, ds + \delta_\lambda t)$$

is an \mathcal{F}_t -submartingale if we consider the special one. In case of deterministic coefficients, ν_t can be understood as a value, which is close to the local growth rate of the certainty equivalent. In case of constant coefficients, the problem can be solved explicitly with zero error and without considering any complementary process. Then ν_t is just a constant and it corresponds to the rate of the exponential growth of the certainty equivalent of the investor's wealth. Note that the certainty equivalent of a random variable X is a constant giving the same expected utility as X .

Remark 1 In order to simplify our notation, we will further assume that the interest rate is zero. Otherwise, we would have to consider the discounted stock market price $\tilde{\mathcal{S}}_t := \mathcal{S}_t \mathcal{B}_t^{-1}$ instead of \mathcal{S}_t and the discounted wealth process $\tilde{\mathcal{W}}_t := \mathcal{W}_t \mathcal{B}_t^{-1}$ instead of \mathcal{W}_t (and $\tilde{\alpha}_t := \alpha_t - r_t$ instead of α_t). It is important to realize that $\tilde{\mathcal{S}}_t$ is an \mathcal{F}_t -Itô process satisfying (1) with $\tilde{\alpha}_t$ instead of α_t and that $\pi_t = \mathcal{N}_t \mathcal{S}_t \mathcal{W}_t^{-1} = \mathcal{N}_t \tilde{\mathcal{S}}_t \tilde{\mathcal{W}}_t^{-1}$. So, the simplification is writing $\mathcal{W}_t, \mathcal{S}_t, \alpha_t$ and writing about wealth, stock market price and rate of return instead of using more general and accurate notation $\tilde{\mathcal{W}}_t, \tilde{\mathcal{S}}_t, \tilde{\alpha}_t$.

3 Zero transaction costs

The purpose of this section is to illustrate our approach in case of zero transaction costs. It is reasonable to consider only such strategies that π_t is a bounded \mathcal{F}_t -progressive process. Then \mathcal{W}_t is an \mathcal{F}_t -Itô process with

$$\begin{aligned} d\mathcal{W}_t &= \mathcal{N}_t d\mathcal{S}_t = \mathcal{W}_t \pi_t [\alpha_t dt + \sigma_t dW_t], & \mathcal{W}_0 &= w_0 > 0, \quad \text{i.e.} \\ \ln \mathcal{W}_t &= \ln w_0 + \int_0^t \sigma_s \pi_s dW_s + (\alpha_s \pi_s - \frac{1}{2} \sigma_s^2 \pi_s^2) ds. \end{aligned}$$

We introduce log-optimal proportion $\theta_t := \sigma_t^{-2} \alpha_t$ as the value of π_t maximizing the logarithmic drift coefficient $\alpha_t \pi_t - \frac{1}{2} \sigma_t^2 \pi_t^2$ of \mathcal{W}_t . Further,

$$\mathcal{U}_\gamma(\mathcal{W}_t) = \exp\left\{\int_0^t \gamma \sigma_s \pi_s dW_s - \frac{\gamma^2}{2} \sigma_s^2 \pi_s^2 ds\right\} \cdot e_\gamma(\ln w_0 + \int_0^t u_{\gamma,s}(\pi_s) ds) \quad (6)$$

is a product of an exponential martingale and a process which is maximal if

$$\forall t \geq 0 \quad \pi_t = \Theta_t := \frac{\theta_t}{1-\gamma}, \quad \text{where} \quad u_{\gamma,t}(x) := \alpha_t x - \frac{1-\gamma}{2} \sigma_t^2 x^2. \quad (7)$$

In case of deterministic coefficients, it is easy to verify that the strategy given by (7) gives the maximal expected utility of the wealth process \mathcal{W}_t , as Θ_t is deterministic process and also the second factor on the right-hand side of (6) is. In case of non-deterministic coefficients and $\gamma < 0$, the strategy does not possess such a property, but it does provided that we are interested in the expected utility of the wealth \mathcal{W}_t at the hitting times of the following process

$$\tilde{\mu}_t := \int_0^t \tilde{\nu}_s ds, \quad \text{where} \quad \tilde{\nu}_s := \frac{1-\gamma}{2} \sigma_s^2 \Theta_s^2 = u_{\gamma,s}(\Theta_s), \quad (8)$$

if $\int_0^\infty \Theta_t^2 dt = \infty$ holds a.s. Put $\tilde{\tau}_h := \inf\{t \geq 0; \tilde{\mu}_t \geq h\}$. Then $E\mathcal{U}_\gamma(\mathcal{W}_{\tilde{\tau}_h}) = e_\gamma(\ln w_0 + h) = \mathcal{U}_\gamma(w_0 e^h)$ holds by Novikov theorem, and it means that the certainty equivalent of $\mathcal{W}_{\tilde{\tau}_h}$ is just $w_0 e^h$.

If we consider a time-changed model, the same approach leads to the same strategy up to the time-change. Note that everytime we write \mathcal{W}_t , we think of the discounted wealth $\tilde{\mathcal{W}}_t$, see remark 1.

4 Robust filtering

This section is devoted to the case when the rate of return cannot be observed directly. We use robust filtering, see [11, Chapter VI. part 3.], based on Bayesian approach in order to obtain the model (1). The problem is that generally, we do not observe the log-optimal proportion θ_t , α_t and W_t .

Let $(\Omega, \mathcal{G}, P, \mathcal{G}_t)$ be a complete filtered probability space with a complete filtration and assume that θ_t is bounded \mathcal{G}_t -Itô process with bounded coefficients such that (1) holds with $\alpha_t = r_t + \sigma_t^2 \theta_t$ and with W_t replaced by a standard \mathcal{G}_t -Brownian motion B_t . Let us assume that we observe just the stock market price \mathcal{S}_t and the process \mathcal{B}_t and denote by \mathcal{F}_t the completed filtration generated by these two processes. Note that \mathcal{F}_t can be also viewed as a complete filtration generated by the processes σ_t, \mathcal{B}_t and $V_t := B_t + \int_0^t \sigma_s \theta_s ds$ as $0 < \sigma_t$ is assumed to be a continuous process.

We are going to introduce a standard \mathcal{F}_t -Brownian motion W_t and a bounded \mathcal{F}_t -Itô process $\hat{\theta}_t$ with bounded coefficients playing the role of the log-optimal proportion such that (1) holds with α_t replaced by $\hat{\alpha}_t = r_t + \sigma_t^2 \hat{\theta}_t$.

Lemma 2 Let $\hat{\theta}_t$ be an \mathcal{F}_t -progressive process such that $\hat{\theta}_t = E[\theta_t | \mathcal{F}_t]$ holds almost surely for every $t \geq 0$ and let us assume that the process $\int_0^t \sigma_u |\theta_u| du$ is integrable. Then W_t given by (9) is a standard \mathcal{F}_t -Brownian motion, where

$$W_t := V_t - \int_0^t \sigma_s \hat{\theta}_s ds = B_t + \int_0^t \sigma_s (\theta_s - \hat{\theta}_s) ds. \quad (9)$$

Proof of lemma is simple and it based on Levy's characterization of a standard \mathcal{F}_t -Brownian motion \hat{W}_t among continuous \mathcal{F}_t -martingales.

Remark 3 Let us assume that $\hat{\theta}_t$ is as in lemma 2 and that (1) holds with W_t replaced by B_t with $\alpha_t = r_t + \sigma_t^2 \theta_t$. Then W_t defined by (9) is a standard \mathcal{F}_t -Brownian motion and

$$\mathcal{S}_t^{-1} d\mathcal{S}_t = \alpha_t dt + \sigma_t dB_t = \hat{\alpha}_t dt + \sigma_t dW_t$$

holds with $\hat{\alpha}_t = r_t + \sigma_t^2 \hat{\theta}_t$. In particular, $\hat{\theta}_t$ plays the role of the log-optimal proportion for the investor with the available information given by \mathcal{F}_t .

We are going to apply the Bayesian approach to the filtering and therefore we need a certain independent variable whose aposteriori distribution we will be searching for. We need also a second independent variable absorbing the observed information. Let us consider the following independent variables

$$\Xi : (\Omega, \mathcal{G}_0) \rightarrow (\mathcal{S}_\Xi, \mathcal{S}_\Xi), \quad \Upsilon : (\Omega, \mathcal{G}_0) \rightarrow (\mathcal{S}_\Upsilon, \mathcal{S}_\Upsilon). \quad (10)$$

Since they are both \mathcal{G}_0 -measurable, they are jointly independent with the standard \mathcal{G}_t -Brownian motion B_t . Further, we need to connect them with the processes that we are interested in and that we observe. We assume that

$$\theta_t = \vartheta_t(\Xi, V, \sigma, \mathcal{B}), \quad \sigma_t = \sigma_t(\Upsilon, V), \quad \mathcal{B}_t = \beta_t(\Upsilon, V) \quad (11)$$

hold for every $t \geq 0$, where ϑ and σ, β are continuous $\mathcal{S}_\Xi \otimes \mathcal{C}_t^3$ and $\mathcal{S}_\Upsilon \otimes \mathcal{C}_t$ -adapted processes, respectively, where \mathcal{C}_t denotes the canonical filtration on the space of all continuous functions on \mathbb{R}^+ denoted by \mathbb{C} . In order to abbreviate the notation, we will denote by \mathbb{V} the triplet of processes (V, σ, \mathcal{B}) and note that this triplet generates \mathcal{F}_t up to the completion.

Lemma 4 Let $\exp\{\frac{1}{2} \int_0^t \sigma_s^2 \theta_s^2 ds\}$ be an integrable process. Then the aposteriori density of Ξ w.r.t. the apriori distribution of Ξ given the history of \mathbb{V}_s up to $t \geq 0$ (i.e. given \mathcal{F}_t) is

$$f_{\Xi | \mathcal{F}_t}(\xi) := \mathcal{E}_t(\xi, \mathbb{V}) / \mathbb{E}_\xi[\mathcal{E}_t(\xi, \mathbb{V})] \quad (12)$$

$$= \exp \left\{ \int_0^t \sigma_u [\vartheta_u(\xi, \mathbb{V}) - \hat{\theta}_u] dW_u - \frac{\sigma_u^2}{2} [\vartheta_u(\xi, \mathbb{V}) - \hat{\theta}_u]^2 du \right\}, \quad (13)$$

where $\mathbb{E}_\xi[\mathcal{E}_t(\xi, \mathbb{V})] := \int \mathcal{E}_t(\xi, \mathbb{V}) \mathcal{L}(\Xi)(d\xi)$ and where

$$\mathcal{E}_t(\xi, \mathbb{V}) := \exp \left\{ \int_0^t \sigma_u \vartheta_u(\xi, \mathbb{V}) dV_u - \frac{1}{2} \sigma_u^2 \vartheta_u^2(\xi, \mathbb{V}) du \right\}. \quad (14)$$

The idea of the proof is the Bayesian approach, and the corresponding technical tools are Itô formula together with stochastic Fubini theorem.

Note that $\mathcal{E}_t(\xi, \mathbb{V})$ is just the wealth process of the investor with the initial wealth 1 paying no transaction costs and applying the corresponding log-optimal strategy, see remark 3, who thinks that $\Xi = \xi$. Then $\mathbb{E}_\xi \mathcal{E}_t(\xi, \mathbb{V})$ is the wealth process of the investor with the initial wealth 1 who distributes his/her initial wealth among the above mentioned investors with the distribution $\mathcal{L}(\Xi)$. Then $f_{\Xi|\mathcal{F}_t}(\xi)$ is the density of the wealth of such an investor of the current distribution of his/her wealth w.r.t. the initial one. In particular, one can regard the distribution of the wealth after a long time of investment as the best information about the investors' knowledge of the market provided that they do not consume and they cumulate their wealth in order to achieve the best ergodic result in case of no transaction costs.

We need to show that $\hat{\theta}_t$ is a bounded \mathcal{F}_t -Itô process with bounded coefficients under certain assumptions and to develop its dynamics.

Lemma 5 Let a_t, b_t be bounded $\mathcal{S}_\Xi \otimes \mathcal{C}_t^3$ -progressive coefficients and

$$d\vartheta_t(\xi, \mathbb{V}) = a_t(\xi, \mathbb{V}) dt + b_t(\xi, \mathbb{V}) dV_t. \quad (15)$$

Further assume that $\theta_t, \sigma_t, \mathcal{B}_t$ satisfy (11) with ϑ, σ bounded. Then the dynamics of $\hat{\theta}_t := \mathbb{E}_\xi [\vartheta_t(\xi, \mathbb{V}) f_{\Xi|\mathcal{F}_t}(\xi)] = E[\theta_t|\mathcal{F}_t]$ is given by

$$d\hat{\theta}_t = \sigma_t \widehat{\text{var}}_t(\theta) dW_t + \hat{a}_t dt + \hat{b}_t dV_t + \sigma_t \widehat{\text{cov}}_t(\theta, b) dt, \quad \text{where} \quad (16)$$

$$\hat{a}_t := \mathbb{E}_\xi [a_t(\xi, \mathbb{V}) f_{\Xi|\mathcal{F}_t}(\xi)] = E[a_t(\Xi, \mathbb{V})|\mathcal{F}_t], \quad (17)$$

$$\hat{b}_t := \mathbb{E}_\xi [b_t(\xi, \mathbb{V}) f_{\Xi|\mathcal{F}_t}(\xi)] = E[b_t(\Xi, \mathbb{V})|\mathcal{F}_t], \quad (18)$$

$$\widehat{\text{var}}_t(\theta) := \mathbb{E}_\xi \{\vartheta_t(\xi, \mathbb{V}) [\vartheta_t(\xi, \mathbb{V}) - \hat{\theta}_t] f_{\Xi|\mathcal{F}_t}(\xi)\} = \text{var} [\theta_t|\mathcal{F}_t], \quad (19)$$

$$\widehat{\text{cov}}_t(\theta, b) := \mathbb{E}_\xi \{b_t(\xi, \mathbb{V}) [\vartheta_t(\xi, \mathbb{V}) - \hat{\theta}_t] f_{\Xi|\mathcal{F}_t}(\xi)\} = \text{cov} [\theta_t, b_t(\Xi, \mathbb{V})|\mathcal{F}_t]. \quad (20)$$

In particular, if ϑ_t is a Lipschitz process, i.e. $b_t = 0$, we have that

$$d\hat{\theta}_t = \sigma_t \widehat{\text{var}}_t(\theta) dW_t + \hat{a}_t dt. \quad (21)$$

Lemma 5 can be obtained with the help of Itô formula and stochastic Fubini theorem.

5 Almost optimal strategy

Let us denote $[\hat{\alpha}_t(\bar{\alpha}_t, \bar{\beta}_t)\hat{\beta}_t]$ the strategy that buys the stock whenever the position π_t reaches the value $\hat{\alpha}_t$ (or is lower) so that the value of the position jumps onto the value $\bar{\alpha}_t$ and similarly that sells the stock whenever the position reaches the value $\hat{\beta}_t$ (or higher) so that the value of the position jumps onto the value $\bar{\beta}_t$.

The construction of the almost optimal strategy is technical so that we can prove its properties without additional technical difficulties. In order to avoid technical details we only write that $\hat{\alpha}_t = \hat{\Theta}_t - \omega_t$, $\hat{\beta}_t = \hat{\Theta}_t + \omega_t$ and that

$$\bar{\alpha}_t = \hat{\Theta}_t - \omega_t + \tau\omega_t^2 + o(\omega_t^2), \quad \bar{\beta}_t = \hat{\Theta}_t + \omega_t - \tau\omega_t^2 + o(\omega_t^2),$$

where $\tau > 0$ is a parameter. In order to be able to prove that a certain strategy is almost optimal, we have chosen

$$\omega_t = \sqrt[3]{\frac{3\lambda \left[\sigma_t \Theta_t (1 - \Theta_t) - \frac{d\langle \Theta, W \rangle_t}{dt} \right]^2 + \frac{d\langle \Theta \rangle_t}{dt} - \left[\frac{d\langle \Theta, W \rangle_t}{dt} \right]^2 + \sigma_t^2 \lambda^a}{4 \sigma_t^2 (1 - \gamma)}},$$

where we consider $a = 1$ if the numerator of the second fraction is bounded away from zero. If this is not the case, we put $a = 2/7$. If $a = 1$, we put $q = 1$ and if $a = 2/7$ then $q = 6/7$.

In the logarithmic case, i.e. if $\gamma = 0$ we are able to show that

$$E \ln \hat{\mathcal{W}}_\rho - E \ln \mathcal{W}_\rho \geq O(\lambda) + O(\lambda^q) E \rho$$

whenever ρ is a bounded \mathcal{F}_t -stopping time, where $\hat{\mathcal{W}}$ and \mathcal{W} stand for the wealth processes corresponding to the special strategy and to a strategy from a wide class of admissible strategies, respectively.

If $\gamma < 0$, we are only able to find $\varpi_\lambda = O(\lambda^q)$ and a system of \mathcal{F}_t -stopping times $\tau_h := \inf\{t \geq 0, \int_0^t \nu_s ds + \varpi_\lambda t \geq h\}$ such that

$$\limsup_{\lambda \rightarrow 0^+} \frac{1}{\lambda^q} \limsup_{h \rightarrow \infty} \frac{e_\gamma^{-1} E e_\gamma(\ln \mathcal{W}_{\tau_h}) - e_\gamma^{-1} E e_\gamma(\ln \hat{\mathcal{W}}_{\tau_h})}{h} < \infty$$

holds with ν_t defined below provided that ν_t is a non-negative process.

$$\nu_t := \frac{1-\gamma}{2} \sigma_t^2 (\Theta_t^2 - \omega_t^2) + \frac{r}{2} \Lambda^2 \mathbf{1}_{[|\Theta_t| < \Lambda]} d\langle \theta \rangle_t / dt,$$

where $\Lambda := (\frac{3\lambda}{4L})^{1/3}$ and L is a certain constant, which is not specified here. Further, if there is a continuous function g positive at zero such that $d\langle \theta \rangle_t \geq g(\theta_t) dt$, we are able to show that $\nu_t \geq 0$ holds for $r > 0$ is large enough. We only remark that the constant L is such that $0 < \omega_t \leq \Lambda$ holds for every $t \geq 0$.

References

- [1] Akian M., Menaldi J.L., Sulem A. (1996). *On an investment-consumption model with transaction costs*. SIAM J. Control Optim. **34(1)**, 329–364.
- [2] Akian M., Sulem A., Taksar M.I. (2001). *Dynamic optimization of long-term growth rate for a portfolio with transaction costs and logarithmic utility*. Mathematical Finance **11**, 153–188.

- [3] Davis M., Norman A. (1990). *Portfolio selection with transaction costs*. Math. Oper. Res. **15**, 676–713.
- [4] Dostál P. (2006). *Optimal trading strategies with transaction costs paid only for the first stock*. Acta Universitatis Carolinae, Mathematica & Physica **47(2)**, 43–72.
- [5] Dostál P. *Investment strategies in the long run with proportional transaction costs and HARA utility function*. to appear in Quantitative Finance.
- [6] Janeček K., Shreve S.E. (2004). *Asymptotic analysis for optimal investment and consumption with transaction costs*. Finance and Stochastics **8**, 181–206.
- [7] Janeček K., Shreve S.E. *Futures trading with transaction costs*. submitted.
- [8] Magill M.J.P., Constantinides G.M. (1976). *Portfolio selection with transaction costs*. J. Econ. Theory **13**, 245–263.
- [9] Merton R.C. (1971). *Optimum consumption and portfolio rules in a continuous-time model*. J. Econ. Theory **3**, 373–413 [Erratum **6**, 213–214 (1973)].
- [10] Morton A.J., Pliska S. (1995). *Optimal portfolio management with fixed transaction costs*. Math. Finance **5(4)**, 337–356.
- [11] Rogers L.C.G., Williams D. (2000). *Diffusions, Markov processes and martingales*. (2nd Edition) Cambridge Cambridge University Press.
- [12] Shreve S., Soner H.M. (1994). *Optimal investment and consumption with transaction costs*. Ann. Applied Probab. **4**, 609–692.

Acknowledgement: The work is partly supported by GAČR 201/04/P257 and a part of the research project MSM 0021620839 financed by MSMT.

Address: P. Dostál, MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8 – Karlín

E-mail: petr.dostal@mff.cuni.cz

GENERALIZED ZERO RANGE PROCESS AS A TRAFFIC MODEL

Lucie Fajfrová

Keywords: Zero range process, attractiveness, invariant measures.

Abstract: In the paper, a model of conservative particle system, which generalises a well known zero range process, is studied. The generalisation consists in allowing jumps of more than one particle in one moment. We describe what this generalisation means in the context of modeling a traffic flow.

Abstrakt: Příspěvek je věnován konzervativnímu částicovému systému, který zobecňuje známý částicový systém zvaný Zero-range proces. V tomto zobecnění povolíme přeskok více než jedné částice v jednom kroku. Proces použijeme pro modelování dopravního toku, kde má takové zobecnění velmi dobrý význam.

1 Introduction

We are interested in particle systems with zero range or exclusion interactions [1, 8]. A particle system is a large system of indistinguishable particles which occupy sites of a graph, typically lattice \mathbb{Z}^d or for example binary trees [4]. The evolution of such a system in time is given by movement of particles between sites and this movement is influenced by interactions among particles. In case of zero range and exclusion interactions, the whole system is conservative, i.e. there are no births or deaths. Exclusion models allow at most one particle per site and interactions between two particles arise if one particle attempts to jump to the site occupied by the second one. Such jumps are excluded. Zero range models allow arbitrary number of particles per site and interactions arise just among particles at the same site, the particles are queueing. These systems are models for queuing systems, networks or traffic. From the theoretical point of view they are Markov processes with continuous time with a state space given by allowable particle configurations.

Particularly, the one dimensional ($d = 1$) exclusion process can be used as a simple model for a traffic flow. Each particle (a car) can move forward only if the site ahead is unoccupied. If zero range models are employed we gain wider possibilities how to describe different speeds of columns of various lengths. Both models were used as models for a traffic flow in [7].

In this contribution, we focus on a generalisation of zero range processes with allowance of multiple jumps. This generalisation would permit to model, in the traffic context, splitting of columns. Our aim is to study stationary states of this process and also a closely connected problem of the existence of a coupling process for the given dynamics.

Previously, stationary states for zero range dynamics generalised by multiple jumps were studied by Evans [2, 3] in the framework of finite volume systems, i.e. particles live only on a finite interval $\{1, \dots, L\} \subset \mathbb{Z}$.

In a special case, when jumps of particles have a constant rate - independent of the number of particles at sites and also independent of the amount of jumping particles, a multiple jump model was treated in [9]. The model from paper [9], called *stick process* was studied in the context with Ulam's problem, i.e. evaluation of a limit theorem for the longest increasing subsequence of a random permutation of n symbols.

2 Definition of model

Let us consider a particle system living on one dimensional lattice \mathbb{Z} , each site $i \in \mathbb{Z}$ is occupied by an arbitrary number $\eta(i)$ of particles and the whole particle configuration $\eta = (\eta(i) : i \in \mathbb{Z})$ is a state of our process $(\eta_t)_{t \geq 0}$. We suppose that the particle system is evolving in time in the sense given by series of independent Poisson processes, where each of them rules movement of particles at just one site i , and exactly events of these Poisson processes are only possible moments for jumps of particles from one site to another. We will consider here the totally asymmetric case, it means that jumps are possible only between neighbours on \mathbb{Z} and only in one direction (let us say from the left to the right). In classical zero range processes, the jump of at most one particle in one moment is possible. In generalised - *multiple jumps* - zero range processes (MJ-ZRP), the following jumps may occur in time t of a Poisson event at i :

k -many particles from total amount $\eta_{t-}(i)$ leaves site i and moves to site $i + 1$.

The rate of this jump is equal to

$$g(k, \eta_{t-}(i)) \tag{1}$$

where g is a nonnegative function on $\mathbb{N} \times \mathbb{N}$, $g(k, \alpha) = 0$ if $k > \alpha$.

The classical zero range process (ZRP) is then a special case with

$$g(k, \alpha) = I_{[k=1]}r(\alpha) \tag{2}$$

for a nonnegative function r on \mathbb{N} , $r(0) = 0$. So only one particle can jump in time t of a Poisson event at i and a rate of the jump is $r(\eta_{t-}(i))$. A meaning of the rate is following: looking at site i in time s we will wait for a jump of one particle from i an exponential time with mean $1/r(\eta_s(i))$. We can imagine an exponential clock at each site instead of the Poisson process. Note that always when a jump happens some of exponential clocks have to change their means according to new configuration η .

A totally asymmetric *multiple jumps* - *zero range process* on \mathbb{Z} can be defined as a Markov process $(\eta_t)_{t \geq 0}$ with a state space $X \subset \mathbb{N}^{\mathbb{Z}}$ given by infinitesimal generator

$$(\mathbb{L}f)(\eta) = \sum_{i \in \mathbb{Z}} \sum_{k \in \mathbb{N}^+} g(k, \eta(i)) (f(\eta^{k,i}) - f(\eta)) \quad (3)$$

for every $\eta \in X$ and every cylinder function $f : X \rightarrow \mathbb{R}$.

Notation and remark on definition:

1. We denoted by $\eta^{k,i}$ a changed configuration after a multiple jump

$$\eta^{k,i}(j) = \begin{cases} \eta(i) - k & \text{if } j = i \\ \eta(i+1) + k & \text{if } j = i+1 \\ \eta(j) & \text{otherwise.} \end{cases}$$

2. A function $f : \mathbb{N}^{\mathbb{Z}} \rightarrow \mathbb{R}$ is called *cylinder function*, if there exists a finite $K \subset \mathbb{N}$ such that $f(\eta)$ depends only on $\eta|_K = (\eta(j) : j \in K)$ for each $\eta \in \mathbb{N}^{\mathbb{Z}}$.

3. The reason, why we can not consider the whole product space $\mathbb{N}^{\mathbb{Z}}$ as a state space, is that even if we assume bounded function g there are particle configurations that could induce arrival of infinitely many particles to the same site in a finite time. This is the case of configurations with too many particles laying on the left halfline near $-\infty$. Hence we allow only following set of configuration $X = \{\eta \in \mathbb{N}^{\mathbb{Z}} : \lim_{n \rightarrow -\infty} \frac{1}{n^2} \sum_{i=n}^{-1} \eta(i) = 0\}$, see [9].

Let us state here for completeness a generator of the classical zero range process (totally asymmetric):

$$(\mathbb{L}_c f)(\eta) = \sum_{i \in \mathbb{Z}} r(\eta(i)) (f(\eta^{1,i}) - f(\eta)) \quad (4)$$

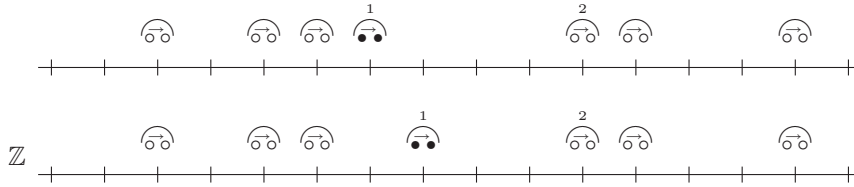
for every $\eta \in \mathbb{N}^{\mathbb{Z}}$ and every cylinder function $f : \mathbb{N}^{\mathbb{Z}} \rightarrow \mathbb{R}$. No restriction on the state space are needed if we assume r bounded. For a case of unbounded r , see [1].

3 Traffic interpretation

In the framework of *finite volume*, zero range process was used as a model for traffic flow in work [7]. Particle systems in finite volume consider only finite configurations of particles living on an interval $[1, \dots, L]$ instead of \mathbb{Z} , usually with periodic boundary condition (site 1 is assumed as the neighbour on the right from site L). In this paper, we are interested in models in infinite volume, defined in (3).

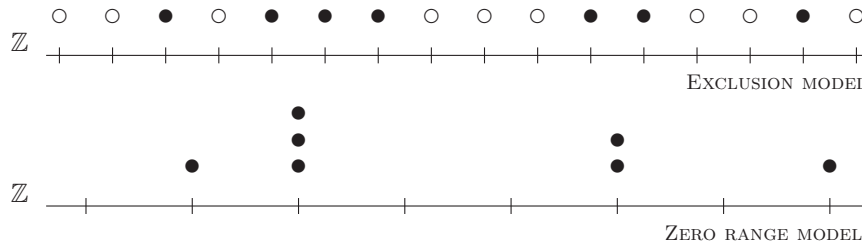
Let us interpret particles as cars and one dimensional lattice \mathbb{Z} as a road on which cars move in one direction (to the right, by our definition). So each $i \in \mathbb{Z}$ is either occupied by a car or is empty. The car at site i can move to site $i+1$ only if site $i+1$ is empty. If there is an interval $[i, i+1, \dots, j]$ of sites which are occupied by cars we will say that there is a column (convoy) of cars. This simple model can be described by a totally asymmetric exclusion process, which refers to evolution of a configuration $(\eta(i) : i \in \mathbb{Z})$, where

$\eta(i) = 1$ if site i is occupied or $\eta(i) = 0$ if it is empty. A jump of a particle from site i to $i + 1$ has rate $\mathbb{I}_{[\eta(i)=1, \eta(i+1)=0]}$, i.e. a constant rate of jumps from an occupied site to its empty right neighbour.



Therefore, if cars have empty space ahead of them they move fluently with constant rate 1 (car 1 in the picture above). If, on the other hand, the traffic is heavy then next site could be occupied and then the car has to wait (car 2 in the picture above).

But we would like if our traffic model covers the fact that the speed of a single car, a short column and a long columns differs. From this reason we use instead of the simple exclusion process the classical zero range process (4). Note that the models can be transformed one by one in a simple way demonstrated on a following picture.



Now the first car of column i moves with a rate (speed) $r(\eta(i))$ where $\eta(i)$ is a number of cars of i -th column. In this context, it is natural to assume that r is nonincreasing. Since ZRP with decreasing rate function r is not attractive (attractiveness is defined in the next section), a lot of works on ZRP excluded this case. ZRP with decreasing r was studied in details in [6].

The generalisation of zero range process defined by (3) allows to consider situations that k particles of $\eta(i)$ move together in one moment, which means that a column of length k breaks away from i -th columns with a rate $g(k, \eta(i))$.

4 Coupling process

In this section, we investigate the problem of attractiveness of MJ-ZRP (3). Since the attractiveness is closely related to ordering of particle configurations, let us start with some notation.

Let us consider the following partial order on state space $X \subset \mathbb{N}^{\mathbb{Z}}$: $\eta \leq \zeta$ if and only if $\eta(i) \leq \zeta(i)$ for every $i \in \mathbb{Z}$, $\eta, \zeta \in X$. We say that a function f on X is *monotone* if $f(\eta) \leq f(\zeta)$ whenever $\eta \leq \zeta$.

We say that a particle system is *attractive* if for every bounded, monotone continuous function f and every time $t > 0$, function $S_t f$ is again bounded, monotone continuous function, where $(S_t f)(\xi)$ for each $\xi \in X$ is the expected value of $f(\eta_t)$ under assumption that the process $(\eta_t)_{t \geq 0}$ have started from configuration ξ at time 0. In another words, an attractive system preserves ordering in the following sense. If we imagine two initial configurations η, ζ which are ordered, $\eta \leq \zeta$, and then we let them evolve by the same dynamics and control both of them by the same set of Poisson processes (same realisation of exponential clocks), the configurations of these systems in arbitrary time will be again ordered, $\eta_t \leq \zeta_t$.

The technique, we have just now used to explain what attractiveness means, is called *coupling technique*. It is very usual tool and, specially, the construction of a *coupling process* (i.e. particle system (η_t^1, η_t^2) on $X \times X$ where each its marginal η_t^i is the original process) which preserves ordering of its marginals (i.e. $\eta_0^1 \leq \eta_0^2$ implies $\eta_t^1 \leq \eta_t^2$ for every $t > 0$) acts as a proof of attractiveness. However it is not the aim of this paper to present results about coupling process. We focus only on conditions on rate function g under which the process is or is not attractive. A thorough characterisation of coupling rates for a big class of conservative process, MJ-ZRP among them, can be found in [5].

The attractiveness is an important property and a lot of results on particle systems were proved only using the fact that a process is attractive. On the other hand, if a process is not attractive then it could have some special properties. At the end of the previous section, we said that using ZRP as a traffic model it is natural to assume nonincreasing speed function r . And it is well known fact, see e.g. [1], that in this case (except the constant speed function) ZRP is not attractive. In [6], cases of decreasing speed function r are studied for which so called *condensation* occurs. They showed that there exists a critical density of particles and if we start with a configuration with a supercritical density, a mass of particles will concentrate at one single site. If we interpret this condensation phenomena in the traffic context we obtain a critical density of traffic flow beyond which traffic jam occurs.

In the following proposition we examine single cases of rate function g of MJ-ZRP and state conditions on g so that the process is attractive.

Proposition 4.1. Let us consider MJ-ZRP given by generator (3). We distinguish following cases:

$$g(k, \alpha) = h(k) \mathbb{I}_{[k \leq \alpha]} \quad (5)$$

$$g(k, \alpha) = r(\alpha) \mathbb{I}_{[k \leq \alpha]} \quad (6)$$

$$g(k, \alpha) = r^*(\alpha) r^*(\alpha - 1) \dots r^*(\alpha - k + 1) \mathbb{I}_{[k \leq \alpha]}. \quad (7)$$

for some functions h , r and r^* .

The process is attractive if and only if following conditions on function h , r and r^* , respectively, are satisfied:

- $h(k)$ nonincreasing for $k \geq 1$, nonnegative,
- $r(\alpha)$ nonincreasing & $\alpha r(\alpha)$ nondecreasing for $\alpha \geq 1$, nonnegative, and
- $0 < r^*(\alpha) \leq 1$ & $\sum_{i=1}^k r^*(k) \dots r^*(i) \leq \frac{r^*(\alpha+1)}{1-r^*(\alpha+1)}$ for every $1 \leq k \leq \alpha$,

respectively.

This result is a simple consequence of a general theorem presented in [5] where necessary and sufficient conditions for a general class of conservative particle systems are proved. Note that a sufficient condition on r^* to have the process attractive is to assume r^* positive, bounded by 1 and nondecreasing.

An interesting problem now could be to study the condensation phenomena for generalised zero range process (3) with such rate function g for which we have obtained by Proposition 4.1 the process is not attractive. To study this problem we need to know how invariant measures in mentioned cases of the rate function look like.

5 Invariant measures

This part is devoted to the study of invariant measures of MJ-ZRP (3). Invariant measures are stationary states of a Markov process which means: if the distribution of the initial particle configuration is just a stationary distribution then looking at the system in arbitrary fixed time we can observe again the same distribution as at the start.

A measure μ on $X \subset \mathbb{N}^{\mathbb{Z}}$ is called invariant measure for the process with generator \mathbb{L} if

$$\int \mathbb{L}f \, d\mu = 0 \quad (8)$$

for every cylinder function $f : X \rightarrow \mathbb{R}$.

Note that (8) is other, more useful formulation of a standard formula for invariant measures

$$\int (S_t f)(\xi) \, \mu(d\xi) = \int f(\xi) \, \mu(d\xi), \text{ for every continuous, bounded } f,$$

since generator \mathbb{L} is given explicitly for a majority of particle systems. In case of particle systems there is usually large set of invariant measures indexed by a real parameter. An exhaustive description of all invariant measures for given dynamics is then to set exactly how \mathcal{I}_e , the set of extremal invariant measures, looks like. Since the problem of finding \mathcal{I}_e is nontrivial, a complete description of \mathcal{I}_e is known only for basic models. Sometimes, a simpler task is to look for invariant measures which are in addition translation invariant,

$$\text{i.e. } \mu(\{\eta : \eta(i) \in A_i, i \in I\}) = \mu(\{\eta : \eta(i+j) \in A_i, i \in I\}) \text{ for each } j \in \mathbb{Z}.$$

A result on invariant measures for MJ-ZRP we want to present is following.

Theorem 5.1. Let us consider MJ-ZRP given by generator (3). If

$$g(k, \alpha) = \frac{h(k)r(\alpha)}{r(\alpha - k)} \text{ for every } 1 \leq k \leq \alpha, \alpha \geq 1, \quad (9)$$

for some h nonnegative, r positive function, then product measures ν^φ on X ,

$$\nu^\varphi(\{\eta : \eta(i) = n\}) = Z_\varphi \frac{\varphi^n}{r(n)} \text{ for every } i \in \mathbb{Z}, n \in \mathbb{N}, \quad (10)$$

with a normalising constant Z_φ , are invariant for (3) for every $\varphi \in (0, c)$.

Let us discuss now the particular cases which we studied in the previous section. We can see that cases (5) and (7) are covered by Theorem 5.1 and product invariant measures ν^φ have the following form:

$$\begin{aligned} \nu^\varphi(\eta(i) = n) &= Z_\varphi \varphi^n / (r^*(n)r^*(n-1)\dots r^*(1)), & \text{in case (7),} \\ \nu^\varphi(\eta(i) = n) &= \varphi^n (1 - \varphi), & \text{in case (5).} \end{aligned}$$

Hence invariant measures ν^φ in case (5) do not depend on the exact form of function h . Note that the same measures are also invariant for a classical (totally asymmetric) ZRP (4) with a constant rate function $r(\alpha) = I_{[\alpha > 0]}$.

The situation in case (6) is no more analogical to classical ZRP (4) which has product invariant measures, $\nu^\varphi(\eta(i) = n) = Z_\varphi \frac{\varphi^n}{r(n)r(n-1)\dots r(1)}$, see e.g. [1]. In case (6) of MJ-ZRP, invariant and translation invariant measures are no more product ones.

Theorem 5.2. MJ-ZRP with rate function $g(k, \alpha) = r(\alpha)$, case (6), has no product translation invariant measures which are invariant with respect to (4), except a trivial case when r is constant.

A similar problem was treated in [3] in the framework of finite volume particle systems. They proved that finite volume MJ-ZRP has product invariant measures if and only if rate function g can be formulated as in (9). So Theorem 5.1 generalises one implication of this result for infinite volume MJ-ZRP and Theorem 5.2 generalises the second implication for one particular case of infinite volume MJ-ZRP.

Sketch of proof of Theorem 5.1. One can show that following condition on rate function g :

$$g(k, k+l) \frac{\mu(\eta(i) = k+l)}{\mu(\eta(i) = l)} = c(k), \text{ for each } k \geq 1,$$

where c is a function independent of $l \geq 0, i \in \mathbb{Z}$, is sufficient for (8) to be satisfied for a product measure μ on X . The proof is done, if we verify the above conditions with g as given in (9) and $\mu = \nu^\varphi$ defined in (10).

Sketch of proof of Theorem 5.2. We find necessary conditions on a product, translation invariant measure μ on X for (8) to be satisfied. Let us denote $\pi(n) = \mu(\eta(i) = n)$. Then for arbitrary $n \in \mathbb{N}$, $(l_1, \dots, l_n) \in \mathbb{N}^n$

$$\begin{aligned}
0 &= \sum_{i=2}^n \sum_{k=1}^{l_i} \left(r(l_{i-1} + k) \frac{\pi(l_{i-1} + k)}{\pi(l_{i-1})} \frac{\pi(l_i - k)}{\pi(l_i)} - r(l_i) \right) + \\
&+ \sum_{k=1}^{l_1} \sum_{m \geq 0} \left(r(m + k) \frac{\pi(m + k)}{\pi(m)} \frac{\pi(l_1 - k)}{\pi(l_1)} - r(l_1) \right) \pi(m) + \\
&+ \sum_{k \geq 1} \left(r(l_n + k) \frac{\pi(l_n + k)}{\pi(l_n)} - \sum_{m \geq 0} r(m + k) \pi(m + k) \right)
\end{aligned}$$

are desired necessary conditions. One can show that these conditions does not hold if r is not a constant function.

References

- [1] Andjel E.D. (1982). *Invariant measures for the zero range process*. Ann. Probab. **10**, 525–547.
- [2] Evans M.R., Hanney T. (2005). *Nonequilibrium statistical mechanics of the zero-range process and related models*. J. Phys. A: Math. Gen. **38**, R195.
- [3] Evans M.R., Majumdar S.N., Zia R.K.P. (2004). *Factorized steady states in mass transport models*. J. Phys. A: Math. Gen. **37**, L275–L280.
- [4] Fajfrová L. (2006). *Infinite queueing system with tree structure*. Kybernetika **42**, 5, 585–604.
- [5] Gobron T., Saada E. (2008). *Attractiveness and coupling for conservative particle systems*. Preprint 2008. <http://arxiv.org/abs/0903.0316v1>
- [6] Großkinsky S., Schütz G.M., Spohn H. (2003). *Condensation in the zero range process: stationary and dynamical properties*. Journal of Statistical Physics **113**, 389–410.
- [7] Kaupuzs J., Mahnke R., Harris R.J. (2005). *Zero-range model of traffic flow*. Physical Review E **72**, 056125.
- [8] Liggett T.M. (1999). *Stochastic interacting systems: contact, voter and exclusion process*. Springer-Verlag, Berlin.
- [9] Seppäläinen T. (1996). *A microscopic model for the Burgers equation and longest increasing subsequences*. Electronic J. Probab. **1**, Paper 5, 1–51.

Acknowledgement: This work was supported by the Grant Agency of the Czech Republic under Grant No. 201/06/1323 and MŠMT ČR 1M0572.

Address: L. Fajfrová, Institute of Information Theory and Automation of AS CR, Pod Vodárenskou věží 4, 182 08 Prague 8, Czech Republic

E-mail: fajfrova@utia.cas.cz

KOZIOLŮV-GREENŮV MODEL SE ZLEVA USEKNUTÝMI POZOROVÁNÍMI

Michal Friesl

Klíčová slova: Náhodné cenzorování, Koziolův-Greenův model, neparametrické bayesovské odhady, gama proces, Channing House Data.

Abstrakt: Příspěvek pojednává o neparametrickém bayesovském odhadu funkce spolehlivosti v modelu s proporcionálním cenzorováním. Protože standardně se nepředpokládá, že by rozdělení cenzoru mohlo souviset s rozdělením doby života, musíme tvar neparametrických bayesovských odhadů pro tuto situaci upravit. Na minulém Robustu jsme se zabývali odhady z dat pouze cenzorovaných, tentokrát přidáme useknutí zleva.

Abstract: The paper deals with nonparametric Bayes estimator of the reliability function in a model with proportional censorship and left truncation. The paper complements our Robust'06 paper where no truncation was assumed.

1 Úvod

Uvažujme dobu života — nezápornou náhodnou veličinu X , jejíž pozorování může být zprava cenzorováno náhodným časem Y . To znamená, že známe jen dobu ukončení pozorování a k tomu víme, zda jsme pozorovali sledovanou veličinu X , či pouze cenzor,

$$Z = X \wedge Y \quad \text{a} \quad I = I_{[X \leq Y]}.$$

Zároveň jsou pozorování useknutá zleva, tedy pozorování neprobíhá od času 0, ale začíná až v čase A . O případech s $X < A$ nebo $Y < A$ se tak nedozvíme, fakticky pozorování pochází z useknutého rozdělení $(X, Y | X, Y > A)$. Předpokládá se nezávislost sledované veličiny X a cenzorující veličiny Y (a času useknutí A).

Cílem je z podmíněných pozorování veličin (A, Z, I) odhadnout nepodmíněné charakteristiky doby života X , např. funkci spolehlivosti

$$S(x) = 1 - F(x) = P(X > x), \quad x > 0,$$

kde F značí distribuční funkci doby X .

Tato úloha je i v neparametrické bayesovské statistice a pro různé druhy neúplných pozorování prozkoumána, viz např. [3]. V našem případě ovšem máme dodatečný Koziolův-Greenův předpoklad z [4] o proportionalitě cenzorování. Předpokládáme totiž, že funkce spolehlivosti sledované a cenzorující veličiny, S a S_Y jsou svázány vztahem

$$S_Y(t) = (S(t))^\gamma, \quad t > 0, \quad (1)$$

pro nějaké $\gamma > 0$. Proporcionalitou cenzorování je zde myšlena proporcionalita příslušných intenzit poruch. Označíme-li kumulativní intenzitu poruch veličiny X

$$\Lambda(t) = -\ln S(t) = \int_0^t dF(x)/S(x) dx$$

a podobně Λ_Y pro cenzor Y , můžeme předpoklad (1) ekvivalentně zapsat jako

$$\Lambda_Y(t) = \Lambda(t) \cdot \gamma, \quad t > 0.$$

Remark 1.1. *V případě spojitých rozdělání (což ovšem úplně přesně nebude náš případ) znamená tento předpoklad nezávislost indikátoru cenzorování I na pozorovaném čase, tj. pravděpodobnost, že pozorování (ukončené v čase Z) bude necenzorované, na čase Z nezávisí a je $p = P(I = 1) = 1/(1 + \gamma)$.*

Vlastnost proporcionality zůstává zachována i při useknutí zleva, stále máme díky nezávislosti X a Y i u funkcí spolehlivosti podmíněných rozdělání $S_Y(t | Z > a) = (S(t | Z > a))^\gamma$.

Víme-li, že rozdělání cenzoru souvisí s rozdělání doby života, je to z hlediska kvality odhadů výhodou, neboť např. i samotné pozorování cenzorujícího času Y poskytuje informaci o S ; na druhou stranu je třeba odvodit odhady, které tento předpoklad zohlední. V [1] jsme se zabývali modelem bez useknutí zleva a ilustrovali mimo jiné, jak se využití předpokladu projeví na snížení bayesovských rizik odhadů. V [5] jsou studovány vlastnosti neparametrického odhadu, který analogicky jako Abdushukurovův-Chengův-Linův (ACL) odhad vychází z neparametrického odhadu funkce spolehlivosti veličiny Z , zde ale počítaného z useknutých dat, viz (4).

2 Odhady

Hodláme najít neparametrický bayesovský odhad funkce spolehlivosti S , tzn. nepředpokládáme, že by bylo známo byť jen z jaké rodiny rozdělání sledované veličiny pochází. Parametrem, který při bayesovském přístupu chápeme jako náhodnou veličinu, je zde tedy nikoliv jedna hodnota (popř. hodnoty dvě, tři, . . .), nýbrž celá funkce spolehlivosti ($S(t), t > 0$) jako nekonečněrozměrný parametr.

Tradičním apriorním modelem pro náhodné rozdělání S jsou zprava neutrální procesy, či ekvivalentně, využijeme-li k formulaci kumulativní intenzity Λ , předpokládáme, že $(\Lambda(t), t > 0)$ je neklesající proces s nezávislými přírůstky. V dalším budeme počítat konkrétně s gama procesem, který u přírůstků procesu Λ předpokládá gama rozdělání

$$\Lambda(s, t) = \Lambda(t) - \Lambda(s) \sim G(n_0, n_0 \Lambda_0(s, t))$$

kde Λ_0 je nějaká pevně zvolená intenzita poruch; pro jednoduchost nechť Λ_0 je spojitá. Zde Λ_0 slouží jako apriorní střední hodnota, $E \Lambda(s, t) = \Lambda_0(s, t)$, zatímco parametr n_0 řídí rozptyl $\text{var} \Lambda(s, t) = \Lambda_0(s, t)/n_0$.

Remark 2.1. Nemůžeme čekat, že by platilo $ES(t) = e^{-\Lambda_0(t)}$, k této hodnotě se ovšem $ES(t) = E e^{-\Lambda(t)} = (1 + 1/n_0)^{-n_0 \Lambda_0(t)}$ blíží při $n_0 \rightarrow \infty$, kdy limitně apriorní rozdělení Λ je soustředěno do nenáhodné intenzity Λ_0 .

Pro koeficient proporcionality γ nepožadujeme žádné specifické apriorní rozdělení, necht' má hustotu $\pi(\gamma)$ vzhledem k nějaké míře μ na $(0, \infty)$. Předpokládáme ale, že parametr γ je nezávislý s Λ .

Necht' pozorovaná data spočívají v náhodném výběru rozsahu n z rozdělení trojice (A, Z, I) . K popisu aposteriorního rozdělení budeme potřebovat následující označení. Uspořádáním pozorovaných časů $A_1, Z_1, \dots, A_n, Z_n$ podle velikosti a po odstranění případných opakujících se hodnot dostaneme na časové ose posloupnost už navzájem různých časových okamžiků

$$0 = T_0 < T_1 < \dots < T_N < T_{N+1} = \infty, \quad (2)$$

kde jsme přidali případně $T_0 = 0$ a $T_{N+1} = \infty$. Pro každý okamžik T_j definujeme

$$\begin{aligned} U_j &= \# \text{ necenzorovaných pozorování se } Z_k = T_j, \\ C_j &= \# \text{ cenzorovaných pozorování se } Z_k = T_j, \\ N_j &= \# \text{ pozorování s } A_k \leq T_j \text{ a } Z_k > T_j \end{aligned}$$

Dále definujeme

$$\begin{aligned} M_j^m(\gamma) &= n_0 + N_j(1 + \gamma) + m, \quad j = 0, \dots, N, \quad \text{a} \\ c_j^m(\gamma) &= \sum_{k=0}^{U_j} \sum_{\ell=0}^{C_j} (-1)^{k+\ell} \binom{U_j}{k} \binom{C_j}{\ell} \ln \left(\frac{M_j^m(\gamma) + C_j}{M_j^m(\gamma) + C_j + k + \ell \gamma} \right), \quad (3) \\ q_j^m(\gamma) &= (M_{j-1}^m(\gamma))^{-n_0 \Lambda_0(T_{j-1}, T_j)} c_j^m(\gamma), \quad j = 1, \dots, N, \end{aligned}$$

$m = 0, 1$, přičemž píšeme $M_j(\gamma) = M_j^0(\gamma)$, $c_j(\gamma) = c_j^0(\gamma)$ a $q_j(\gamma) = q_j^0(\gamma)$.

Tvrzení 1. Pro aposteriorní rozdělení parametrů vychází:

1. při daném γ má kumulativní intenzita Λ nezávislé přírůstky

– na intervalech mezi body dělení (2) s gama rozdělením

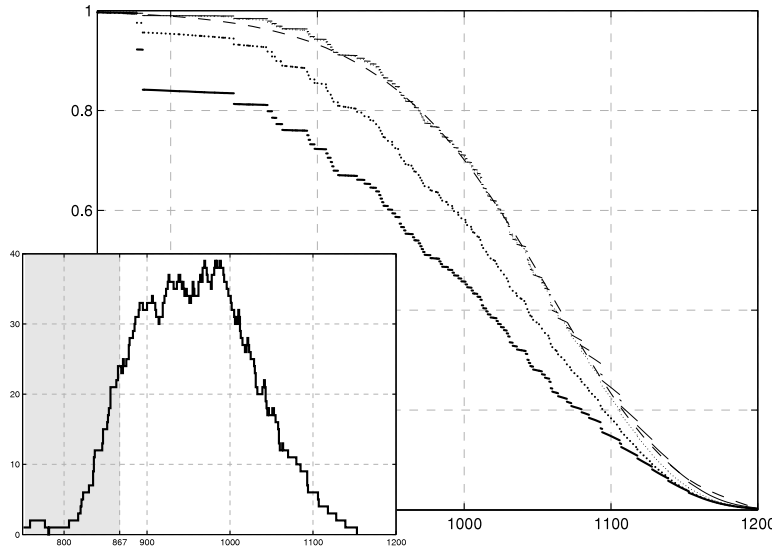
$$(\Lambda(s, t) \mid \text{data}, \gamma) \sim G(M_{j-1}(\gamma), n_0 \Lambda_0(s, t)), \quad (s, t) \subset (T_{j-1}, T_j),$$

– v bodech dělení T_j shodných s některým Z_i (tj. s $U_j + C_j > 0$) má Λ skok $\Delta \Lambda(T_j)$ s hustotu

$$x^{-1} e^{-(M_j(\gamma) + C_j)x} (1 - e^{-x})^{U_j} (1 - e^{-\gamma x})^{C_j} / c_j(\gamma), \quad x > 0,$$

2. parametr γ má hustotu

$$\pi(\gamma \mid \text{data}) \propto \left(\prod_{j=1}^N q_j(\gamma) \right) \pi(\gamma), \quad \gamma > 0, \quad \text{vzhledem k } \mu.$$



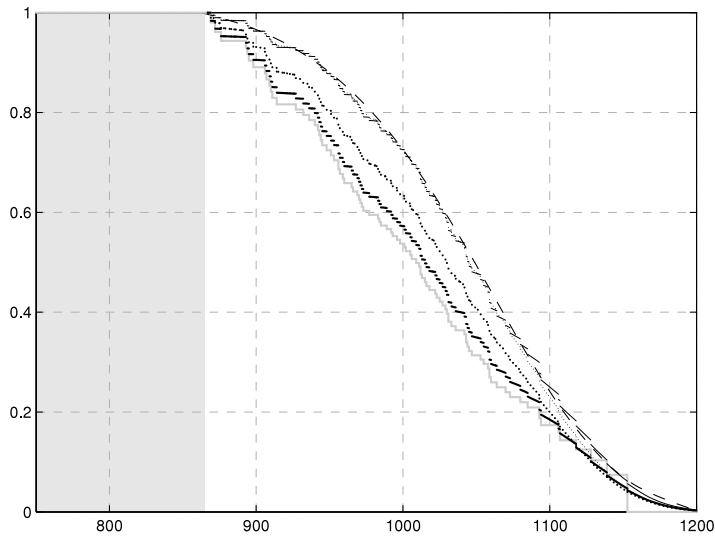
Obrázek 1: Neparаметrické bayesovské odhady pro Channing House Data. Čárkovaně použitá $\exp(-\Lambda_0)$, plnou čarou odhady při $n_0 = 10$, tečkovaně při $n_0 = 50$ (tlustě odhad z tvrzení 2, tenče když useknutí ignorováno). Vlevo počty jedinců sledovaných v daném věku.

Odhlédneme-li od parametru γ , změní se u přírůstků v aposteriorním rozdělení oproti apriornímu parametr měřítka, střední hodnota přírůstků na intervalech se sníží. Na druhou stranu přibudou skoky, jejichž rozdělení je dáno pouze konfigurací uspořádání pozorování, nikoliv jejich konkrétní polohou v čase.

Tvrzení 2. Bayesovský odhad funkce spolehlivosti $S(t)$ při kvadratické ztrátové funkci je

$$\begin{aligned} E(S(t) \mid \text{data}) = & \left[\int \left(\prod_{j < i} q_j^1 \right) \left(\prod_{j \geq i} q_j(\gamma) \right) \left(\frac{M_{i-1}(\gamma)}{M_{i-1}^1(\gamma)} \right)^{n_0 \Lambda_0(T_{i-1}, s)} \pi(\gamma) d\mu(\gamma) \right] \\ & \times \left[\int \left(\prod_{j=1}^N q_j(\gamma) \right) \pi(\gamma) d\mu(\gamma) \right]^{-1}, \quad t \in \langle T_{i-1}, T_i \rangle. \end{aligned}$$

Můžeme konstatovat, že výsledky jsou vlastně stejné jako v modelu bez useknutí zleva, „jen“ s jinak definovanými N_j , které zde berou v úvahu useknutí zleva.



Obrázek 2: Odhad podmíněné funkce spolehlivosti.

3 Příklad

Na minulém Robustu jsme porovnávali neparametrické bayesovské odhady (bez useknutí zleva) mj. na příkladu dat týkajících se délky života 97 mužů zaznamenaných v důchodcovské komunitě Channing House, viz např. [2]. Tato pozorování ale vznikla jako zleva uskenutá časem příchodu do komunity (tento údaj je u každého pozorování k dispozici), což jsme minule ignorovali. Dnes se podíváme, jak se na odhadech z těchto dat projeví, když useknutí zleva správně vezmeme do úvahy.

V následujících ukázkách jako referenční rozdělení uvažujeme stejně jako v [1] rozdělení Weibullovo s kumulativní intenzitou

$$\Lambda_0(t) = (t/\theta)^b, \quad t > 0,$$

s parametry $\theta = 1071$ a $b = 15,9$, kde jako parametr tvaru budeme případně místo b uvažovat hodnoty $2b$ a $b/2$, resp. místo parametru měřítka θ hodnotu $0,9\theta$ (viz obr. 3–5). Dále budeme uvažovat $n_0 = 50$ a 10 . Volba apriorního rozdělení pro parametr proporcionality γ nemá na výsledné odhady velký vliv, tedy alespoň pokud předem neomezíme obor hodnot γ . Prezentované grafy počítají s rovnoměrným rozdělením na diskretních hodnotách $\gamma = 1/p - 1$, $p = 0,1,0,2, \dots, 0,9$.

Nejprve na obr. 1 můžeme porovnat odhad z Tvzení 2 s odhady prezentovanými minule, které useknutí ignorují, tj. jako kdyby sledování všech jedinců začalo v čase 0 (v definici N_j by se použilo $A_k = 0$, $k = 1, \dots, n$).

Můžeme si všimnout relativně velkých skoků u prvních časů pozorování (v časech 777, 781 na začátku zobrazeného grafu funkce) zejména při menší

hodnotě parametru n_0 , kdy zvyšujeme váhu pozorování. To, jak asi očekáváme, souvisí s problémem počtu pozorování v jednotlivých intervalech. Skok našeho odhadu funkce spolehlivosti v časech pozorování T_j vzniká vynásobením její předchozí hodnoty faktorem, který má např. v případě okamžiku s necenzorovaným pozorováním ($U_j = 1, C_j = 0$) hodnotu vycházející z

$$E(e^{-\Delta\Lambda(T_j)} \mid \text{data}, \gamma) = \frac{\ln(1 + 1/(a + 1))}{\ln(1 + 1/a)},$$

kde $a = M_j(\gamma) = n_0 + N_j(1 + \gamma)$. Je-li pak a malé (z důvodu malého počtu pozorování N_j , viz levý graf v obr. 1, a malého $n_0 = 10$), může být tento výraz o dost menší než 1. Např. pro $a = 10, 11, 12$ vychází 0,91, 0,92 a 0,93.

Remark 3.1. *Podobnou vlastnost má ovšem i neparametrický odhad představený v [5]*

$$\hat{S}_{LTACL}(t) = \left(\prod_{i, Z_i \leq t} \left(1 - 1 / \sum_j I_{[A_j \leq Z_i \leq Z_j]} \right) \right)^{\hat{p}}, \quad \hat{p} = \sum U_j / n, \quad (4)$$

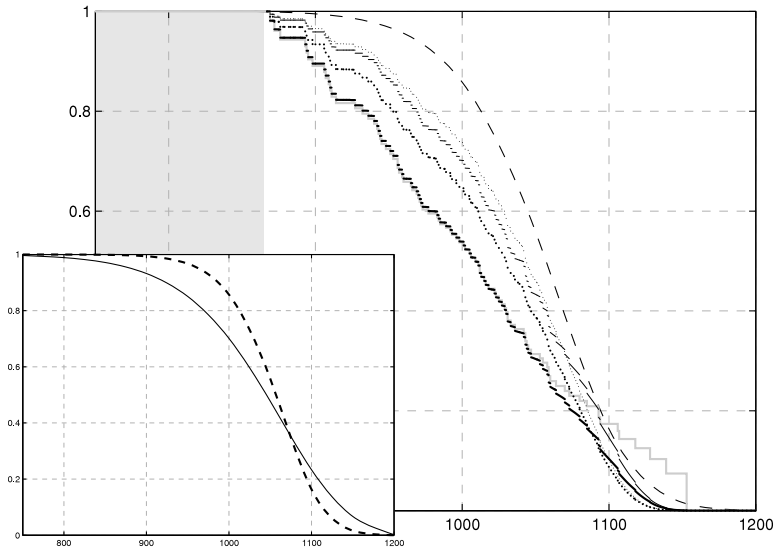
u kterého se může dokonce vyskytnout násobící faktor 0. Pro naše data se toto stane už u druhého nejmenšího z časů Z_i (mezi časy $T_4 = Z_{(2)}$ a $T_5 = A_{(3)}$ nebyl sledován žádný jedinec), odhad zůstává na intervalu $\langle T_4, \infty \rangle = \langle 781, \infty \rangle$ už nulový.

Proto se v dalším zaměříme na odhad funkce spolehlivosti podmíněného rozdělení

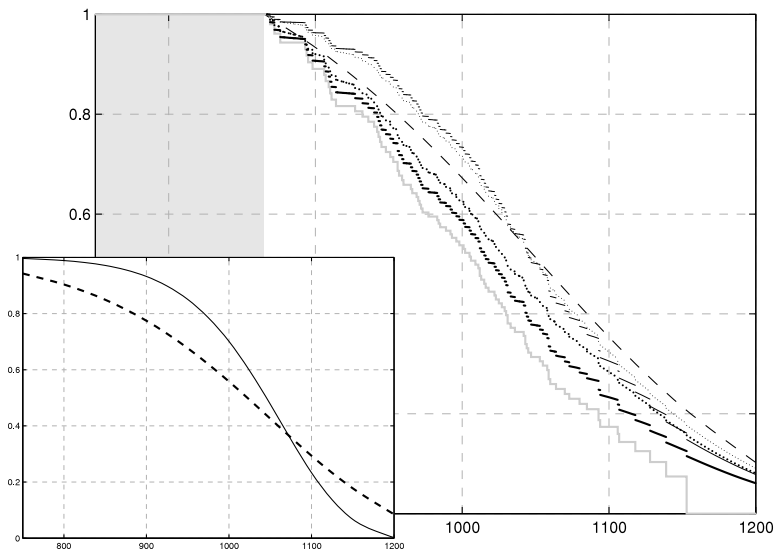
$$P(X > t \mid X > t_0) = e^{-\Lambda(t_0, t)}, \quad t > 0,$$

kde t_0 je zvoleno jako okamžik, kde již je k dispozici více pozorování. Tak jako jiní autoři u těchto dat budeme počítat s $t_0 = 867$, u prvního zahrnutého okamžiku T_j tak bude činit počet pozorování $N_j = 24$. Obr. 2 ukazuje situaci s referenční apriorní intenzitou Λ_0 . Po započtení vlivu useknutí se odhady posunují směrem k podmíněné verzi odhadu \hat{S}_{LTACL} , znázorněné jako šedá schodovitá funkce.

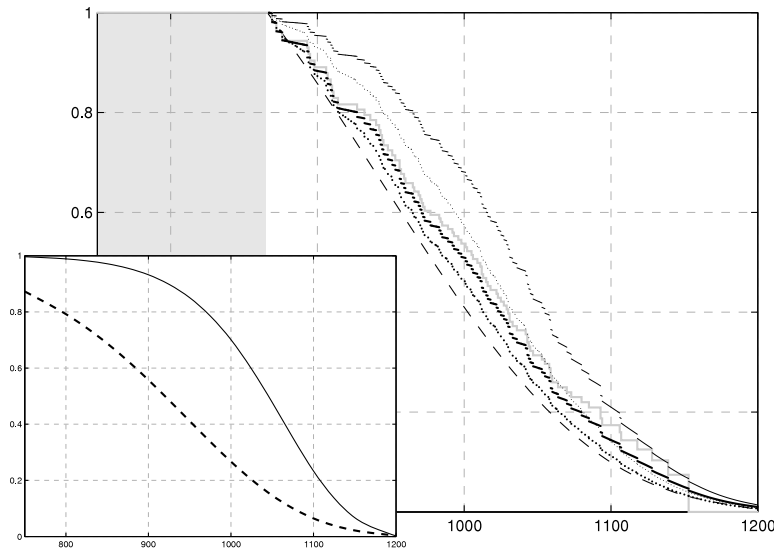
Na obr. 3, 4 a 5 se nakonec můžeme podívat, jak se na odhadech projeví změna tvaru apriorní střední hodnoty $\Lambda_0(t)$. Předpokládáme postupně rozdělení soustředěné na interval užší, širší a na interval rozšířený směrem doleva. Příslušná $\exp(-\Lambda_0(t))$ je zakreslena vždy čárkovanou čarou v malém obrázku, ve velkém pak spolu s podmíněnými odhady i její podmíněná verze.



Obrázek 3: Pro Λ_0 s parametry $(\theta, 2b)$.



Obrázek 4: Pro Λ_0 s parametry $(\theta, b/2)$.



Obrázek 5: Pro Λ_0 s parametry $(0, 90, b/2)$.

Literatura

- [1] Friesl M. (2006). *Porovnání neparametrických bayesovských odhadů při cenzorování*. In ROBUST 2006 (Antoch J. a Dohnal G., eds.), JČMF, Praha, 83–90.
- [2] Hyde J. (1977). *Testing survival under right censoring and left truncation*. Biometrika **64** (no. 2), 225–230.
- [3] Kim Y. (1999). *Nonparametric Bayesian estimators for counting processes*. Ann. Statist. **27**, no. 2, 562–588.
- [4] Koziol J. A., Green S. B. (1976). *A Cramér-von Mises statistic for randomly censored data*. Biometrika **63** (no. 3), 465–474.
- [5] Pawlitschko J. (2000). *Estimation in the Koziol-Green model with left truncated observations*. Sankhyā Ser. A **62**, no. 1, 67–79.

Poděkování: Tato práce byla podporována grantem MSM 4977751301.

Adresa: M. Friesl, FAV ZČU, KMA, Univerzitní 22, 306 14 Plzeň

E-mail: friesl@kma.zcu.cz

MODEL PRO NÁHODNÉ SJEDNOCENÍ KRUHŮ SE VZÁJEMNÝMI INTERAKCEMI

Kateřina Helisová, Jesper Møller

Klíčová slova: Booleovský model, interakce, Markovský proces, model zárodku a zrna, náhodná množina, proces kruhů.

Abstrakt: Příspěvek se zabývá modelem náhodné množiny dané konečným sjednocením kruhů, mezi nimiž se vyskytují vzájemné interakce. Tento model je popsán hustotou pravděpodobnosti vzhledem k Booleovskému modelu, která závisí na geometrických charakteristikách (např. plocha, obvod apod.) dané množiny. Popsány jsou zde hlavně metody odhadu parametrů hustoty, testování jejich významnosti a kontroly vhodnosti modelu.

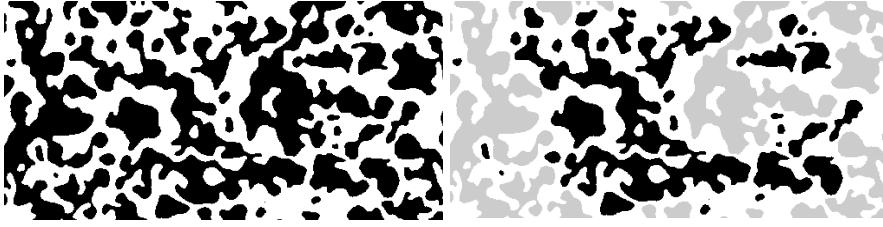
Abstract: The contribution concerns a model of a random set given by a finite union of interacting discs. The model is described by a density with respect to a Boolean model. This density depends on geometrical characteristics (e.g. area or perimeter) of the given set. Methods for estimating the parameters of the density, testing their relevance and checking the fit of the model are shown.

1 Úvod

Mnoho jevů v přírodě (keře, stromy v lese apod.) je modelováno sjednocením kompaktních množin v prostoru \mathbf{R}^2 . Ve stochastické geometrii se pro tyto modely používá název *modely zárodku a zrna*. Název je odvozen z konstrukce těchto modelů. Jednotlivé objekty reprezentovány kompaktními množinami tvoří základní stavební prvky, tzv. *zrna*, jejichž tvar může být náhodný. Umístění v prostoru dané posunem každého zrna K_i o vektor u_i je také náhodné a bod u_i reprezentující toto umístění se nazývá *zárodkem*. Sjednocení zárodků je bodový proces v prostoru \mathbf{R}^2 a náhodná množina popsaná tímto modelem je pak sjednocení $\cup\{u_i + K_i\} = \cup\{u_i + x : x \in K_i\}$.

Pokud jsou tvary jednotlivých zrn navzájem nezávislé a nezávisí ani na zárodcích, mluvíme o *Booleovských modelech*. Často však jednotlivá zrna interagují s jinými, což je pak třeba při tvorbě modelu vzít v úvahu.

Tento článek je rešerší [2] a [3] zaměřených na model, v němž jsou zrna tvořena kruhy s náhodnými poloměry a interakce jsou dány geometrickými charakteristikami celého sjednocení. Model jsme aplikovali na data popisující vřesový porost (viz Obrázek 1). Průřez keře vřesu totiž můžeme považovat za kruh a z dřívějších analýz těchto dat jsou interakce mezi keři patrné. Jelikož navíc tvar hustoty, jíž je model popsán, závisí na celém sjednocení a nikoliv na jednotlivých kruzích, máme vyřešen také problém, že z dat nejsou hranice jednotlivých keřů pozorovatelné.



Obrázek 1: Keře vřesu označené černě, vpravo s šedě vyznačenými komponentami protínající hranici pozorovacího okna.

2 Model

2.1 Bodový proces daný hustotou

Nechť (Ω, \mathcal{F}, P) je pravděpodobnostní prostor. Označme N systém lokálně konečných podmnožin \mathbf{R}^d , $\mathcal{N} = \sigma(\{\mathbf{x} \in N : \#(\mathbf{x} \cap A) = m\} : A \in \mathcal{B}, m \in \mathbf{N}_0)$, kde \mathcal{B} značí omezené Borelovské množiny, a uvažujme *bodový proces* X na \mathbf{R}^d jako měřitelné zobrazení z (Ω, \mathcal{F}) do (N, \mathcal{N}) . Mějme *Poissonův bodový proces* Y s mírou intenzity μ (tj. proces splňující, že pro systém $\{A_n\}$ disjunktních množin v \mathbf{R}^d jsou počty bodů v těchto množinách $Y(A_n)$ nezávislé náhodné veličiny a pro každou $A \subset \mathbf{R}^d$ takovou, že $\mu(A) < \infty$, má $Y(A)$ Poissonovo rozdělení s parametrem $\mu(A) = EY(A)$). Pak při označení $\Pi(F) = P(Y \in F)$, $F \in \mathcal{N}$, je X *bodový proces daný hustotou f vzhledem k Poissonovu procesu Y* , jestliže

$$P(X \in F) = \int_F f(\mathbf{x}) \Pi(d\mathbf{x}).$$

2.2 Proces kruhů daný hustotou

Pro konstrukci modelu interagujících kruhů si označme $b = b(z, r)$ kruh se středem v bodě $z \in \mathbf{R}^2$ a poloměrem $r \in (0, \infty)$. Ztotožníme-li b s bodem $x = (z, r) \in \mathbf{R}^2 \times (0, \infty)$, pak sjednocení $\cup_{i \in I} b_i = \cup_{i \in I} b(z_i, r_i)$, $I \subseteq \mathbf{N}$, můžeme ztotožnit s bodovým procesem na $\mathbf{R}^2 \times (0, \infty)$.

Za vztažný proces si zvolíme Poissonův bodový proces Y (tj. vztažný Booleovský model je náhodná množina daná sjednocením kruhů odpovídajících procesu Y) s mírou intenzity $\rho(z) dz q(dr)$ na $\mathbf{R}^2 \times (0, \infty)$, kde ρ značí funkci intenzity bodového procesu středů a q je rozdělení poloměrů kruhů.

Výsledným modelem pak bude sjednocení kruhů $\mathcal{U}_{\mathbf{X}}$ odpovídající bodovému procesu \mathbf{X} , který je absolutně spojitý vzhledem k Poissonovu procesu Y a je vzhledem k němu daný hustotou $f(\mathbf{x})$ pro konfiguraci $\mathbf{x} = \{x_1, \dots, x_n\}$.

Budeme navíc předpokládat, že \mathbf{X} je konečný bodový proces definovaný na $S \times (0, R)$, kde $S \subset \mathbf{R}^2$ je omezená množina taková, že $\int_S \rho(z) dz > 0$, a $R < \infty$. Hustotu pak předpokládáme ve tvaru

$$\begin{aligned}
f_{\theta}(\mathbf{x}) &= \frac{1}{c_{\theta}} \exp(\theta \cdot T(\mathcal{U}_{\mathbf{x}})) \\
&= \frac{1}{c_{\theta}} \exp(\theta_1 A(\mathcal{U}_{\mathbf{x}}) + \theta_2 L(\mathcal{U}_{\mathbf{x}}) + \theta_3 N_{cc}(\mathcal{U}_{\mathbf{x}}) + \theta_4 N_h(\mathcal{U}_{\mathbf{x}})),
\end{aligned} \tag{1}$$

kde c_{θ} je normovací konstanta, $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$ je vektor parametrů, \cdot značí skalární součin a $T(\mathcal{U}_{\mathbf{x}}) = (A(\mathcal{U}_{\mathbf{x}}), L(\mathcal{U}_{\mathbf{x}}), N_{cc}(\mathcal{U}_{\mathbf{x}}), N_h(\mathcal{U}_{\mathbf{x}}))$ je vektor geometrických charakteristik sjednocení kruhů $\mathcal{U}_{\mathbf{x}}$ odpovídajících konfiguraci \mathbf{x} , přičemž $A(\mathcal{U}_{\mathbf{x}})$ značí plochu, $L(\mathcal{U}_{\mathbf{x}})$ obvod, $N_{cc}(\mathcal{U}_{\mathbf{x}})$ počet spojitých komponent a $N_h(\mathcal{U}_{\mathbf{x}})$ počet děr v tomto sjednocení.

Jelikož interakce v procesu daném hustotou (1) závisí na vektoru geometrických charakteristik T , mluvíme o něm jako o *T-interakčním procesu*.

2.3 Markovská vlastnost modelu

Definujme *Papangelouovu podmíněnou intenzitu* jako

$$\lambda_{\theta}(\mathbf{x}, v) = f_{\theta}(\mathbf{x} \cup \{v\}) / f_{\theta}(\mathbf{x}).$$

T -interakční proces nazveme *Markovským* vzhledem k reflexivní relaci \sim (nazývané *sousedství*), jestliže $\lambda_{\theta}(\mathbf{x}, v)$ závisí na \mathbf{x} pouze prostřednictvím $\{u \in \mathbf{x} : u \sim v\}$, tj. sousedů v v konfiguraci \mathbf{x} . Obvykle se předpokládá, že relace \sim závisí pouze na porovnávání bodech. Pokud uvažujeme relaci \sim závislou i na zbytku konfigurace \mathbf{x} , mluvíme o *Markovském procesu nejbližšího souseda*. Pro oba typy procesů platí Hammersley-Cliffordova věta (viz [4]).

Jelikož pro T -interakční proces lze hustotu psát ve tvaru

$$f_{\theta}(\mathbf{x}) = \frac{1}{c_{\theta}} \prod_{K \in \mathcal{K}(\mathcal{U}_{\mathbf{x}})} \exp(\theta_1 A(K) + \theta_2 L(K) + \theta_3 + \theta_4 N_h(K)),$$

kde $\mathcal{K}(\mathcal{U}_{\mathbf{x}})$ je množina všech spojitých komponent $\mathcal{U}_{\mathbf{x}}$, pak pokud definujeme relaci \sim jako příslušnost k téže spojitě komponentě, je podle Hammersley-Cliffordovy věty tento model Markovským procesem (nejbližších sousedů).

3 Analýza dat

3.1 Data

Pro aplikaci jsme si vybrali data poprvé prezentována v [1] znázorňující (černou barvou) keře vřesu na ploše 10×20 m v oblasti Jädraås ve Švédsku (viz Obrázek 1 vlevo). Data jsme získali v rozlišení 250×508 pixelů, tj. 1 pixel odpovídá přibližně čtverci o straně délky 4cm.

3.2 Okrajové efekty

Jelikož pozorovací okno W je menší než plocha, na které se nachází sbíraná data, nastává problém s chybějící informací o chování komponent ležících

mimo okno, které mohou chování uvnitř okna ovlivnit (tzv. *problém s okrajovými efekty*). Tento problém však lze vyřešit použitím Markovské vlastnosti popsané v kapitole 2.3.

Rozdělme \mathbf{X} na procesy $\mathbf{X}^{(a)}$, $\mathbf{X}^{(b)}$, $\mathbf{X}^{(c)}$ odpovídající spojitým komponentám $\mathcal{U}_{\mathbf{X}}$ ležícím postupně (a) celé v pozorovacím W , (b) částečně v okně W a částečně mimo něj a (c) ležící celé mimo W . Označíme-li $\mathbf{x}^{(b)}$ realizaci $\mathbf{X}^{(b)}$, pak z Markovské vlastnosti plyne, že podmíněně při $\mathbf{X}^{(b)} = \mathbf{x}^{(b)}$ jsou $\mathbf{X}^{(a)}$ a $\mathbf{X}^{(c)}$ vzájemně nezávislé a navíc podmíněně rozdělení $\mathbf{X}^{(a)}$ závisí na $\mathbf{x}^{(b)}$ pouze prostřednictvím $V = W \cap \mathcal{U}_{\mathbf{x}^{(b)}}$.

Toto dělení procesu \mathbf{X} je znázorněno na datech na Obrázku 1 vpravo.

Abychom se vyhnuli problémům s okrajovými efekty, budeme za data považovat pouze realizaci procesu $\mathbf{X}^{(a)}$, tj. komponenty ležící celé uvnitř W , a za pozorovací okno množinu $W \setminus V = W \setminus \mathcal{U}_{\mathbf{x}^{(b)}}$. Hodnoty charakteristik těchto dat jsou $A = 45.6$ (m²), $L = 190.0$ (m), $N_{cc} = 32$ a $N_h = 2$.

3.3 Vztažné procesy

Jedinou podmínkou pro volbu vztažného procesu je požadavek, aby výsledný model byl vzhledem k němu absolutně spojitý. Rozhodli jsme se zkoumat, jak volba vztažného procesu ovlivňuje hodnoty a významnost parametrů v modelu, proto jsme jako vztažné procesy zvolili kvůli porovnání tři Booleovské modely s různými intenzitami počtu kruhů ρ a různými náhodnými poloměry Q . Těmito procesy jsou

(R1): $\rho = 2.45$ a Q má rozdělení $N(0.26, 0.16^2)$ zúžené na interval $[0, 0.50]$,

(R2): $\rho = 2.45$ a Q má rovnoměrné rozdělení na intervalu $[0, 0.53]$,

(R3): $\rho = 1.16$ a Q má rovnoměrné rozdělení na intervalu $[0, 0.53]$.

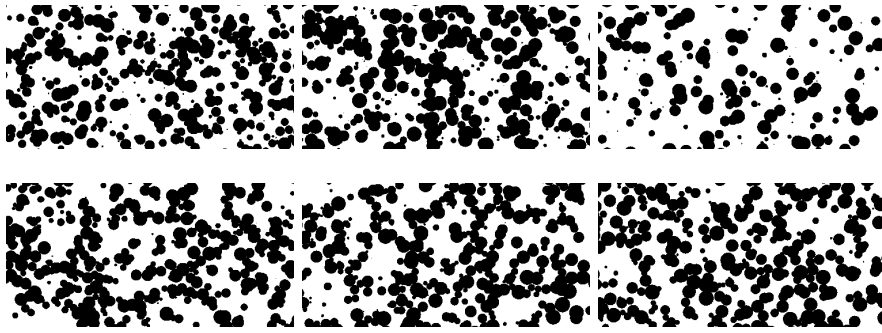
Rozdělení poloměrů byla volena tak, aby jejich horní hranice nebyla menší než maximálnímu poloměru keře, který je dle předešlých analýz těchto dat 0.5m. Konkrétní hodnoty vycházejí rovněž z předchozích analýz. Simulace těchto procesů jsou ukázány na Obrázku 2.

3.4 Odhad parametrů a Waldův test

K získání $\hat{\theta}$ odhadu parametru $\theta = (\theta_1, \dots, \theta_4)$ jsme použili metodu maximální věrohodnosti založenou na simulacích (viz [4]). Mějme pozorování $\mathcal{U}_{\mathbf{x}}$ a pišme hustotu ve tvaru $f_{\theta}(\mathbf{x}) = h_{\theta}(\mathbf{x})/c_{\theta}$. Logaritmicko-věrohodnostní funkce je

$$\begin{aligned} l(\theta) &= \log h_{\theta}(\mathbf{x}) - \log c_{\theta} \\ &= \theta_1 A(\mathcal{U}_{\mathbf{x}}) + \theta_2 L(\mathcal{U}_{\mathbf{x}}) + \theta_3 N_{cc}(\mathcal{U}_{\mathbf{x}}) + \theta_4 N_h(\mathcal{U}_{\mathbf{x}}) - \log c_{\theta}. \end{aligned} \quad (2)$$

Jelikož ale c_{θ} nemá explicitní vyjádření, nelze tuto funkci maximalizovat přímo. Proto je potřeba místo (2) maximalizovat pro pevný vektor θ_0 poměr věrohodností $l(\theta) - l(\theta_0) = \log(h_{\theta}(\mathbf{x})/h_{\theta_0}(\mathbf{x})) - \log(c_{\theta}/c_{\theta_0})$, který lze aproximovat pomocí



Obrázek 2: Simulace vztažných Booleovských modelů (R1)-(R3) (nahore zleva doprava) a (A, L, N_{cc}) -interakčních modelů vzhledem ke vztažným procesům (R1)-(R3) (dole rovněž zleva doprava).

$$l(\theta) - l(\theta_0) = \log(h_\theta(\mathbf{x})/h_{\theta_0}(\mathbf{x})) - \log \frac{1}{n} \sum_{m=0}^{n-1} h_\theta(Z_m)/h_{\theta_0}(Z_m),$$

kde Z_m jsou realizace z hustoty f_{θ_0} získané pomocí MCMC simulací.

Pro testování významnosti parametrů jsme použili Waldův test (viz [4]), který testuje hypotézu $H : \theta M = 0$, kde M je $p \times k$ -rozměrná matice hodnoty k a p je délka vektoru θ . Waldova statistika $(\hat{\theta}M)(M^T j(\hat{\theta})^{-1}M)(\hat{\theta}M)^T$ pak má asymptoticky χ_k^2 -rozdělení, kde $j(\hat{\theta}) = \text{Var}_\theta T(X)$ je kovarianční matice geometrických charakteristik, kterou lze získat ze simulací procesu z hustoty s parametrem $\hat{\theta}$.

Pro speciální volbu, kdy M je matice rozměru $p \times 1$ se všemi prvky rovnými nule kromě i -tého, který je roven 1, platí $H' : \theta_i = 0 \iff H : \theta M = 0$.

Výsledky odhadů parametrů a hodnoty příslušné Waldovy statistiky jsou zobrazeny v následující tabulce.

	θ_1	θ_2	θ_3	θ_4
(R1)	-2.14	0.89	-1.78	-1.01
Wald	7.45	17.89	48.96	2.14
(R2)	-4.81	1.17	-2.26	-0.69
Wald	37.04	29.77	83.66	1.01
(R3)	-3.67	1.62	-2.25	-0.13
Wald	17.01	46.67	73.01	0.04

Jelikož kritická hodnota $\chi_1^2(0.95) = 3.842$, můžeme pro všechny tři konstruované modely považovat parametr θ_4 za nulový a tudíž N_h z hustoty vypustíme. Pro nové modely, jejichž hustoty obsahují pouze A , L a N_{cc} , jsme provedli nové odhady parametrů a zjistili jsme, že jejich hodnoty se příliš neliší od hodnot pro původní model, jak je vidět i z následující tabulky.

	θ_1	θ_2	θ_3
(R1)	-2.33	0.92	-1.77
Wald	9.54	21.01	46.89
(R2)	-4.91	1.18	-2.25
Wald	38.02	32.33	78.78
(R3)	-3.71	1.64	-2.25
Wald	17.04	47.11	73.89

Jelikož se na klasické 5% hladině jeví všechny parametry jako významné, považujeme (A, L, N_{cc}) –interakční modely s uvedenými parametry za výsledné modely. Jejich simulace jsou ukázány na Obrázku 2. Je vidět, že oproti vztažným Booleovským modelům se v simulacích (A, L, N_{cc}) –interakčních modelů nevyskytuje tolik samostatných kruhů, čímž se vztažné modely od dat nejvíce odlišovaly. Z tohoto pohledu se (A, L, N_{cc}) –interakční procesy datům podobají mnohem více než vztažné Booleovské modely.

4 Testování vhodnosti modelu

K testování byly použity sumární statistiky náhodných množin uvedeny např. v [6] a statistiky založené na modifikaci množiny popsané blíže v [5].

Nechť $\mathbf{A} \subset \mathbf{R}^2$ je množina pozorovaná v (omezeném) okně $W \subset \mathbf{R}^2$ a G je množina pixelů digitálního záznamu \mathbf{A} .

4.1 Normovaná sférická kontaktní distribuční funkce

Sférická kontaktní distribuční funkce je distribuční funkce náhodné veličiny popisující vzdálenost od libovolného bodu ležícího mimo náhodnou množinu k nejbližšímu bodu této množiny. Matematicky je definována jako $H(r) = P(D \leq r | D > 0)$, $r \geq 0$, kde $D = \inf\{r \geq 0 : \mathbf{A} \cap b(0, r) \neq \emptyset\}$. Její neparametrický odhad je dán vztahem

$$\hat{H}(r) = \frac{\sum_{u \in G} \mathbf{1}[u \notin \mathbf{A}, u + b(0, r) \subset W, (u + b(0, r)) \cap \mathbf{A} \neq \emptyset]}{\sum_{u \in G} \mathbf{1}[u \notin \mathbf{A}, u + b(0, r) \subset W]}.$$

Pro větší přehlednost se využívá *normovaná* kontaktní distribuční funkce

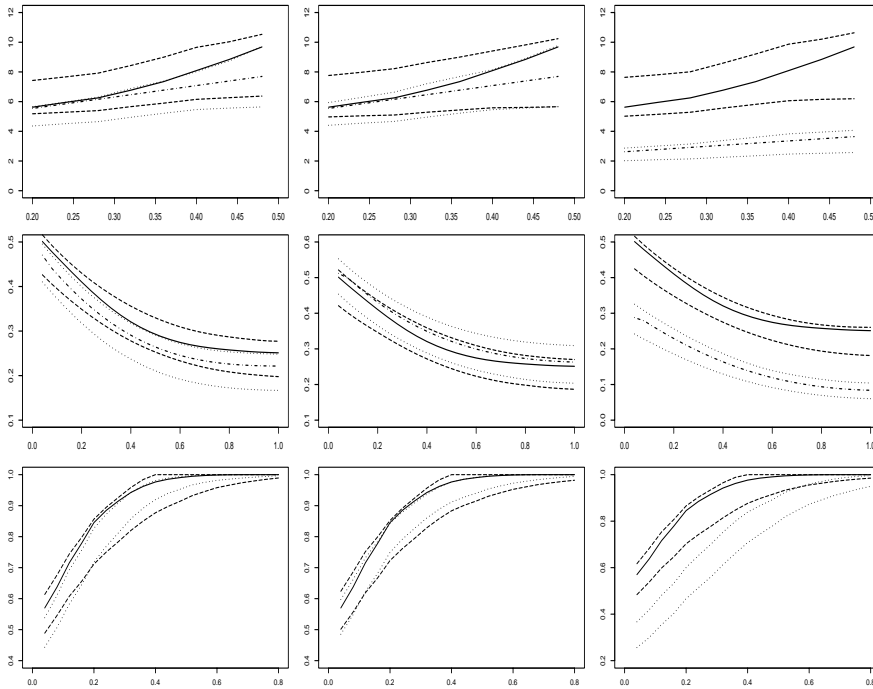
$$T(r) = -\frac{1}{r} \log(1 - H(r)),$$

jejíž tvar pro vztažné Booleovské modely je $T(r) = 2\rho\pi EQ + \rho\pi r$.

4.2 Kovarianční funkce

Kovarianční funkce je definována jako $C(r) = P(u \in \mathbf{A}, v \in \mathbf{A})$ pro $u, v \in \mathbf{R}^2 : \|u - v\| = r$. Jejím neparametrickým odhadem je

$$\hat{C}(r) = \frac{\sum_{u, v \in G} \mathbf{1}[\|u - v\| = r, \{u, v\} \subset \mathbf{A}]}{\sum_{u, v \in G} \mathbf{1}[\|u - v\| = r]}.$$



Obrázek 3: Sloupce odpovídají zleva doprava procesům (R1)-(R3), řádky shora dolů statistikám $T(r)$, $C(r)$ a $d(r)$. Na jednotlivých grafech jsou znázorněny jejich odhady získané z dat (plná křivka), 2.5 % a 97.5 % obálky získané ze simulací vztažného Booleovského modelu (tečkovaná křivka), 2.5 % a 97.5 % obálky získané ze simulací (A, L, N_{cc}) -interakčního modelu (čárkovaná křivka) a - s výjimkou dilatace - teoretické funkce pro Booleovský model (čerchovaná křivka).

Pro Booleovské modely má pak funkce tvar

$$C(r) = 2p - 1 + (1 - p)^2 \exp\left(\rho E\left[2Q^2 \arccos \frac{Q}{2r} - \frac{r}{2} \sqrt{4Q^2 - r^2}\right]\right).$$

4.3 Dilatace množiny

Nechť $\mathbf{A}_{\oplus r} = \cup_{u \in \mathbf{A}} b(u, r)$ je zvětšení a $\mathbf{A}_{\ominus r} = \{u : b(u, r) \subseteq \mathbf{A}\}$ zmenšení množiny \mathbf{A} o hodnotu r . Dilatace množiny \mathbf{A} je definována jako funkce

$$d(r) = \frac{|\mathbf{A}_{\oplus r} \cap W_{\ominus r}|}{|W_{\ominus r}|}.$$

Porovnání výše popsaných statistik získaných z dat s teoretickými funkcemi pro vztažné Booleovské modely a s 2.5% a 97.5% obálkami získanými ze simulací (viz [4]) vztažného a (A, L, N_{cc}) -interakčního modelu je ukázáno

na Obrázku 3. Z grafů je stejně jako při vizuálním porovnání patrné, že (A, L, N_{cc}) –interakční modely popisují data lépe než vztažné Booleovské modely. Navíc vzhledem k tomu, že všechny uvedené statistiky odhadnuté z dat leží v 2.5% – 97.5% obálkách, můžeme vyslovit závěr, že vytvořený proces je modelem dobře popisujícím analyzovanou množinu.

Literatura

- [1] Diggle P.J. (1981). *Binary mosaics and the spatial pattern of heather*. *Biometrics* **37**, 531 – 539.
- [2] Møller J., Helisová K. (2008). *Power diagrams and interaction processes for unions of discs*. *Advances in Applied Probability* **40**, 321 – 347.
- [3] Møller J., Helisová K. (2008). *Likelihood inference for unions of interacting discs*. Přípravuje se.
- [4] Møller J., Waagepetersen R.P. (2004). *Statistical Inference and Simulation for Spatial Point Processes*. Chapman and Hall/CRC, Boca Raton.
- [5] Ripley B.D. (1988). *Statistical Inference for Spatial Processes*. Cambridge University Press, Cambridge.
- [6] Stoyan D., Kendall W.S., Mecke J. (1995). *Stochastic Geometry and Its Applications*. Wiley, Chichester.

Poděkování: Tato práce byla podporována granty GAČR 201/06/0302 a GAČR 201/05/H007, výzkumným záměrem MSM6840770038 a grantem 272-06-0442 „Point process modelling and statistical inference“ (Danish Natural Science Research Council).

Adresa: K. Helisová, MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8 – Karlín; FEL ČVUT, Katedra matematiky, Technická 2, 166 27 Praha 2 – Dejvice;

J. Møller, Aalborg University, Department of Mathematical Sciences, Frederik Bajers Vej 7G, DK-9220 Aalborg, Denmark

E-mail: helisova@karlin.mff.cuni.cz, helisova@math.feld.cvut.cz, jm@math.aau.dk

VÝPRAVY DO HLUBIN DAT

Daniel Hlubinka

Klíčová slova: Hloubka dat, kvantilové kontury, směrové kvantily, zobecněná poloprostorová hloubka

Abstrakt: V článku je představena hloubka dat jako zobecnění pojmu uspořádání pro účely mnohorozměrné statistické analýzy. Pro náhodné vektory jsou zopakovány hlavní typy hloubek spolu s jejich vlastnostmi, výhodami i nevýhodami. V druhé části článku jsou diskutovány dvě alternativní definice hloubky, které umožňují oproti běžným hloubkám zvýraznit lokální vlastnosti rozdělení.

Abstract: The data depth is introduced as a generalization of ranking to be used in the multivariate statistical analysis. The most common definitions of depth are recalled together with their properties and both advantages and disadvantages. In the second part of the paper two alternative definitions of depth are discussed. These definitions allow to emphasize, in contrast to most usual depths, the local behaviour of the underlying distribution.

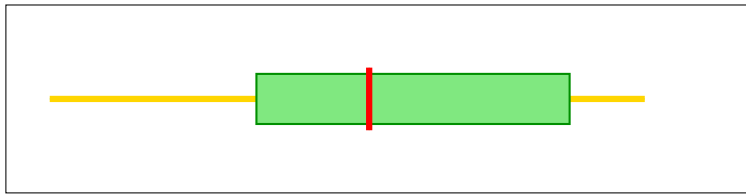
1 Kvantil a problém uspořádání

Ve statistické analýze se občas objeví chvíle, kdy potřebujeme hovořit o oblasti spolehlivosti, predikční oblasti, kritické oblasti. V jednorozměrném případě si obvykle můžeme pomoci kvantily, ve většině případů nejde o složitý problém. Nemusíme ale dlouho vzpomínat, abychom si uvědomili, že kvantil *tak jak jej známe* je pojem úzce vázaný na jednorozměrnou náhodnou veličinu. Většina z nás bude velmi váhat s definicí kvantilu pro dvou- a více-rozměrná rozdělení. Jednou z výjimek je mnohorozměrné normální rozdělení, kde existují přirozeně definované elipsoidy (definované pomocí střední hodnoty a varianční matice), které plně uspokojí naše požadavky na uvedené oblasti spolehlivosti. Jak se ale obejít bez předpokladu normality? V tomto článku budeme hledat nějaký vhodný neparametrický přístup k pojmu kvantil ve více rozměrech. Napřed si ale připomeňme, co víme o jednorozměrných kvantilech a mohlo by při našem hledání být užitečné.

Pro jednorozměrnou náhodnou veličinu X je kvantil definován různými způsoby. Nejčastěji tak, že hodnoty náhodné veličiny kvantil q_α rozděluje na menší a větší; přitom platí zároveň

$$P[X \leq q_\alpha] \geq \alpha, \quad P[X \geq q_\alpha] \geq 1 - \alpha,$$

a hodnota $\alpha \in [0, 1]$ je předem pevně zvolená. Kvantil v \mathbb{R}^p musí být p -rozměrný vektor. Zde je jádro problémů pro náhodné vektory—v prostoru \mathbb{R}^p není přirozené lineární uspořádání. My ale musíme být schopni porovnat



Obrázek 1: Inspirace krabicovým diagramem.

každé dvě hodnoty náhodného vektoru, každé dva vektory, jinak není možné definovat oblast spolehlivosti analogicky jednorozměrnému případu.

Pro jednorozměrnou náhodnou veličinu existuje nástroj, který nám může být inspirací i pro náhodné vektory. Podívejme se na známý krabicový diagram, jednoduchý obrázek patřící do *obyčejné* popisné statistiky. V něm pomocí jednorozměrných kvantilů zobrazíme podstatnou informaci o rozdělení. Zaujme nás *prostřední* bod, *medián*, *vnitřní krabíčka*, ohraničená dolním a horním kvantilem, a *tykadla—vousy*, ukazující k nejvyšší a nejnižší hodnotě¹. Náš zrak je zde veden nejenom zleva doprava, od malých hodnot k větším, ale spíš zprostřed ke krajům, od středových hodnot ke vnějším. Každé dvě hodnoty tak mohou být porovnány s ohledem na *vzdálenost* od středu, přičemž pojem *vzdálenosti* je pevně svázán s rozdělením; v jednorozměrném případě je přímo definován pomocí kvantilů.

Zde se potkávají dva koncepty. Na jedné straně je to *míra odlehlosti od středu*². Na druhé straně to je naopak určitá *obklopenost* ostatními hodnotami náhodné veličiny. Pro medián je polovina realizací náhodné veličiny větších a polovina menších, je tedy *uprostřed*, nebo *hluboko v rozdělení*. Pro dolní kvartil *jen čtvrtina* realizací náhodné veličiny je menších, je tedy z této strany mnohem méně obklopen rozdělením než medián, je *méně hluboko*.

Takto se budeme dívat i na mnohorozměrné náhodné vektory. Každá hodnota může být s jinou porovnána s ohledem na *odlehlost* od středu. Budeme proto posuzovat, jak *hluboko* se určitý bod nachází v daném rozdělení. Neboli zkoumáme, jak je bod ve výběrovém prostoru *obklopen* ostatními body s ohledem na pravděpodobnostní rozdělení. Tomuto konceptu budeme říkat *hloubka*, *hloubka dat*³. Funkce, které bodům ve výběrovém prostoru přiřazují jejich hloubku, nazveme *hloubky*, případně *hloubkové funkce*. Takových funkcí je jistě celá řada. V příspěvku nejprve pohlédneme na hloubku trochu z nadhledu, ujasníme si čím jsou hloubkové funkce inspirovány, jak se dají třídit do skupin, jaké jsou jejich žádoucí a nežádoucí vlastnosti. Také si představíme několik známých i méně známých hloubkových funkcí i s jejich základními výhodami a nevýhodami. Ve druhé části článku se budeme věnovat dvěma konceptům hloubky, které v současnosti podrobněji zkoumáme.

¹Obvykle se usekávají v určité vzdálenosti od krabíčky, ale to každý dobře zná

²V angličtině se většinou používá výraz *outlyingness*

³Anglický termín je *depth*, *data depth*. O hloubce *dat* hovoříme proto, že původně tento koncept vyšel z popisné statistiky a byl aplikován na náhodné výběry.

2 Kvantil v říši více rozměrů

Již jsme uvedli, že pro dvou- a vícerozměrné náhodné vektory místo uspořádaní hodnot podle velikosti uijeme lineární kvaziuspořádání podle *odlehlosti*⁴. Formálně postupujeme takto. Označme \mathcal{P} soubor všech pravděpodobnostních rozdělení na prostoru \mathbb{R}^p .

Definice 1. *Hloubkovou funkcí rozumíme zobrazení $H : \mathcal{P} \times \mathbb{R}^p \rightarrow \mathbb{R}^+$.*

Bod $x \in \mathbb{R}^p$ je pro rozdělení $P \in \mathcal{P}$ hlubší než bod $y \in \mathbb{R}^p$, pokud $H(P, x) \geq H(P, y)$. Bod $d \in \mathbb{R}^p$ je pro rozdělení $P \in \mathcal{P}$ nejhlubší, pokud $H(P, d) \geq H(P, x)$ pro všechna $x \in \mathbb{R}^p$. Označme $d_{P,H} = H(P, d)$ hloubku nejhlubšího bodu.

Nebude-li třeba zdůrazňovat rozdělení P budeme psát krátce $H(x)$ místo $H(P, x)$ a d_H místo $d_{P,H}$, případně napíšeme ještě stručněji pouze d .

Dva různé body mohou mít stejnou hloubku, proto jde o kvaziuspořádání a každé dva body jsou vzájemně porovnatelné s ohledem na svou hloubku, proto se jedná o *lineární* kvaziuspořádání.

Nejhlubší bod budeme chápat jako mnohorozměrnou obdobu mediánu. Nemusí být dán jednoznačně, stejně jako nemusí být jednoznačný medián v jednorozměrném případě. Díky hloubce lze definovat *centrální* oblasti pro dané rozdělení a *kvantilovou konturu*.

Definice 2. *Pro dané rozdělení P a hloubkovou funkci H definujeme a hlubokou oblast*

$$R_{P,H}(a) := \{x : H(P, x) \geq a\}, 0 \leq a \leq d_{P,H}.$$

Hranice oblasti $R_{P,H}$ se nazývá a -hloubkovou konturou a budeme ji značit $c_{P,H}(a)$. V případě, kdy nemůže dojít k záměně, značíme a hlubokou oblast zkráceně $R(a)$, a hloubkovou konturu obdobně $c(a)$.

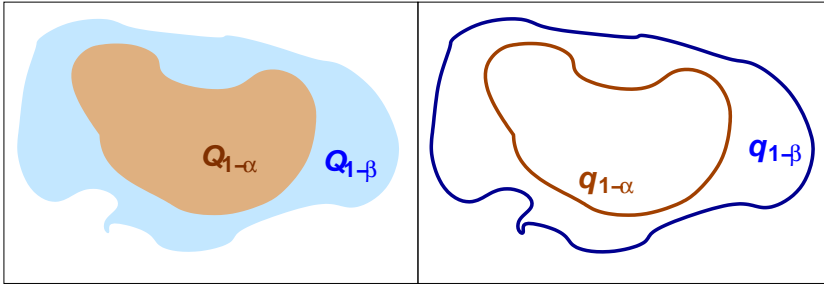
Jak je vidět z definice, a hluboké oblasti, $0 \leq a \leq d_{P,H}$ jsou *úrovňové množiny* hloubkové funkce H . Problematika úrovňových množin je velmi aktuální tematikou současné matematiky a i my se k ní v tomto článku stručně vrátíme o něco později.

Poznámka 1. *Pro jednorozměrnou náhodnou veličinu X a její rozdělení můžeme hloubku bodu x definovat*

$$H(x) = 1 - 2|F(x) - 1/2|, \quad (1)$$

kde $F(x)$ je distribuční funkce X . Hloubka mediánu \tilde{X} je zřejmě 1 a hloubka klesá v obou směrech od mediánu. Body se stejnou hodnotou distribuční funkce mají stejnou hloubku a a hlubokou oblastí je množina

⁴Též můžeme říkat uspořádání ve smyslu vnitřní/vnější, nebo centrální/odlehlý, ...



Obrázek 2: Centrální oblast (vlevo) a kvantilová kontura. $0 < \beta < \alpha < 1$.

$$R(a) = \{x : a/2 \leq F(x) \leq 1 - a/2\}.$$

V případě spojitě a rostoucí distribuční funkce je

$$c(a) = \{F^{-1}(a/2), F^{-1}(1 - a/2)\},$$

a zřejmě

$$P[R(a)] = 1 - a, \quad \text{pro všechna } a \in [0, 1]. \quad (2)$$

Obecně samozřejmě nemusí platit, že pro zvolenou hloubku H je hloubka nejhlubšího bodu rovna 1 pro všechna $P \in \mathcal{P}$. Nemusí ani platit (a pro většinu klasických hloubek opravdu neplatí) vztah (2) $P[R_{P,H}(a)] = 1 - a$. Proto je nezbytné definovat centrální oblasti a kvantilové kontury rozdělení P v hloubce H .

Definice 3. *Definujeme*

$$h(1 - \alpha) = h_{P,H}(1 - \alpha) = \sup\{a : P[R_{P,H}(a)] \geq 1 - \alpha\}.$$

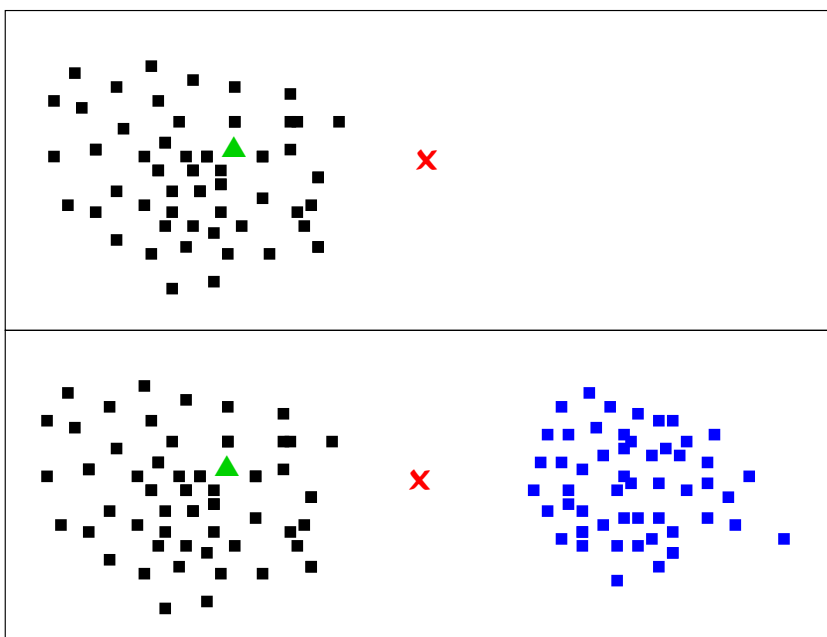
Oblast $Q_{P,H}(1 - \alpha) = R_{P,H}(h_{P,H}(1 - \alpha))$ nazveme $1 - \alpha$ centrální oblastí.

Hranici $q_{P,H}(1 - \alpha)$ oblasti $Q_{P,H}(1 - \alpha)$ nazveme $1 - \alpha$ kvantilovou konturou.

Jak je vidět z definice oblasti $Q(1 - \alpha)$ a hodnoty $h(1 - \alpha)$, bude k jejich určení obvykle nezbytné najít *všechny* úrovně množiny hloubkové funkce H a jejich pravděpodobnosti. To pochopitelně *není* snadný úkol. Později si ukážeme, že existuje možnost definovat specifickou hloubku tak, aby hodnota hloubky a pravděpodobnost a hluboké oblasti byly svázané předpisem

$$P[H(P, X) \geq \alpha] = 1 - \alpha \quad \text{pro všechna } \alpha \in [0, 1], \quad (3)$$

čímž se velmi usnadní hledání požadovaných $1 - \alpha$ centrálních oblastí.



Obrázek 3: Je bod ▲ hluboko? Je bod × hluboko?

Klasický jednorozměrný kvantil umožňuje vytvářet centrální množiny ve tvaru intervalů. Někdy jsou upřednostňovány úrovně množiny hustoty, které mohou být tvořeny sjednocením intervalů. Ve vyšších rozměrech se nám však, díky bohatší geometrii, otvírá svět rozmanitějších⁵ tvarů. Čím více rozměrů, tím bohatší tvary centrálních oblastí můžeme uvažovat. Přesto se dnes hlavní proud výzkumu v oblasti hloubky omezuje jen na hloubkové funkce, pro něž jsou centrální oblasti (úrovně množiny) konvexní.

Je dobré si připomenout, že každá redukce informace o rozdělení, v našem případě do několika centrálních oblastí, nám jiným způsobem poukazuje na podstatu náhody v uvažovaném modelu (pokusu). Každá hloubka svým způsobem zvýrazní určité rysy rozdělení, zatímco ostatní vlastnosti rozdělení hodnotu hloubky v podstatě neovlivní. Každá hloubka má také své výpočetní aspekty; jak při výpočtu teoretické hloubky bodu pro dané rozdělení (vesměs musíme používat numerické metody), tak i při výpočtu hloubky bodů v náhodném výběru. Podívejme se na obrázek 3. Na obou obrázcích bychom o bodu ▲ prohlásili, že je hluboko. Bod × je na horním obrázku zřejmě mimo data, co však na dolním?

Uděláme-li si na dolním diagramu v obrázku 3 marginální rozdělení (projekce do os), bude v obou případech bod × blízko mediánu. A dopadneme tak

⁵Mnohdy také jen těžko představitelných.

i v případě projekce na libovolnou přímku. Není proto možné jednoznačně říci, zda \times je či není hluboko. Tento příklad nám ukazuje, že ponětí o tom, který bod je hluboko a který není hluboko v rozdělení, závisí na tom, co od hloubky očekáváme.

3 Tisíc podob hloubky

Příklad 1. Hloubka (1) jednorozměrné náhodné veličiny X definovaná pomocí distribuční funkce F_X jako $H(P, x) = 1 - 2|F(x) - 1/2|$ pro $x \in \mathbb{R}$ je funkcí bodu a pravděpodobnostní míry. Na druhou stranu je zřejmé, že pravděpodobnostní míra popisuje vnější prostředí, jinak do definice hloubkové funkce nepřispívá. Nemění totiž samotnou podstatu hloubky. Pro všechna rozdělení je koncept H shodný.

Motivační příklad 1 nás navádí k tomu, že hloubková funkce by měla vycházet z nějaké obecné myšlenky, přičemž konkrétní pravděpodobnostní rozdělení tento koncept nemění. Jak by takový přístup k hloubkové funkci měl vypadat? Serfling v práci [45] navrhuje řadu *žádoucích vlastností* hloubky, potažmo hloubkové funkce.

3.1 Serflingův průvodce vlastnostmi hloubky

- I. *Afinní invariance hloubky.* Označme A nějakou afinní transformaci \mathbb{R}^p . Má platit $H(P \circ A, A(x)) = H(P, x)$. Hloubka nezávisí na souřadnicovém systému, hloubkové a kvantilové kontury jsou afinně ekvivariantní.
- II. *Maximalita ve středu.* Je-li rozdělení P v nějakém smyslu symetrické kolem $s \in \mathbb{R}^p$, pak $H(P, s)$ je maximální. Neboli s je nejhlubším bodem pro P .
- III. *Přenos symetrie.* Je-li rozdělení P v nějakém smyslu symetrické kolem $s \in \mathbb{R}^p$, pak $H(P, x)$ je odpovídajícím způsobem symetrické kolem s .
- IV. *Pokles podél paprsků.* Je-li s nejhlubší bod pro P , pak funkce $H(P, x)$ je nerostoucí na všech polopřímkách s počátkem v s (ve směru od s).
- V. *Zanedbatelnost v nekonečnu.* Platí $H(P, x) \rightarrow 0$ kdykoliv $\|x\| \rightarrow \infty$.
- VI. *Spojitosť v prostoru.* Zobrazení $x \mapsto H(P, x)$ je spojitě, nebo shora polospojité. Díky této vlastnosti jsou úrovněvé množiny H a tedy i a hluboké oblasti a $1 - \alpha$ centrální oblasti uzavřené.
- VII. *Spojitosť na rozděleních.* Zobrazení $P \mapsto D(P, x)$ je spojitě vzhledem ke slabé konvergenci měr.
- VIII. *Kvazikonkavita.* Úrovněvé množiny (a hlubkové oblasti, $1 - \alpha$ centrální oblasti) jsou konkávní.

Uvedeným vlastnostem se budeme v článku průběžně věnovat. V tuto chvíli se pozastavíme jen u požadavků na symetrii. V bodech II a III se hovoří o symetrii *v nějakém smyslu*, jejich splnění je proto závislé na formě uvažované symetrie⁶. V bodě III je dobré specifikovat, jak se symetrie rozdělení P má přenést na (a na jakou) symetrii hloubkové funkce H .

3.2 Račte si vybrat ...

Na příkladech známých i méně známých hloubek si ukážeme si, že v jednotlivých definicích lze vysledovat několik základních přístupů k mnohorozměrným rozdělením, na kterých jsou hloubkové funkce založené. Kromě definice a některých základních vlastností se pokusíme alespoň trochu pohled⁷ na silné a slabé stránky studované hloubky.

3.2.1 Poloprostorová hloubka Sluší se začít neklasičtější ze všech hloubek (matkou hloubek). Jde o koncept, který představil ve svém článku [47] John Tukey v roce 1975.

Definice 4. *Bud' P rozdělení na \mathbb{R}^p . Definujme poloprostorovou hloubku⁸ bodu x jako*

$$H(P, x) = \inf\{P(D) : D \text{ je uzavřený poloprostor, } x \in D\}. \quad (4)$$

Zřejmě stačí uvažovat poloprostory, pro něž je x na hranici.

Tato definice je docela názorná. Jde o poměrně jednoduché zobecnění kvantilu. V jednorozměrném případě máme pouhé dva směry; hodnoty větší než x a hodnoty menší než x . V mnohorozměrném případě je směrů pocho-pitelně nespočetně.

V dvourozměrném prostoru lze poloprostorovou hloubku názorně definovat také pomocí *projekcí* na přímku. Každá přímka p procházející bodem x je tímto bodem rozdělena na dvě polopřímky p_1 a p_2 . Označme P_p rozdělení vzniklé projekcí pravděpodobnostní míry P na přímku p . Hloubka bodu x je v tomto případě

$$H(P, x) = \inf\{\min\{P_p(p_1), P_p(p_2)\} : p \text{ je přímka, } x \in p\}^9.$$

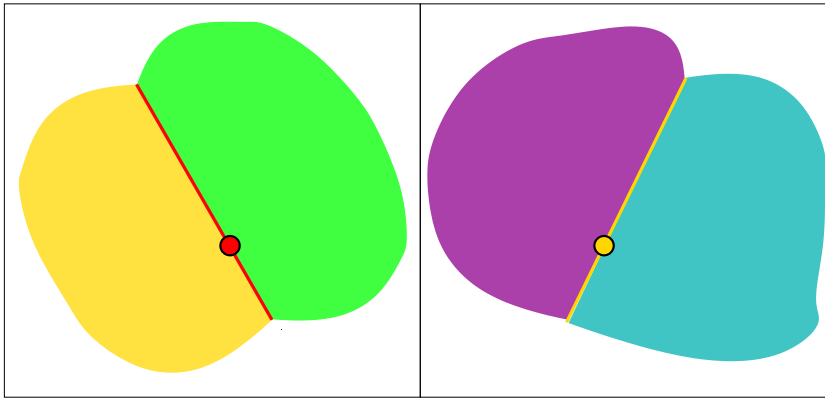
Je zřejmé, že v této definici nijak nezohledňujeme vzdálenosti ostatních bodů od x a svým způsobem ani jejich polohu vůči x . Tím se někdy zbytečně ochuzujeme o důležitou informaci. Najdeme-li totiž poloprostor D takový,

⁶Symetrií lze uvažovat celou řadu; může jít o klasickou středovou symetrii, osovou symetrii, středově-angulární symetrii, eliptickou symetrii, ale třeba i o symetrii vůči otáčení atd.

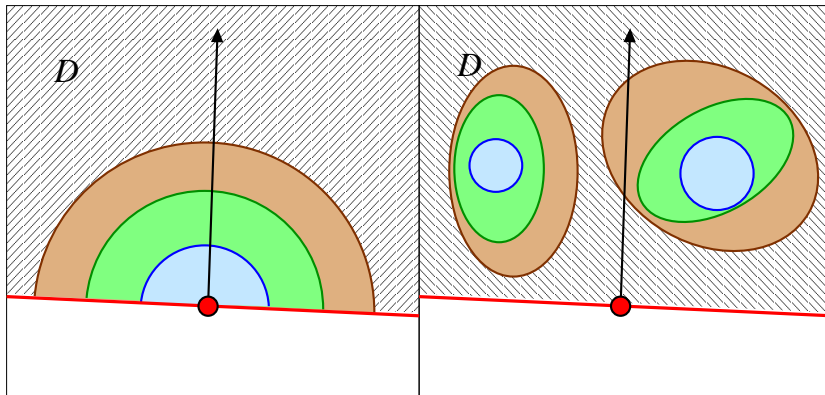
⁷značně subjektivní

⁸Anglicky *halfspace depth*

⁹Čtenář si může sám domyslet, co by nastalo, kdyby místo *všech* přímek obsahujících x uvažoval pro každý bod jen dvě přímky rovnoběžné s osami.



Obrázek 4: Hledání poloprostorové hloubky.



Obrázek 5: Poloprostorová hloubka využívá jen částečně tvar nosiče rozdělení.

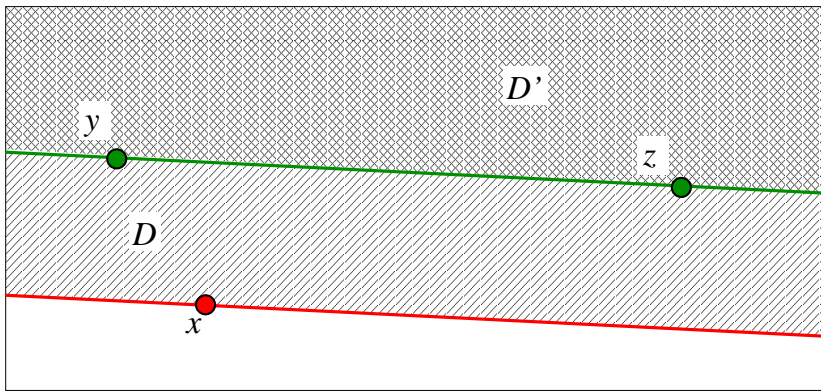
že $x \in D$ a $P(D)$ je minimální, nevíme, jak vypadá rozdělení na D . Je rozdíl, je-li rozdělení na D symetrické kolem normálového vektoru, nebo velmi nesymetrické (viz obrázek 5).

Na druhou stranu, pro náhodný výběr (empirickou míru) platí, že znalost *všech* kvantilových kontur charakterizuje tuto empirickou míru [46].

Zaměříme se na některé základní vlastnosti poloprostorové hloubky. Afinní invariance vyplývá z faktu, že poloprostor po afinní transformaci zůstane poloprostorem. Vzájemná poloha bodů se tak, z *pohledu definice poloprostorové hloubky*, nemění.

Nechť absolutně spojitě rozdělení P je středově symetrické kolem středu s^{10} , pak s je nejhlubším bodem. Jakýkoliv uzavřený podprostor na jehož hranici leží s musí mít pravděpodobnost $1/2$, takže hloubka $H(P, s) = 1/2$;

¹⁰Platí tedy $P(s + A) = P(s - A)$ pro každý náhodný jev A .



Obrázek 6: Poloprostorová hloubka klesá směrem od nejhlubšího bodu a je kvazikonkávní.

větší hloubku nemůže mít žádný bod pro libovolné spojitě rozdělení. Pro spojitá rozdělení se souvislým nosičem lze dokázat i jednoznačnost nejhlubšího bodu. V diskrétním případě (speciálně pro empirické míry) se může stát, že existuje konvexní množina nejhlubších bodů.

Pro středově symetrická rozdělení platí symetrie hloubky kolem středu s , neboli

$$H(P, (s + x)) = H(P, (s - x)) \text{ pro všechna } x \in \mathbb{R}^p.$$

Monotonie a kvazikonkavita poloprostorové hloubky jsou také přímým důsledkem používání poloprostorů. Na obrázku 6 je znázorněn bod x a D , poloprostor s nejmenší pravděpodobností, na jehož hranici x leží. Zřejmě

$$\max\{H(P, y), H(P, z)\} \leq P(D') \leq P(D) = H(P, x) \leq 1/2$$

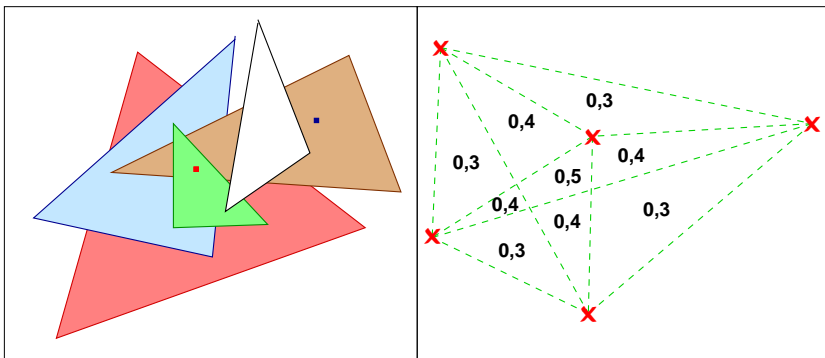
a pro spojitá rozdělení se souvislým nosičem dokonce platí ostré nerovnosti

$$H(P, x) > 0 \Rightarrow \max\{H(P, y), H(P, z)\} < H(P, x).$$

Tím jsme dokázali pokles hloubky podél polopřímek s počátkem v nejhlubším bodě. Podobné úvahy pomohou i k důkazu konvexity centrálních oblastí.

Poloprostorová hloubka splňuje další z požadovaných vlastností, konzistenci. Označíme-li $H(P_n, x)$ hloubku bodu x vzhledem k empirické míře P_n , platí $H(P_n, x) \rightarrow H(P, x)$ skoro jistě. Čtenáře odkazujeme na klasický článek [11], díky kterému se poloprostorová hloubka stala velice populární. V uvedené práci je diskutována také robustnost poloprostorové hloubky a další zajímavé pokročilé vlastnosti.

Velkou nevýhodou poloprostorové hloubky zůstává její výpočetní náročnost. I počítání hloubky byt' jen jediného bodu je s rostoucí dimenzí a s rostoucím počtem pozorování časově velmi náročné. Nejčastěji používané algoritmy vyvinul Rousseeuw se spolupracovníky [42, 43].



Obrázek 7: Náhodné simplex v \mathbb{R}^2 a simplexová hloubka.

Na druhou stranu zajímavé rozšíření poloprostorové hloubky přináší *regresní hloubka*. Místo hloubky bodů je definována (poloprostorová) hloubka regresní přímky a tím i *nejhlubší regrese* spolu s centrálními oblastmi regresních přímek. Jde o zajímavý alternativní robustní přístup k regresním úlohám. Detailní rozbor regresní hloubky lze nalézt v článku [44].

3.2.2 Simplexová hloubka Tato hloubka se objevuje v článkách Reginy Y. Liu(ové) [29, 30]. Jde o hloubku, kterou můžeme zařadit do třídy hloubek založených na pravděpodobnosti pokrytí bodu *náhodnou oblastí určitého tvaru*. V případě simplexové hloubky je tím předepsaným tvarem *simplex*, čili konvexní obal $p + 1$ bodu $x_1, \dots, x_{p+1} \in \mathbb{R}^p$.

Definice 5. *Bud' P rozdělení na \mathbb{R}^p . Náhodným simplexem rozumíme útvar vzniklý konvexním obalením bodů $V_i, i = 1, \dots, p + 1$, kde V_i jsou nezávislé náhodné vektory s rozdělením P . Definujme simplexovou hloubku¹¹ bodu x jako*

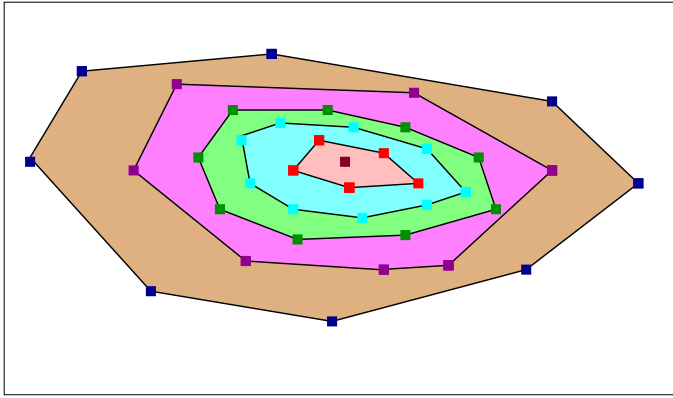
$$H(P, x) = P[x \in S], \text{ kde } S \text{ je náhodný simplex.} \quad (5)$$

Místo simplexu můžeme brát i jiné útvary definované nezávislým výběrem k bodů z rozdělení P . Tak můžeme uvažovat náhodné koule, elipsoidy, mnohoúhelníky.

Náhodné simplex mají před ostatními tvary zásadní výhodu. Po afinní transformaci zůstává simplex simplexem, což například pro koule a elipsoidy neplatí. Proto hloubka založená na simplexech afinně invariantní je, zatímco ostatní obdobné hloubky afinně invariantní *být nemusí* (a typicky nejsou).

Ačkoliv jsou hloubky založené na pravděpodobnosti pokrytí náhodným tvarem, podobně jako poloprostorová hloubka, velmi názorné, jejich vlastnosti se tak snadno nedokazují. Asymptotické chování simplexové hloubky

¹¹Anglicky *simplicial depth*



Obrázek 8: Loupání slupek.

je shrnuto v článku [13]. Simplexová hloubka je kvazikonkávní, konzistentní, afinně invariantní, spojitá a má spoustu dalších vlastností.

Na druhou stranu je simplexová hloubka výpočetně velmi náročná. V teoretických rozdělení je téměř nemožné ji určit. Pro empirickou míru jde o výpočetně náročný úkol. Možná i proto je simplexová hloubka poněkud ve stínu slavnější a oblíbenější poloprostorové hloubky. Přesto její vývoj neustává, jak lze zjistit například v článku [4].

3.2.3 Konvexní slupky Přísně vzato nejde o hloubku, protože neexistuje definice pro spojitá rozdělení, ale pouze pro výběry ze spojitých rozdělení, tedy pro diskrétní rovnoměrné rozdělení. Jak již název napovídá, *do hloubky dat* se vypravíme skrze odstraňování konvexních obalů výběru¹².

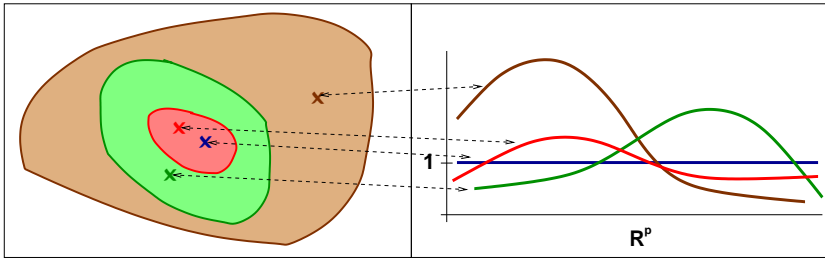
Definice 6. *Bud' P_n empirická míra příslušná náhodnému výběru X_1, \dots, X_n o rozsahu n z rozdělení P . Konvexní obal hloubky $1/n$, označme jej C_1 , je konvexním uzávěrem extrémálních bodů množiny $\mathbf{X} = \{X_1, \dots, X_n\}$. Konvexní obal hloubky k/n , $2 \leq k \leq n$, označme jej C_k , je konvexním uzávěrem extrémálních bodů množiny $\text{int}C_{k-1}$, kde $\text{int}C$ označuje (relativní) vnitřek množiny C . Definujme*

$$H(P_n, x) = k/n \iff x \in C_k \setminus C_{k+1}. \quad (6)$$

Přechod na obecná diskrétní rozdělení, nebo na teoretické spojitá rozdělení není u konvexních slupek jasný. To je ovšem podstatná nevýhoda. Slabou útechou může být, že se výpočetně jedná o výrazně jednodušší verzi hloubky, než jsou simplexová hloubka a poloprostorová hloubka.

Z podstaty definice plyne, že centrální oblasti vymezené konvexními slupkami jsou konvexní. Konvexní slupky jsou také afinně ekvivalentní. O kon-

¹²Odtud se vzal anglický název *convex hull peeling*.



Obrázek 9: Zonoidy a jejich definující funkce.

zistenci se, pochopitelně, hovořit nedá. Podobně problematické je mluvit o symetrii, protože náhodný výběr ze symetrického rozdělení tvoří symetrickou empirickou míru s pravděpodobností nula¹³.

3.2.4 Zonoidy Jde o příklad hloubky, která místo mediánu zobecňuje střední hodnotu. Zonoidy definovali Koshevoy a Mosler v článku [24].

Definice 7. *Bud' $\alpha \in [0, 1]$. Množina*

$$Z_\alpha(P) = \{x \in \mathbb{R}^p : \exists g : \mathbb{R}^p \rightarrow [0, 1/\alpha], E_P g(X) = 1 \text{ a } E_P X g(X) = x\} \quad (7)$$

se nazývá α zonoidem.

Zonoidová hloubka bodu x je definována

$$H(P, x) = \inf\{\alpha : x \in Z_\alpha(P)\} \quad (8)$$

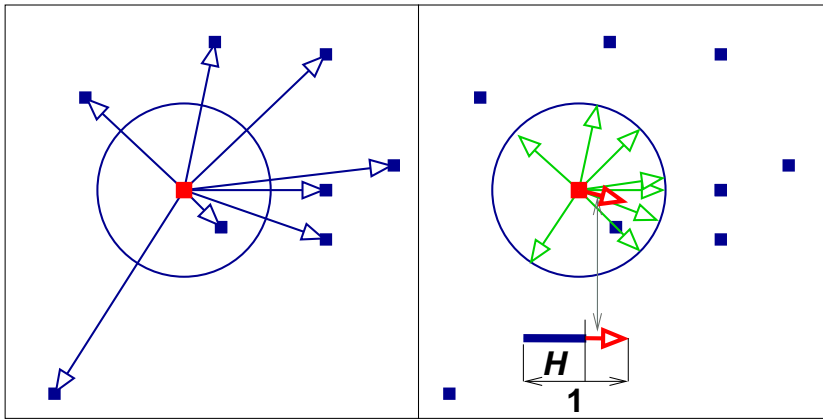
Poprvé se setkáváme s nerobustní definicí hloubky. V definici zonoidu předpokládáme existenci prvního momentu EX . Existuje-li konečná střední hodnota, je vždy nejhlubším bodem. Neexistuje-li konečná střední hodnota, nelze obecně zonoidy definovat.

Pro $\alpha = 1$ je zřejmě jedinou volbou $g \equiv 1$ s.j. a $Z_1 = EX$. Je-li $\alpha_1 < \alpha_2$, pak zjevně (viz obrázek 9) platí $Z_{\alpha_1} \supset Z_{\alpha_2}$. Snadno lze nahlédnout, že zonoidy jsou konvexní. Stačí použít konvexní kombinaci $ag_1(x) + (1-a)g_2(x)$ nějakých dvou funkcí $g_1, g_2 : \mathbb{R}^p \rightarrow [0, 1/\alpha]$. Zonoidová hloubka je tedy kvazikonkávní a navíc, jak lze ověřit přímo z definice, afinně invariantní.

Empirická zonoidová hloubka se dá poměrně rychle spočítat pro konkrétní bod. Je ale mnohem složitější určit vlastní zonoid pro dané α . Algoritmus pro dimenzi $p = 2$ lze najít v [14], jeho výpočetní náročnost pro rozsah výběru n je $O(n^2 \log n)$. Pro vyšší dimenze jde prakticky o nemožný úkol, stejně jako pro teoretická rozdělení.

Další vlastnosti a aplikace zonoidů lze najít v článku [38] či v knize [37].

¹³Je samozřejmě pravda, že symetrická empirická míra vede na symetrickou hloubku—taková situace ale asi v historii lidstva nebyla pozorována



Obrázek 10: Určování směrů a výpočet L_1 hloubky.

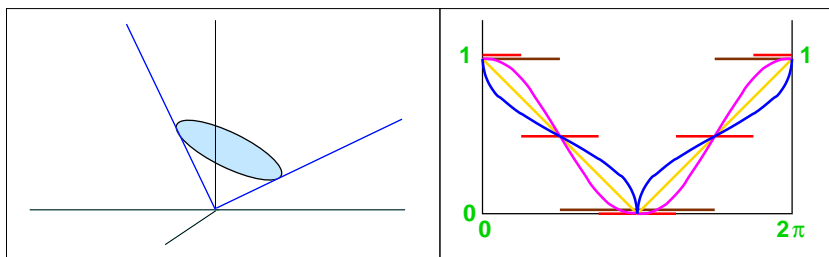
3.2.5 L_1 hloubka Výpočetní náročnost spojená s předchozími definicemi hloubky a s tím spojené obtíže s určováním centrálních oblastí vedly k definicím nových hloubek. Ty často nesplňují úplně všechny klasické požadavky na hloubku (viz část 3.1), ale zato jsou výpočetně mnohem snazší. Jednou z nich je L_1 hloubka uvedená v článku [49]. Podobně jako v případě poloprostorové hloubky nezávisí L_1 hloubka bodu x na vzdálenostech ostatních bodů. Na rozdíl od poloprostorové hloubky ale závisí na jejich vzájemném postavení.

Na obrázku 10 vidíme postup pro určení empirické L_1 hloubky. Formálně L_1 hloubku definujeme pomocí délky *průměrného směřového vektoru* k ostatním bodům.

Definice 8. Pro zvolený bod $x \in \mathbb{R}^p$ určíme střední hodnotu $E(X - x) / \|X - x\|$. L_1 hloubkou bodu x nazveme hodnotu

$$H(P, x) = 1 - \|s\|, \text{ kde } s = E \frac{X - x}{\|X - x\|} = \int_{\mathbb{R}^p} \frac{u - x}{\|u - x\|} P(du). \quad (9)$$

Výpočet hloubky je velmi snadný. Není nutné pro každý bod procházet *všechny* možnosti, stačí spočítat střední hodnotu jednoduché transformace náhodného vektoru. Asi nejviditelnější ztrátou za tuto jednoduchost je ztráta afinní invariance hloubky. Po afinní transformaci se může změnit vzájemné postavení bodů; v tomto případě jde o změnu úhlů mezi dvojicemi vektorů. Další, možná ještě horší, vlastností L_1 hloubky je fakt, že ani vně konvexního uzávěru nosiče rozdělení není hloubka nulová. Nulová hloubka znamená, že s pravděpodobností jedna *všechny* hodnoty náhodného vektoru leží *v jediném směru* od uvažovaného bodu. To ale obvykle není možné, takže nulovou hloubku obvykle nemá žádný bod a platí pouze limitní vztah $H(x) \rightarrow 0$,



Obrázek 11: Zobecnění kvantilu na mnohorozměrný M-kvantil. Vlevo ztrátová funkce, vpravo ukázky volby směrových funkcí pro \mathbb{R}^2 .

jestliže $\|x\| \rightarrow \infty$. Velmi nevhodná je tato vlastnost pro náhodné vektory, jejichž nosič není celý prostor.

Pro empirické míry mohou centrální oblasti L_1 hloubky být i nekonvexní. Rozdíl mezi centrální oblastí a jejím konvexním obalem je obvykle zanedbatelný. Proto autoři této metody doporučují použít místo centrálních oblastí jejich *konvexní obaly*¹⁴.

3.2.6 M-kvantily V souvislosti s mnohorozměrnými kvantily se nabízí zobecnění známé definice jednorozměrného kvantilu. Zavedme funkci

$$\varrho(x) = (1 - \alpha)l(x < 0)|x| + \alpha l(x > 0)|x|, x \in \mathbb{R}.$$

Hodnota q_α minimalizující výraz $E\varrho(X - q_\alpha)$ je α kvantilem rozdělení náhodné veličiny X ¹⁵. Připomeňme, že q_α není definováno jednoznačně, ale může jít jak o jediný bod, tak o interval.

Funkci ϱ zobecníme do více rozměrů. Cílem je získat jakýsi „nakloněný polokružel“, obdobu funkce ϱ . Místo indikátorů $l(x < 0)$ a $l(x > 0)$ budeme uvažovat *směrové funkce* $\phi_s(x)$, $s \in \mathbb{R}^p$, $\|s\| = 1$. Funkce $\phi_s(x)$ je funkcí úhlu mezi x a s .

Také vzdálenost l v \mathbb{R}^p můžeme volit různě, nejčastěji se vyskytují \mathcal{L}_1 , \mathcal{L}_2 a \mathcal{L}_∞ vzdálenosti.

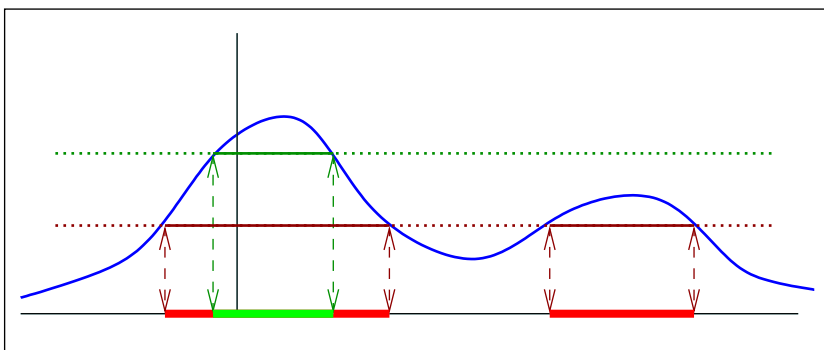
Definujme α M-kvantil pro $\alpha \in [0, 1/2]$ *ve směru* s jako libovolnou hodnotu $q_{\alpha,s}$ minimalizující

$$q_{\alpha,s}(P) = \arg \min_{q \in \mathbb{R}^p} E_P [(1 - 2\alpha)\phi_s(X - q) + \alpha]l(X - q).$$

Volbou $\alpha = 1/2$ získáme mnohorozměrný medián \tilde{x} , se zmenšující se hodnotou α se pro pevné s α M-kvantil vzdaluje od mediánu. Připomeňme, že takto definovaný medián (a potažmo M-kvantily) není afinně ekvivariantní.

¹⁴Jak uvidíme dále, s touto nucenou konvexitou se nedokážeme úplně ztotožnit. Stejně jako s *žádoucí* vlastností hloubky VIII—v tom ale nejsme ojedinelí.

¹⁵Existuje-li první moment náhodné veličiny X



Obrázek 12: Úroňové množiny.

Na M-kvantily je navázána hloubka

$$H(P, x) = \sup\{\alpha : x \notin \{q_{\beta,s}(P); \|s\| = 1, \beta \geq \alpha\}, \alpha \in [0, 1/2]\}. \quad (10)$$

M-kvantily se pravděpodobně poprvé objevují v práci [2]. Poté jsou rozvíjeny v článkách [7, 3] a také u nás v příspěvku [19] a v diplomové práci [17]. Velmi důkladný rozbor teoretických vlastností M-kvantilů a obecnějších M-parametrů mnohorozměrných rozdělení lze najít v článku [23].

Připomeňme, že M-kvantil je funkcí směru s a necentrality α . Praktické pokusy ukazují, že empirické M-kvantily $q_{\alpha,s}$ pro pevné s a pro $\alpha \in [0, 1/2]$ neleží na přímce od středu. V mnoha případech existují dva různé směry s a t a hodnota α takové, že empirické kvantily \tilde{x} , $q_{\alpha,s}$ a $q_{\alpha,t}$ leží na přímce, tedy kvantilová kontura může být zacyklená.

3.2.7 Množiny minimálních objemů Poslední z mnoha definic hloubek, které si ukážeme je založena na *úroňových množinách*. Připomeňme, že pro libovolnou funkci $f : \mathbb{R}^p \rightarrow \mathbb{R}$ definujeme úroňové množiny¹⁶ funkce f

$$L_a = L_a(f) = \{x \in \mathbb{R}^p : f(x) \geq a\}, a \in \mathbb{R}.$$

Hlavní výhoda úroňových množin je zřejmá. Buď X p -rozměrný náhodný vektor s absolutně spojitým rozdělením P s hustotou $\psi(x)$. Uvažujme-li úroňovou množinu $L_a(\psi)$ hustoty ψ , pak platí

$$\lambda_p(L_a) = \min\{\lambda_p(B); P(B) = P(L_a)\},$$

kde λ_p je p -rozměrná Lebesgueova míra. Proto hovoříme o *množinách minimálního objemu*¹⁷ v nějaké třídě množin, v tomto případě mezi množinami s danou pravděpodobností.

Hloubka bodu může být definována jednoduše jako hodnota hustoty v daném bodě. Je však vhodnější definovat hloubku tak, aby byla v rozmezí $[0, 1]$.

¹⁶Anglický termín je *level-sets*

¹⁷Minimum volume sets

Definice 9. *Bud' P absolutně spojitě rozdělení pravděpodobností s hustotou ψ . Označme $L_a, a \geq 0$ úrovně množiny hustoty ψ . Definujme hloubku*

$$H(P, x) = 1 - \min\{P(L_a); x \in L_a\} = 1 - P(L_{\psi(x)}).$$

Má-li rozdělení P jeden modus (případně více stejně velkých modů, které dohromady mají pravděpodobnost 0), pak modus je nejhlubším bodem a jeho hloubka je 1. Body mimo nosič hustoty mají zřejmě hloubku 0. Jak je zřejmé i ze schematického obrázku 12, nemusí být centrální množina minimálního objemu ani souvislá, natož konvexní.

Empirická verze úrovněových množin je většinou založena na vhodném odhadu hustoty. Od vlastností těchto odhadů se odvíjejí i vlastnosti empirické hloubky, zejména možná afinní invariance a asymptotické vlastnosti. Tomuto problému byla věnována řada prací, například [1, 40, 35].

3.3 Dvě tváře hloubky: třídění podle definice a podle vlastností

Hloubková funkce může být založena na několika různých nápadech, přičemž jsou vždy zdůrazněny jiné globální a jiné lokální vlastnosti rozdělení. Mezi jednotlivými přístupy k hloubce nelze zcela libovolně přecházet, jedna není speciálním případem druhé, těžko by se hledalo zobecnění zastřešující různé přístupy k hloubce. Dá se tedy říci, že při definici hloubky můžeme sledovat pravděpodobnost nějakých útvarů s referenčním bodem x (poloprostorová hloubka), pravděpodobnost, že x je pokryto nějakou množinou (simplexová hloubka), polohu x vůči ostatním bodům (L_1 hloubka, konvexní slupky), řešit optimalizační úlohu (M-kvantily), vázat hloubku přímo na hustotu v daném bodě (úrovněové množiny), nebo si vymyslet spoustu dalších konceptů. V každé této třídě lze volit různé varianty základní definice. Místo náhodných simplexů může jít o náhodné elipsoidy, místo poloprostorů o jiné útvary (kužele, pásy, ...), zohlednit vzdálenost, zavést váhy.

Kromě roztržidění hloubek na základě jejich definice lze také posuzovat jejich jednotlivé vlastnosti. Pro stejnou třídu hloubek se chování hloubkové funkce může velmi lišit podle zvolené varianty definice. Typickým příkladem bývá afinní invariance, méně často kvazikonkavita. Také konzistence empirické hloubky se odvíjí od faktu, zda použitá definice splňuje předpoklady některého ze stejnoměrných zákonů velkých čísel (pro asymptotickou normalitu pak nějaké funkcionální centrální limitní věty), které mohou být nalezeny například v [12, 48].

3.4 Jaké vlastnosti má mít ideální hloubka?

Výše jsme uvedli Serflingův seznam vlastností, které by měla splňovat každá hloubka, hloubková funkce.

Začneme bodem I, afinní invariancí. Vzhledem k tomu, že hloubka poukazuje na centrální oblasti a kvantilové kontury, což jsou parametry polohy mnohorozměrného rozdělení, je požadavek afinní invariance přirozený. Dodejme

ještě, že v jednorozměrném případě jsou kvantily ekvivariantní vůči ryze monotónním transformacím (potažmo na kvantilech založená hloubka (1) je invariantní). Obdoba takové vlastnosti není ve více rozměrech obecně možná¹⁸.

Afinní invariance až na výjimky platí pro otočení, překlopení, posunutí. Problém nastává, jestliže zvolená transformace mění různě měřítko jednotlivých os. Tím se totiž mění úhly mezi vektory a hloubka jejíž definice nějakým způsobem úhly mezi vektory zahrnuje, nemůže být vůči této transformaci invariantní. Typickým příkladem je L_1 hloubka, ale i M-kvantily.

Asi největším zklamáním je skutečnost, že přirozený kandidát na mnohorozměrný medián, hodnota minimalizující střední \mathcal{L}_1 odchylku $E\|X - a\|_1$ není afinně ekvivariantní¹⁹. Proto existují různé postupy, jak definovat afinně ekvivariantní medián založený na minimální střední \mathcal{L}_1 odchylce. Častým postupem je standardizace dat, výpočet mediánu a následná zpětná transformace, viz například [5].

Pokud tedy hloubková funkce není afinně invariantní, doporučuje se začít výpočet hloubky vhodnou standardizací dat²⁰ a určením kvantilových kontur pro standardizovaná data. Poté pomocí zpětné transformace určíme hloubku a kvantilové kontury pro původní rozdělení. Možných řešení je spousta, většina je založena na takové transformaci, aby standardizovaná data měla shodný parametr měřítka pro všechna marginální rozdělení a aby byla nekorelovaná. Této problematice se věnují například články [5, 6, 18].

Vlastnosti II a III hovoří o *nějaké* symetrii rozdělení kolem bodu $s \in \mathbb{R}^p$. Bod, kolem kterého je rozdělení symetrické má být nejhlubší. To je určitě rozumný požadavek. Někdy ale jsme před těžkou volbou, zda bod, který není bodem symetrie, ale přesto k němu má v nějakém smyslu blízko, má či nemá být nejhlubším bodem. Tato situace je ilustrována na obrázku 13. Obrázky znázorňují nosiče absolutně spojitěho rozdělení, přičemž předpokládáme, že hustota klesá stejně ve všech směrech od vyznačeného modu. V levém sloupci vidíme dvě situace, kde rozdělení není středově symetrické, ale přesto bychom oba vyznačené body mohli považovat za nejhlubší. V pravém sloupci obrázku 13 jsou dvě středově symetrická rozdělení pro která vyznačené body bez váhání označíme za nejhlubší.

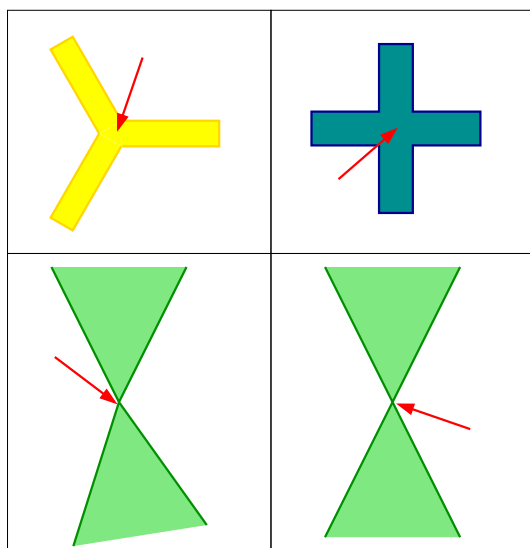
Vlastnost IV, klesání podél paprsků, probereme spolu s konvexitou, se kterou částečně souvisí. Zanedbatelnost hloubky v nekonečnu, vlastnost V, je pro rozdělení náhodného vektoru také přirozeným požadavkem.

Z našeho hlediska je nejvíce diskutabilní vlastnost VIII, kvazikonkavita hloubky. Jedná se o velmi oblíbenou vlastnost, která je jako pozitivní vyzdvihována u nejčastěji používaných hloubek. Je ale zřejmé, že pro úrovně množiny je konvexita centrálních oblastí dosažitelná jen ve speciálních případech. Takže *diktát konvexity* není tak úplný, jak se zdá.

¹⁸Opět se potýkáme s neexistencí lineárního uspořádání a tedy s nemožností definovat rozumně ryzí monotonii zobrazení z \mathbb{R}^p do \mathbb{R} .

¹⁹Afinně ekvivariantní není ani, na první pohled též přirozený, medián po složkách, tedy vektor marginálních mediánů.

²⁰Což bývá vhodné afinní vzájemně jednoznačné zobrazení



Obrázek 13: Nejhlubší bod a střed symetrie.

V práci [10] autoři navrhují uvažovat unimodální hvězdicovité rozdělení a nahradit konvexní konfidenční množiny množinami hvězdicovitými. Připomeňme, že absolutně spojitě unimodální rozdělení s modem v bodě 0 je hvězdicovité kolem nuly, jestliže pro jeho hustotu f platí

$$0 \leq s \leq t < \infty \Rightarrow f(tx) \leq f(sx)$$

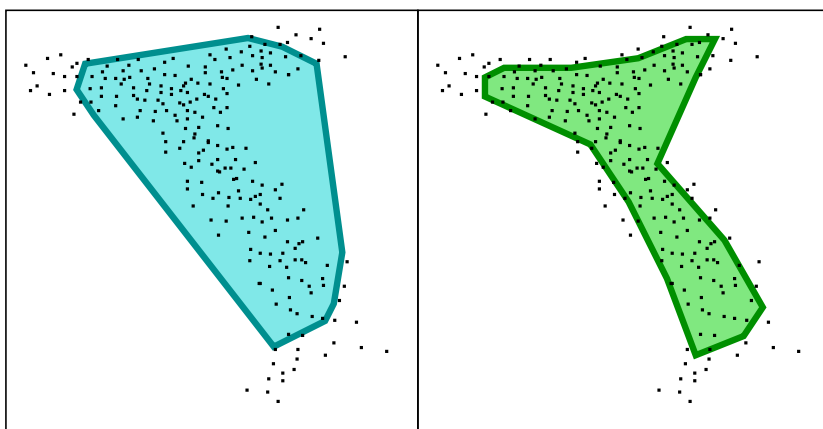
pro všechny body $x \in \mathbb{R}^p$. Množina M je hvězdicovitá vůči bodu x , centru hvězdicovité množiny M , jestliže platí

$$y \in M \Rightarrow \overline{xy} \subset M,$$

kde \overline{xy} značí úsečku s koncovými body x a y . Autoři [10] navrhují najít *nejvhodnější* tvar konfidenční množiny C mezi všemi hvězdicovitými množinami se společným centrem (modem rozdělení) a poté konstruovat centrální oblasti *zvětšováním a zmenšováním* optimální množiny C . Za zmínku stojí, že všechny takto zkonstruované konfidenční množiny mají stejný tvar.

Samozřejmě nechceme zpochybňovat oprávněnost konvexních centrálních oblastí. Jsou případy, kdy se jedná o nezpochybnitelnou volbu. Například pro mnohorozměrné normální rozdělení bychom čekali, že všechny definice hloubky povedou ke známým (a stejným) elipsoidům. Domníváme se ale, že by kvazikonkavita hloubky, tedy konvexita centrálních oblastí, měla být požadována *jen v odůvodněných případech*²¹. Na obrázku 14 je znázorněno,

²¹Jinak jsme sváděni k domněnce, že kvazikonkavita hloubky je pouze z *nouze ctnost*, protože vychází u všech hloubek z *hlavního proudu*



Obrázek 14: Konvexní a hvězdovitá množina.

jak nuceně konvexní tvary centrálních oblastí mohou obsahovat zbytečně mnoho *hluchých* míst. V určitých případech nezbývá než uznat, že mnohorozměrné prostory nabízejí bohatší geometrii než jednorozměrný. Proto se přidáváme k názoru, že je vhodné uvažovat bohatší množinu tvarů, než jsou konvexní množiny.

Opustíme-li tedy požadavek konvexity centrálních oblastí, musíme se zamyslet zda tento požadavek nějak nahradíme. DasGupta a spol. navrhnou používat hvězdicovité množiny, tedy zachovat alespoň vlastnost IV, pokles hloubky podél paprsků (navrhují ale použít jeden tvar množiny pro všechny centrální oblasti daného rozdělení, což je hodně omezující požadavek vzniklý snad z výpočetně-optimalizačních důvodů). Na druhé straně stojí úrovně množiny, které mohou mít zcela libovolný tvar, nemusí být ani souvislé, nemusí být dokonce ani otevřené, ani uzavřené (nesplňují tedy apriori požadavky IV, VI a VIII). Z analogie k jednorozměrným kvantilům se přirozeným požadavkem ukazuje *souvislost* centrálních oblastí, alespoň v případech unimodálních rozdělení²². V každém případě pro obecná rozdělení se nezdá být vhodné omezovat možné tvary centrálních oblastí. Spíše je podstatné vybrat hloubkovou funkci, která z rozdělení *zachová a zvýrazní* ty vlastnosti, které zkoumáme²³.

4 Směrové hloubky

Chceme-li rozšířit možné tvary centrálních oblastí na hvězdicovité množiny, musíme se zamyslet nad definicí hloubky, která tuto vlastnost zaručí. Výše jsme uvedli, že v článku [10] se hledá optimální tvar hvězdicovité množiny

²²Tato vlastnost je pro unimodální rozdělení úrovněvými množinami pochopitelně splněna.

²³Volba metody je dost subjektivní—to by nás nemělo ve statistice překvapovat.

pro dané rozdělení. Existuje však definice hloubky, která přímo vede na hvězdicovité množiny a navíc α centrální oblasti mohou mít pro různá α velmi odlišné tvary. Jedná se o směrovou hloubku.

4.1 Směrová hloubka na přímkách

Ve svém článku [52] definuje Wei směrové kvantily pomocí rozdělení na přímkách procházejících *středem*.

Definice 10. *Uvažujme absolutně spojitě rozdělení pravděpodobnosti P na \mathbb{R}^p a jeho hustotu f . Buď S mnohorozměrný medián rozdělení P . Pro libovolnou přímku l tvaru $l = S + tx$, $t \in \mathbb{R}$, kde $x = (x_1, x_2, \dots, x_p)$ je vektor délky 1 a $x_p \geq 0$ (l prochází bodem S , x směřuje na „sever“) definujme podmíněnou hustotu $f_l(t)$ na \mathbb{R} jako hustotu podmíněného rozdělení reálného parametru t za podmínky, že $X = S + tx$ (X leží na přímce l). Pro (jednorozměrné) rozdělení s hustotou f_l najdeme $\alpha/2$ kvantil $q_{\alpha/2}(l)$ a $1-\alpha/2$ kvantil $q_{1-\alpha/2}(l)$. Body $S + q_{\alpha/2}(l)x$ a $S + q_{1-\alpha/2}(l)x$ tvoří α směrové kvantily pro přímku l . Množina všech směrových α kvantilů tvoří α kvantilovou konturu.²⁴*

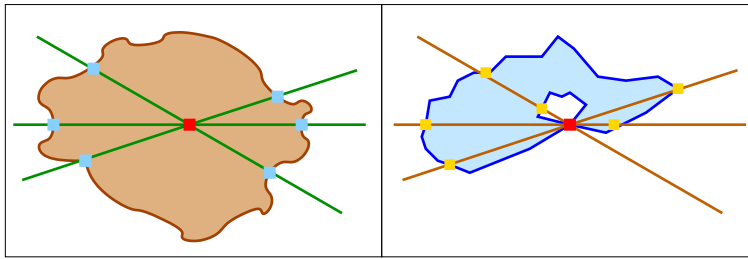
Všimněme si, že směrové kvantily samy o sobě *nedefinují střed*. Ten musí být dodán zvenčí.

Samotná definice směrových kvantilů, tedy pomocí podmíněných rozdělení na přímkách, je afinně invariantní. Přímka zůstane přímkou a pořadí bodů na přímce se buď zachová, nebo převrátí. Pro afinní invarianci je tedy rozhodující afinní invariance navrženého středu. Wei navrhuje použít mnohorozměrný medián, který obecně afinně invariantní není, proto doporučuje nejprve provést standardizace diskutované výše.

Směrové kvantily na přímce mají jednu zásadní nevýhodu. Pro jednotlivé přímky je medián $q_{1/2}(\cdot)$ totožný se středem S jen velmi výjimečně (v případě středově symetrických rozdělení). Může se proto snadno stát, že oba kvantily, $q_{\alpha/2}(l)$ a $q_{1-\alpha/2}(l)$ leží vůči středu S na stejné straně přímky l . Tím dochází k nevhodnému *zacyklení* kvantilových kontur, což je jen velmi obtížně vysvětlitelný fenomén. Zejména postavení středu S je v tomto případě značně podivné.

Další otázkou je, jak počítat směrové kvantily pro empirická rozdělení. V tomto případě skoro všechny přímky neobsahují žádné pozorování. Pro určení směrových kvantilů proto musíme použít nějakou lokální metodu. Nabízí se nahradit kvantily na přímce nějakým odhadem kvantilu na okolí zkoumané přímky. Jinou možností je nejprve odhadnout rozdělení, například pomocí jádrového odhadu hustoty, a poté spočítat směrové kvantily pro odhad hustoty. V každém případě na volbě metody závisí, zda empirické směrové kvantily budou afinně invariantní, případně konzistentní.

²⁴Autor doufá, že zde přesně vystihl význam trochu vágněji napsané definice, kterou lze najít v původním článku.



Obrázek 15: Nezacyklená a zacyklená kvantilová kontura.

Pokud nedochází k zacyklení kvantilových kontur, jsou směrové centrální množiny přirozeně hvězdovitě. V opačném případě pochopitelně nikoliv. V článku [52] autor navrhuje *zacyklenou oblast přidat* do centrální množiny, čímž se odstraní (!) uvedený defekt a centrální množina je hvězdovitá. Střed S v každém případě zůstává na hranici všech zacyklených kvantilových kontur.

4.2 Polární směrové kvantily—cykly u nás nevedeme

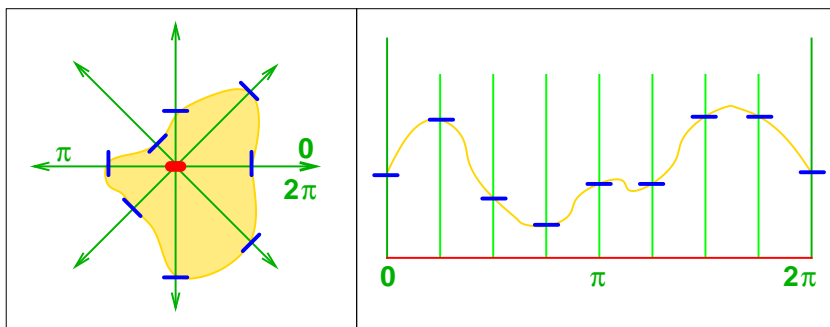
V diplomové práci [17] se objevuje jiná definice směrového kvantilu. Tentokrát se nebudeme dívat na přímky procházející středem S , ale na polopřímky, kterým je S počátkem.

Definice 11. *Uvažujme absolutně spojitě rozdělení pravděpodobností P na \mathbb{R}^p s hustotou f . Buď S mnohorozměrný medián rozdělení P . Pro libovolnou polopřímku m tvaru $m = S + tx$, $t \geq 0$, kde $\|x\| = 1$ definujme podmíněné rozdělení $f_m(t)$ jako podmíněné rozdělení P za podmínky, že X leží na polopřímce m . Pro (jednorozměrné) rozdělení s hustotou f_m najdeme $1 - \alpha$ kvantil $q_{1-\alpha}(m)$. Body $S + q_{1-\alpha}(m)x$ tvoří $1 - \alpha$ směrové kvantily pro polopřímku m . Množina všech směrových $1 - \alpha$ kvantilů tvoří $1 - \alpha$ kvantilovou konturu.*

Ekvivalentní a názornější je definice *polopřímkového* kvantilu pomocí *polárních souřadnic* vůči středu S .

Uvažujme absolutně spojitě rozdělení pravděpodobností P na \mathbb{R}^p s hustotou f . Buď S mnohorozměrný medián rozdělení P . Převeďme náhodnou veličinu X s rozdělením P na náhodnou veličinu $Z = (D, A)$ vyjadřující hypersférické souřadnice X vůči S . Konkrétně $D \geq 0$ je *vzdálenost* X od S a A určuje *směr*²⁵ X od S . Převod do sférických souřadnic v trojrozměrném Euklidově prostoru lze přehledně nalézt například v [41].

²⁵V hypersférických souřadnicích se jedná o vektor úhlů $\phi_1 \in [0, 2\pi)$, $\phi_i \in [0, \pi)$, $i = 2, \dots, p-1$.



Obrázek 16: Polopřímkový směrový kvantil a jeho výpočet pomocí polárních (hypersférických) souřadnic.

Označme Q rozdělení náhodné veličiny Z , Q je opět absolutně spojité, a označme g hustotu Q vůči Lebesgueově míře. Označme dále g_a hustotu podmíněného rozdělení Q_a vzdálenosti D při podmínce $A = a$.

Definice 12. Polárními kvantily rozdělení P vůči středu S rozumíme kvantily podmíněných rozdělení Q_a .

Polární α kvantil ve směru a , označený $q_{1-\alpha}(a)$, je definovaný jako $1 - \alpha$ kvantil jednorozměrné náhodné veličiny D za podmínky $A = a$.

Po převedení hypersférických souřadnic $(q_{1-\alpha}(A), A)$ zpět do kartézských dostaneme stejné kvantilové kontury jako v definici 11. Z praktických důvodů dáváme přednost definici 12, protože názorně ukazuje cestu k výpočtu, viz obrázek 16.

Stejně jako v případě přímkového směrového kvantilu, i polární směrový kvantil je afinně ekvivariantní, je-li afinně ekvivariantní volba středu S . Díky polárním souřadnicím je možné poměrně snadno počítat podmíněné kvantily $q_{1-\alpha}(a)$ ve směru a i pro empirická rozdělení. Tato definice přímo vybízí k použití regresních kvantilů, které v tomto případě musí být ve správném smyslu periodické. Více o tomto přístupu lze nalézt v diplomové práci [26], případně v tomto svazku v příspěvku [27].

Centrální oblasti polárních kvantilů jsou z definice přirozeně hvězdčovitě vůči středu S .

4.3 Porovnání obou přístupů

V obou případech se směrovými kvantily spojujeme přirozeně definovanou *směrovou hloubku*. Pro přímkové směrové kvantily ji lze pro bod x ležící na přímce $l = S + tz$ definovat jako

$$H(P, x) = \sup\{\alpha : x \in [S + q_{\alpha/2}(l)z, S + q_{1-\alpha/2}(l)z]\}. \quad (11)$$

V případě polárních směrových kvantilů je pro bod x ležící ve směru a od středu S hloubkou hodnota

$$H(P, x) = \sup\{\alpha : \|x - S\| \leq q_{1-\alpha}(a)\}. \quad (12)$$

Jinými slovy, bod x má hloubku $H(P, x)$, jestliže na přímce (polopřímce) spojující x a S leží uvnitř každé $1 - h$ kvantilové kontury pro $h < H(P, x)$.

Velkou výhodou směrových kvantilů je fakt, že lze přímo určovat $1 - \alpha$ centrální oblasti. To vyplývá z faktu, že z podmíněných rozdělání ve směrech (na přímkách, nebo v polárních souřadnicích) určuje obyčejné jednorozměrné kvantily, proto oblast vymezená kvantily pro daný směr má pravděpodobnost $1 - \alpha$ a díky větě o úplné pravděpodobnosti má příslušná centrální oblast pravděpodobnost $1 - \alpha$. Pro empirickou verzi toto platí samozřejmě jen přibližně. Tato vlastnost je velmi výjimečná, u všech ostatních uvedených hloubek je nutné postupovat opačně. Pro úrovněvé množiny hloubkové funkce nejprve určíme pravděpodobnost, za centrální oblast vybereme tu úrovněvou množinu, která má předepsanou pravděpodobnost.

V obou případech se jedná o výpočetně poměrně jednoduché postupy. Domníváme se, že náš postup pomocí polárních souřadnic je dokonce názornější a přímočarý. Jediné, nač musíme upozornit u polárních kvantilů je nutnost *periodicity* (více v [26, 27]).

Směrové kvantily jsou teoreticky afinně ekvivariantní. V obou případech ale tato vlastnost závisí na volbě středu S (viz výše). Afinní invariance hloubky neplatí pro empirické verze, neboť je nutné nahradit přímkou (polopřímku) oblastmi s nenulovou Lebesgueovou mírou. Zvolíme-li takovou oblast pevně, nemůže být afinně ekvivariantní. Proto lze doporučit vhodnou standardizaci náhodného výběru předtím, než přistoupíme k vlastnímu výpočtu hloubky.

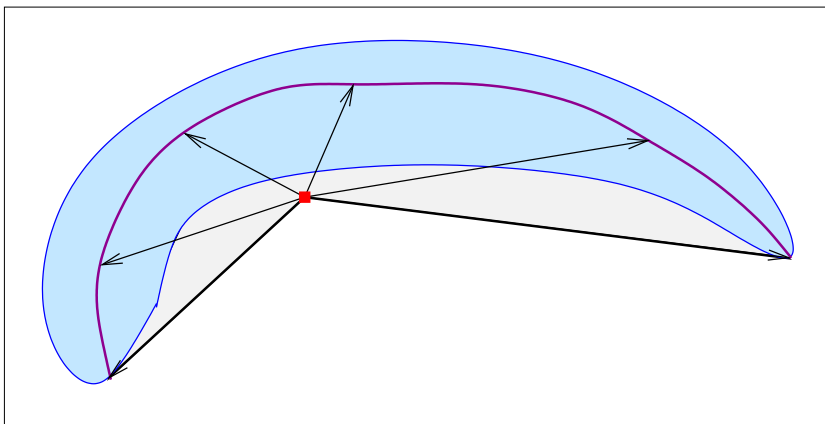
Na polárních směrových kvantilech je dobře vidět, že konzistence hloubky je spjata s konzistencí jednotlivých odhadů. Konzistentní musí být odhad středu, odhad regresních kvantilů a , v případě standardizace, odhad standardizační transformace. V článku [52] je konzistence uvedena jako *zřejmá*.

4.4 Volíme střed

Jak zvolit ten nejvhodnější střed pro směrové kvantily? Wei [52] navrhuje použít afinně ekvivariantní verzi mnohorozměrného mediánu (minimalizujícího střední \mathcal{L}_1 vzdálenost). V tom případě ale potřebujeme předpokládat existenci konečného prvního momentu. Navíc pro nekonvexní nosič rozdělání nemusí mnohorozměrný medián být *uvnitř* nosiče. Ideální kandidát na střed směrových kvantilů by ale *měl být obklopen* oblastí s nenulovou pravděpodobností ve všech směrech. Z podobných důvodů nemusí být vhodným kandidátem ani nejhlubší bod v poloprostorové hloubce, případně v L_1 hloubce.

Zatímco při použití přímkových směrových kvantilů střed mimo nosič rozdělání vede vždy k zacyklení (všech) kvantilových kontur²⁶, pro polární

²⁶K zacyklení může dojít i v případě, kdy střed *je* uvnitř nosiče



Obrázek 17: Polopřímkový směrový kvantil; střed je na hranici všech centrálních oblastí.

směrové kvantily k zacyklení nemůže dojít ani v tomto případě. Přesto přechod k polárním kvantilům není všespásný. Stále zůstává faktem, že polární směrové kvantily jsou vhodné pro *hvězdčovitá* rozdělení. Nemůže sice nastat zacyklení kvantilové kontury, ale střed může stále být na hranici *všech* centrálních oblastí. Na obrázku 17 je proto znázorněno rozdělení *na banánu*, klasický příklad odhalující slabiny mnoha hloubek.

Světle šedá oblast neobsahuje žádné hodnoty náhodné veličiny, přesto je zahrnuta do všech centrálních oblastí. Střed je na hranici všech centrálních oblastí (je obsažen ve všech kvantilových konturách). Problém je zjevně ve volbě středu²⁷. Přitom množina na obrázku 17 hvězdčovitá je vůči mnoha bodům (což sice *neznamená*, že i rozdělení s tímto nosičem je hvězdčovité, ale předpokládáme, že ano), takže by vhodným středem mohl být některý z těchto bodů.

V mnoha situacích by se nejlépe hodil modus rozdělení. Modus se ale nepadně hledá a, což je podstatnější, nemusí být jediný²⁸.

Ve standardních případech symetrických rozdělení je vesměs jedno, kterou hloubku zvolíme pro hledání středu—nejhlubšího bodu. Jsou ale situace, kdy nemůžeme říci, který střed je ten pravý. Snad by stálo za úvahu použít pro hledání vhodného *středu* podobný postup, jaký je v článku [10] navržen pro hledání vhodného *tvaru* hvězdčovité množiny. Vybrat střed (definuje příslušnou hloubku a kvantilové kontury) podle nějakého vhodného kritéria pro (vybrané) centrální oblasti.

²⁷Ale co dělat, když například medián po složkách, nebo poloprostorové nejhlubší bod takto vyjde?

²⁸Modus ovšem může také ležet na hranici nosiče, jako je tomu například pro exponenciální rozdělení

Nezbytnost dosadit zvenčí střed S do definice směrové hloubky může vést ke zvýšení výpočetní náročnosti. Na druhou stranu, ani při použití poloprostorové hloubky není nutné počítat hloubku *všech bodů* k určení nejhlubšího bodu. Mezi rychlé algoritmy výpočtu empirického nejhlubšího bodu patří například postup popsáný v článku Jiřího Matouška [36].

Poznámka 2. *Směrová hloubka se hodí zejména na rozdělení hvězdovitých typu. Oproti kvazikonkávním hloubkám dostaneme širší třídu možných tvarů centrálních množin.*

Přesto v některých případech můžeme potřebovat ještě flexibilnější třídu možných centrálních oblastí. Například pro směsi rozdělení. Nechceme-li přitom používat úrovně množiny, musíme nějakým způsobem lokalizovat některou z klasických definic hloubek.

5 Zobecnění poloprostorové hloubky—vážená hloubka

Původně tato hloubka vznikla při hledání vhodného středu pro směrovou hloubku. Ukázalo se ale, že má některé *nečekané*, ale *zajímavé* vlastnosti, viz [20, 50, 51]. Zejména oceňujeme možnost volit mezi zvýrazněním či potlačněním lokálních a globálních vlastností zkoumaného rozdělení.

5.1 Definice

Nejprve definujeme váhovou funkci.

Definice 13. *Funkce $w_+ : \mathbb{R}^p \rightarrow [0, \infty)$ taková, že $w_+(x) = 0$ pro $x \in \{(x_1, \dots, x_p); x_p < 0\}$ se nazývá váhová funkce hloubky. Označme dále $w_-(x) = w_+(-x)$.*

Z mnoha důvodů se vyplatí používat jen *rozumné* váhové funkce. Například takové, že jejich hodnota závisí na *vzdálenosti* podél osy x_p a na vzdálenosti $x_1^2 + \dots + x_{p-1}^2$ od osy x_p . Tedy

$$w_+(x_1, \dots, x_{p-1}, x_p) = w(x_1^2 + \dots + x_{p-1}^2, x_p), \quad (13)$$

což je požadavek určité *silné symetrie* funkce w_+ okolo osy x_p . Slabším požadavkem je *symetrie* váhové funkce hloubky

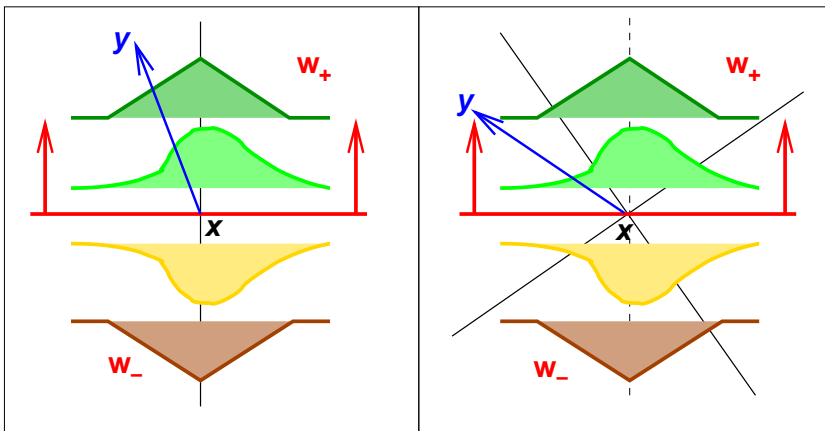
$$w_+(x_1, x_2, \dots, x_{p-1}, x_p) = w_+(-x_1, -x_2, \dots, -x_{p-1}, x_p). \quad (14)$$

Někdy je vhodné předpokládat také *směrovou homogenitu* váhové funkce hloubky

$$w_+(x_1, \dots, x_{p-1}, x_p) = w_+(x_1, \dots, x_{p-1}, kx_p), \text{ pro všechna } k > 0, \quad (15)$$

případně *úhlovou homogenitu* ve tvaru

$$w_+(x_1, \dots, x_{p-1}, x_p) = w_+(kx_1, \dots, kx_{p-1}, kx_p), \text{ pro všechna } k > 0. \quad (16)$$



Obrázek 18: Výpočet vážené hloubky bodu x minimalizací přes všechny ortonormální transformace rozdělení.

Nebudeme-li váhovou funkci specifikovat, budeme dále požadovat její symetrii, po částech spojitost a $w_+(0, \dots, 0, x_p) \geq m > 0$ pro $x_p \geq 0$. Příkladem vhodné váhové funkce může být například *pásová váhová funkce*

$$w_+(x) = \begin{cases} 1 & \sum_1^{p-1} x_i^2 \leq h, x_p \geq 0, \\ 0 & \text{jinak.} \end{cases} \quad (17)$$

Pásová váhová funkce je zřejmě silně symetrická a směrově homogenní ve smyslu (13) a (15).

Vektor $(0, 0, \dots, 0, 1)$ má úlohu *směru*, kterým je zobecněná poloprostorová hloubka orientována. Označme \mathcal{O}_p množinu všech *ortonormálních transformací* (zachovávajících vzdálenosti a úhly) prostoru \mathbb{R}^p .

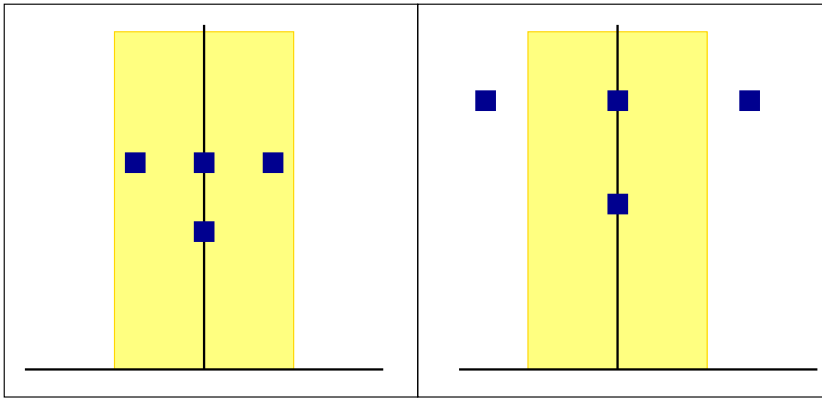
Definice 14. *Bud' P rozdělení pravděpodobnosti na prostoru \mathbb{R}^p a w_+ vhodná váhová funkce hloubky. Definujme hloubku bodu x jako*

$$D(P, x) = \inf_{A \in \mathcal{O}_p} \frac{E_P w_+(A(X - x))}{E_P w_-(A(X - x))}, \quad (18)$$

kte podíl $0/0$ definujeme jako 1.

Na obrázku 18 je znázorněn postup počítání vážené hloubky. V bodě x je počítána hloubka s využitím váhových funkcí w_+ a w_- . Provedeme všechny ortonormální transformace které nezmění bod x (na obrázku je znázorněno otočení bodu y a celého rozdělení) a pro každou takovou ortonormální transformaci A spočítáme podíl vah $E_P w_+(A(X - x)) / E_P w_-(A(X - x))$.

Takto definovaná hloubka je invariantní vůči posunutí i rotacím. Obecně však afinně invariantní není, protože definice váhové množiny v sobě obvyk-



Obrázek 19: Zavedení vah do poloprostorové hloubky kazí afinní invarianci.

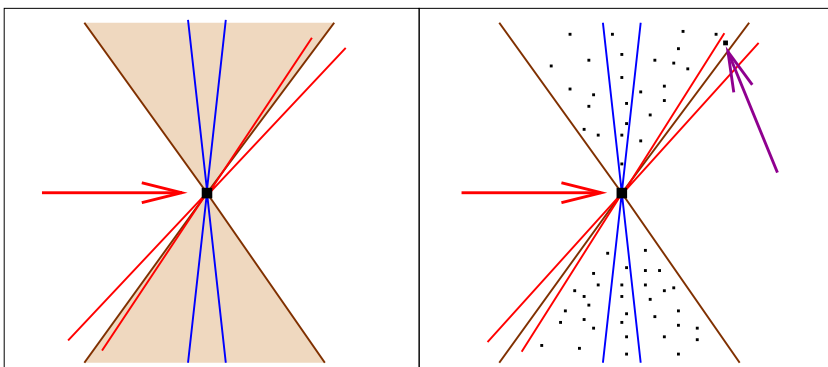
le zahrnuje (být skrytě) úhly mezi vektory. Například pásová hloubková funkce (17) nemůže definovat hloubku invariantní vůči různému přenásobení os, viz obrázek 19. Proto doporučujeme použít opět standardizaci pomocí vhodné transformace a teprve poté volit vhodnou hloubkovou funkci a počítat hloubku.

5.2 Základní vlastnosti a volba hloubkové funkce

Na rozdíl od většiny hloubek nemusí existovat právě jeden bod s největší váženou hloubkou. V tom je vážená hloubka podobná úrovnovým množinám. Centrální oblasti nemusí být ani konvexní, ani hvězdicovité, ba ani souvislé množiny. Na druhou stranu, při volbě vhodné hloubky (symetrické, případě ještě homogenní) lze snadno ukázat, že střed symetrických rozdělání (případně angulárně symetrických rozdělání) je jediným nejhlubším bodem. V případě středově symetrických rozdělání je i vážená hloubka středově symetrická za předpokladu, že váhová funkce hloubky je symetrická.

Pro absolutně spojitá rozdělání a *rozumné* váhové funkce je hloubka bodu konzistentní. Rozumné váhové funkce jsou takové, které jsou dostatečně spojitě (po částech spojitě) a správně měřitelné. příkladem může být funkce (17), případně nějaká její spojitá verze. Toto platí až na některé patologické případy bodů na hranici nosiče hustoty při nevhodně zvolené váhové funkci. Ukázka nekonzistence je na obrázku 20²⁹. Představme si rovnoměrné rozdělání na znázorněné množině. Přirozený střed ve vrcholu kužele teoreticky má, díky symetrii rozdělání, hloubku jedna. Ale pro libovolný náhodný výběr z tohoto rozdělání existuje pozorování y takové, že při vhodném otočení nezasáhne váhová funkce w_+ žádné pozorování, zatímco váhová funkce w_- zasáhne *právě* pozorování y . Empirická hloubka vrcholu kuželu je $0/n^{-1} = 0$ s pravděpodobností jedna, nemůže tak konvergovat k jedné.

²⁹Jde o dost podivný, až patologický příklad



Obrázek 20: Vážená hloubka nemusí být konzistentní ve všech bodech. Teoreticky (vlevo) má vrchol kuželu hloubku jedna. Pro libovolně velký výběr má vrchol s pravděpodobností jedna hloubku nula (vpravo).

Jde zřejmě o velmi umělý příklad. Hloubka všech ostatních bodů v tomto případě konzistentní je. Zde se sešly dvě neblahé skutečnosti; střed (vrchol kuželu) není obklopen nosičem ze všech stran a váhové funkce w_+ s w_- nepokrývají dohromady žádně okolí bodu 0, jen výseč tohoto okolí. Buď se tedy smíříme s tím, že hloubka nemusí být konzistentní na hranici nosiče, nebo zavedeme požadavek *lokálního pokrytí* váhové funkce. Tento požadavek lze vyjádřit formálně

$$\exists r > 0, \varepsilon > 0 \text{ tak, že } \|x\| \leq r, x_p \geq 0 \Rightarrow w_+(x) \geq \varepsilon > 0. \quad (19)$$

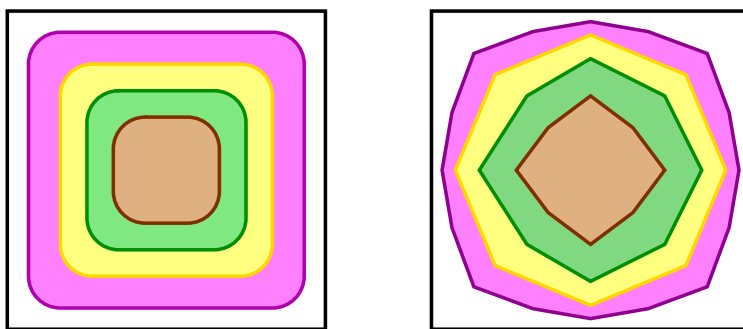
Celkově lze říci, že vážená hloubka oproti poloprostorové hloubce zvýrazňuje lokální vlastnosti rozdělení. Z toho plyne i návod, jak volit váhovou funkci. Volbou

$$w_+(x) = \begin{cases} 1 & x_p \geq 0, \\ 0 & x_p < 0 \end{cases}$$

dostáváme ekvivalent poloprostorové hloubky. Čím více zužujeme nosič hloubkové funkce k ose x_p (viz například vážená hloubka (17)), tím více vzdvihneme lokální vlastnosti rozdělení. Na druhou stranu, empirická vážená hloubka počítá jen s pozorováním, které padne do nosiče váhové funkce. Nespojitá váhová funkce má podobné dopady jako nespojitě jádro v odhadu hustoty. Proto lze, zejména pro menší rozsahy výběru, doporučit váhové funkce spojitě v $\mathbb{R}^{p-1} \times \mathbb{R}^+$.

5.3 Příklady rozdílů mezi poloprostorovou a váženou hloubkou

Zaměřme se na situace, kdy vážená hloubka dává podstatně jiné výsledky než poloprostorová hloubka. Použijme váhovou funkci (17). Připomeňme, že čím



Obrázek 21: Centrální oblasti vážené (vlevo) a poloprostorové (vpravo) hloubky pro rovnoměrné rozdělení na čtverci.

větší hodnotu h v (17) volíme, tím více se vážená hloubka blíží poloprostorové hloubce a naopak.

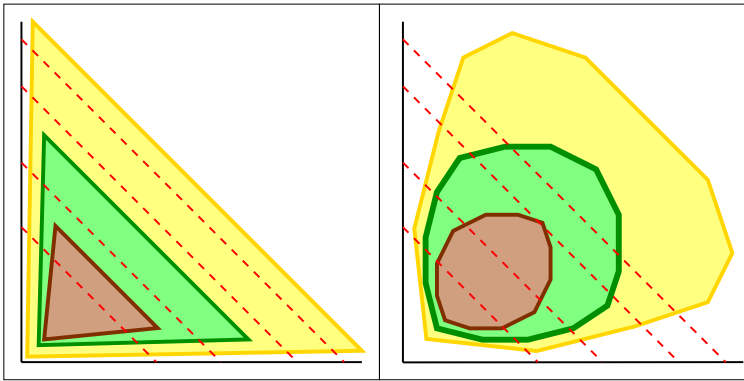
Na obrázku 21 vidíme schematicky znázorněné centrální oblasti pro rovnoměrné rozdělení na čtverci $[0, a]^2$. Volbou váhové funkce (17), kde h je relativně malé vůči délce strany a (například $h = a/20$) dostáváme obrázek vlevo. Na první pohled centrální oblasti vážené hloubky mnohem více kopírují tvar nosiče hustoty. Zaoblení v rozích jsou způsobena volbou šířky pásu h . Standardní poloprostorová hloubka (vpravo) vede na centrální oblasti, které jsou, protože jde o rovnoměrné rozdělení, stejně *velké* v Lebesgueově míře jako odpovídající centrální oblasti vážené hloubky. Tvar centrálních oblastí je hodně proměnlivý a jen pomalu se přibližuje nosiči hustoty. Připomeňme, že úrovnňové množiny jsou v tomto případě jen tři, jmenovitě \emptyset , $[0, a]^2$ a \mathbb{R}^2 .

Rovnoměrné rozdělení není možná úplně typické. Na obrázku 22 máme schematicky naznačeny centrální oblasti pro dvourozměrné rozdělení s exponenciálními marginálami a nezávislými složkami. Pro porovnání jsou doplněny vybrané vrstevnice hustoty (hranice úrovnňových množin).

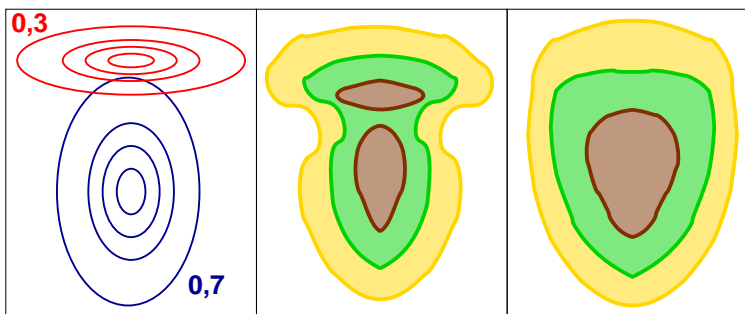
Vážená hloubka opět nabízí centrální oblasti, které jsou mnohem *blíže* struktuře hustoty rozdělení. Tyto centrální oblasti se příliš neliší od úrovnňových množin naznačených v obrázku. Centrální oblasti pro poloprostorovou hloubku jsou *příliš vypouklé* a *zbytečně objemné* (v Lebesgueově míře), podle našeho názoru nejsou moc vhodné³⁰. Na tomto příkladu je pěkně vidět, proč o vážené hloubce tvrdíme, že je *kompromisem mezi úrovnňovými množinami a poloprostorovou hloubkou*.

Poslední ilustrací budiž případ směsi dvou normálních rozdělení. Na obrázku 23 máme schematicky nakreslené vrstevnice dvou normálních rozdělení, které kombinujeme s vahami 0,3 a 0,7. Uprostřed a vlevo jsou naznačeny vrstevnice vážené, případně poloprostorové hloubky a jejich centrální oblasti.

³⁰Nikomlu ale náš názor nevnučujeme. Volba hloubky a posuzování jejích vlastností je velice subjektivní.



Obrázek 22: Centrální oblasti vážené (vlevo) a poloprostorové hloubky pro dvourozměrné exponenciální rozdělení. Čárkovaně jsou naznačeny úrovnňové množiny.



Obrázek 23: Centrální oblasti vážené (uprostřed) a poloprostorové (vpravo) hloubky pro směs dvou normálních rozdělení. Vlevo jsou naznačeny vrstevnice obou složek a jejich váha ve směsi.

Opět vidíme již známý fenomén. Vážená hloubka více kopíruje hustotu, centrální oblasti poloprostorové hloubky jsou konvexní a mají *zbytečně velkou* plochu. Centrální oblasti vážené hloubky naznačují existenci *dvou oblastí*, podobně jako by tomu bylo u úrovnňových množin. Zde se rýsuje možnost použít váženou hloubku pro diskriminační a klasifikační účely.

Přímo se nabízí otázka, zda množina všech kvantilových kontur, nebo ekvivalentně množina všech centrálních oblastí, může charakterizovat rozdělení. Pro poloprostorovou hloubku platí, viz [46], že množina všech *empirických* kvantilových kontur charakterizuje náhodný výběr, potažmo empirickou míru. Podobný výsledek je znám i pro některé další hloubky. Speciálně pro zonoidy [24] a pro simplexovou hloubku [25]. Pro váženou hloubku ještě obdobný výsledek nemáme.

6 Používání hloubky

Hloubka se uplatňuje například v mnohorozměrných obdobách neparametrických testů. Rozsáhlé shrnutí nabízí článek [31], ve kterém se lze dočíst o popisných statistikách, obdobách známých QQ diagramů, nebo o mnohorozměrné variantě krabicových diagramů, slunečních grafech³¹. Jsou definovány pojmy jako měřítko a rozptyl, šikmost, špičatost coby obdoby známých charakteristik jednorozměrné náhodné veličiny. Pomocí grafických metod jsou potom testovány hodnoty těchto charakteristik, případně porovnávány dva výběry. Zde se naskýtá mnoho příležitostí pro odvozování dalších variant a verzí neparametrických mnohorozměrných metod založených na uspořádání hloubkou. Novější výsledky lze nalézt například v článcích [28, 33].

Další z aplikací hloubek je v oblasti klasifikace dat. Velkou výhodou je neparametrická podstata hloubky, čili klasifikační kritéria nejsou založena na (obvyklém) předpokladu normality. Více informací lze získat v článcích [15, 16, 22, 39].

Hloubka je ale zajímavá i pro výzkumníky v oblasti výpočetní geometrie. Zájemcům doporučujeme <http://www.cs.tufts.edu/research/geometry>. Zde lze nalézt také odkaz na software Depth Explorer [21]. Ten pokrývá zejména poloprostorovou hloubku a L_1 hloubku a v současné době by měl už být i ve verzi pro Linux.

Poslední oblastí, kterou zmíníme, je uplatnění hloubky pro funkcionální data, tedy pro nekonečněrozměrné výběrové prostory. V poslední době se objevuje několik prací, zabývajících se touto problematikou, například [8, 9, 34]. I zde je výzkum teprve v počátcích. Proto se domníváme, že v oblasti hloubky a jejích aplikací je spousta zajímavých a nevyřešených problémů, které čekají, až se jich někdo chopí.

Literatura

- [1] Baillo A., Cuesta-Albertas J.A., Cuevas A. (2001). *Convergence rates in nonparametric estimation of level sets*. *Statistics and Probability Letter* **53**, 27–35.
- [2] Breckling J., Chambers R. (1988). *M-quantiles*. *Biometrika* **75**, 761–771.
- [3] Breckling J., Kokic P., Lübke O. (2001). *A note on multivariate M-quantiles*. *Statistics and Probability Letters* **55**, 39–44.
- [4] Burr M.A., Rafalin, E., Souvaine, D.L. (2006). *Simplicial depth: An improved definition, analysis and efficiency for the finite sample case*. In *Data Depth: Robust Multivariate Analysis, Computational Geometry and Applications* (Liu, Serfling and Souvaine, Eds.), American Mathematical Society, 1–16.

³¹Sunburst plot

- [5] Chakraborty B., Chaudhuri P. (1996). *On a transformation and re-transformation technique for constructing an affine equivariant multivariate median*. Proceedings of the American Mathematical Society **124**, 2539–2547.
- [6] Chakraborty B., Chaudhuri P., Oja H. (1998). *Operating transformation retransformation on spatial median and angle test*. Statistica Sinica **8**, 767–784.
- [7] Chaudhuri P. (1996). *On a geometric notion of quantiles for multivariate data*. Journal of the American Statistical Association **91**, 862–872.
- [8] Cuesta-Albertos J.A., Fraiman, R. (2006) *Impartial trimmed means for functional data*. In Data Depth: Robust Multivariate Analysis, Computational Geometry and Applications (Liu, Serfling and Souvaine, Eds.), American Mathematical Society, 121–146.
- [9] Cuesta-Albertos, J. A. and Nieto-Reyes, A. (2008) *A random functional depth*. In Functional and Operatorial Statistics (Dabo-Niang and Ferraty, Eds.), Physica-Verlag, 121–126.
- [10] DasGupta A., Ghosh J.K., Zen M.M. (1995). *A new general method for constructing confidence sets in arbitrary dimensions: with applications* Annals of Statistics **23**, 1408–1432.
- [11] Donoho D., Gasko M. (1992). *Breakdown properties of location estimates based on halfspace depth and projected outlyingness*. Annals of Statistics **20**, 1803–1827.
- [12] Dudley R.M. (2008). *Uniform Central Limit Theorems*. Cambridge.
- [13] Dümbgen L. (1991). *Limit theorems for the simplicial depth*. Statistics and Probability Letters **14**, 119–128.
- [14] Dyckerhoff R. (2000). *Computing zonoid trimmed regions of bivariate data sets*. In Compstat 2000. Proceedings in Computational Statistics (Bethlehem and van der Heijden, Eds.), Physica-Verlag, 295–300.
- [15] Ghosh A.K., Chaudhuri, P. (2005). *On data depth and distribution free discriminant analysis using separating surfaces*. Bernoulli **11**, 1–27.
- [16] Hartikainen A., Oja H. (2006). *On some parametric, nonparametric and semiparametric discrimination rules*. In Data Depth: Robust Multivariate Analysis, Computational Geometry and Applications (Liu, Serfling and Souvaine, Eds.), American Mathematical Society, 61–70.
- [17] Hasil J. (2005). *Problém kvantilu ve více rozměrech*. Diplomová práce, MFF UK.
- [18] Hettmansperger T.P., Randles R.H. (2002). *A practical affine equivariant multivariate median*. Biometrika **89**, 851–860.
- [19] Hlubinka, D. (2002). *O kvantilech ve více rozměrech*. In ROBUST 2002 (Antoch, Dohnal, Eds.).
- [20] Hlubinka, D., Kotík, L. and Vencálek, O. (2009). *Weighted halfspace depth*. To appear in Kybernetika

- [21] Hugg J., Rafalin E., Seyboth K. Souvaine D. (2006). *An experimental study of old and new depth measures*. In Workshop on Algorithms Engineering and Experiments. Lecture Notes in Computer Sciences, Springer, 51–64.
- [22] Jörnsten, R. (2004). *Clustering and classification based on the L_1 data depth*. Journal of Multivariate Analysis **90**, 67–89.
- [23] Koltchinskii, V. I. (1997). *M-estimation, convexity and quantiles*. Annals of Statistics **25**, 435–477.
- [24] Koshevoy G., Mosler K. (1997). *Zonoid trimming for multivariate distributions*. Annals of Statistics **25**, 1998–2017.
- [25] Koshevoy, G. (1997). *Multivariate depths and underlying distributions: A uniqueness property*. Technical report, Universität zu Köln.
- [26] Kotík, L. (2007). *Periodické regresní kvantily*. Diplomová práce, MFF UK.
- [27] Kotík, L. (2009). *Directional quantiles*. In ROBUST 2008 (Antoch, Dohnal, Eds.)
- [28] Li, J. and Liu, R. Y. (2004). *New nonparametric tests of multivariate locations and scales using data depth*. Statistical Science **19**, 686–696.
- [29] Liu, R. Y. (1988). *On a notion of simplicial depth*. Proceedings of the National Academy of Science USA **85**, 1732–1734.
- [30] Liu, R. Y. (1990). *On a notion of data depth based on random simplices*. Annals of Statistics **18**, 405–414.
- [31] Liu R.Y., Parelius J.M., Singh K. (1999). *Multivariate analysis by data depth: Descriptive statistics, graphics and inference (with discussion)*. Annals of Statistics **27**, 783–858.
- [32] Liu R.Y., Serfling R., Souvaine D.L., Editors (2006). *Data Depth: Robust Multivariate Analysis, Computational Geometry and Applications*. American Mathematical Society.
- [33] Liu R.Y., Singh K. (2006). *Rank tests for multivariate scale difference based on data depth*. In Data Depth: Robust Multivariate Analysis, Computational Geometry and Applications (Liu, Serfling and Souvaine, Eds.), American Mathematical Society, 17–36.
- [34] López-Pitado S., Romo J. (2006). *Depth-based classification for functional data*. In Data Depth: Robust Multivariate Analysis, Computational Geometry and Applications (Liu, Serfling and Souvaine, Eds.), American Mathematical Society, 103–120.
- [35] Mason D., Polonik W. (2008). *Asymptotic normality of plug-in level set estimates*. Preprint.
- [36] Matoušek J. (1991). *Computing the centre of planar point set*. In Discrete and Computational Geometry: Papers from the DIMACS Special Year (Goodman, Pollack and Steiger, Eds.), American Mathematical Society, 221–230.
- [37] Mosler K. (2002). *Multivariate Dispersion, Central Regions and Depth; The Lift Zonoid Approach*. Springer.

- [38] Mosler, K. (2003). *Central regions and dependency*. Methodology and Computing in Applied Probability **5**, 5–21.
- [39] Mosler K., Hoberg R. (2006). *Data analysis and classification with the zonoid depth*. In Data Depth: Robust Multivariate Analysis, Computational Geometry and Applications (Liu, Serfling and Souvaine, Eds.), American Mathematical Society, 49–60.
- [40] Polonik W. (1997). *Minimum volume sets and generalized quantile processes*. Stochastic Processes and their Applications **69**, 1–24.
- [41] Rektorys, K. (2003). *Přehled užití matematiky I (7. vyd.)*. Prometheus.
- [42] Rousseeuw, P. and Ruts, I. (1996). *Algorithm AS307. Bivariate location depth*. Applied Statistics **45**, 516–526.
- [43] Rousseeuw P., Struyf A. (1998). *Computing location depth and regression depth in higher dimensions* Statistics and Computing **8**, 193–203.
- [44] Rousseeuw P., Hubert M. (1999). *Regression depth (with discussion)* Journal of the American Statistical Association **94**, 388–433.
- [45] Serfling R. (2006). *Depth functions in nonparametric multivariate inference*. In Data Depth: Robust Multivariate Analysis, Computational Geometry and Applications (Liu, Serfling and Souvaine, Eds.), American Mathematical Society, 1–16.
- [46] Struyf A., Rousseeuw P. (1999). *Halfspace depth and regression depth characterize the empirical distribution*. Journal of Multivariate Analysis **69**, 135–153.
- [47] Tukey J. (1975). *Mathematics and the picturing of data*. In Proceedings of the International Congress of Mathematicians, Vancouver 1974 (R.D. James, Ed.) **2**, 523–531.
- [48] van den Vaart A., Welner J. (2000). *Weak Convergence and Empirical Processes: With Applications to Statistics*. Springer.
- [49] Vardi Y., Zhang C.H. (2000). *The multivariate L_1 median and associated data depth*. Proceedings of National Academy of Science USA **97**, 1423–1426.
- [50] Vencálek O. (2008). *Generalization of halfspace depth*. In WDS '08; „Week of Doctoral Students 2008“ Part I (Šafránková and Pavlů, Eds.), Matfyzpress, 100–105.
- [51] Vencálek O. (2009). *Weighted data depth and its properties*. In ROBUST 2008 (Antoch, Dohnal, Eds.)
- [52] Wei Y. (2008). *An Approach to Multivariate Covariate-Dependent Quantile Contours With Application to Bivariate Conditional Growth Charts*. Journal of the American Statistical Association **103**, 397–409.

Poděkování: Článek vznikl s podporou grantu GAČR 201/08/0486. Práce je součástí výzkumného projektu MSM 0021620839 financovaného MŠMT ČR. Autor děkuje Lukáši Kotíčkovi a Ondřeji Vencálkovi za pomoc a podporu.

Adresa: D. Hlubinka, MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8 – Karlín

E-mail: daniel.hlubinka@mff.cuni.cz

ESTIMATION OF THE SCALE PARAMETER IN BURR TYPE XII DISTRIBUTION

Erika Hönschová

Keywords: L-estimate, Burr type XII distribution, scale estimate

Abstract: An L -estimate of the scale parameter of the Burr type XII distribution, under the assumption that other parameters of this distribution are known, is presented. The parameter c of this distribution is assumed to fulfill the inequality $c > \frac{1}{2}$, which includes the Pareto distribution (the value $c = 1$) as a special case. Under this assumption the asymptotic normality of this estimate is proved and the order of convergence of the remainder term in the asymptotic linearity expansion of the presented statistic is also included.

Abstrakt: V práci je skonštruovaný L -odhad parametra škály Burrovho rozdelenia typu XII, za predpokladu, že ostatné parametre tohto rozdelenia sú známe. Predpokladáme, že parameter c tohto rozdelenia spĺňa podmienku $c > \frac{1}{2}$, čo zahŕňa aj Paretovo rozdelenie (hodnotu $c = 1$) ako špeciálny prípad. Za tohto predpokladu je dokázaná asymptotická normalita tohto odhadu a práca tiež zahŕňa rád konvergencie zvyšku v asymptotickej lineárnej reprezentácii.

1 Introduction

Let X be a random variable with distribution function belonging to the location-scale family of the Burr type XII distribution, which is according to [4] given by

$$F(x, \mu, \sigma, k, c) = 1 - \left(1 + \left(\frac{x - \mu}{\sigma}\right)^c\right)^{-k} \quad \text{for } x \geq \mu, \quad (1)$$

where $k > 0, c > 0, \sigma > 0, \mu \in \mathcal{R}$.

The density function is

$$f(x, \mu, \sigma, k, c) = \frac{kc}{\sigma} \left(1 + \left(\frac{x - \mu}{\sigma}\right)^c\right)^{-k-1} \left(\frac{x - \mu}{\sigma}\right)^{c-1} \quad \text{for } x \geq \mu. \quad (2)$$

The Burr type XII distribution has been applied in studies of household income, insurance risk, reliability analysis etc, e.g. by Tadikamalla [9], Embrechts and Schmidli [1], McDonald [8].

The topic of this paper is the estimation problem of the scale parameter σ based on order statistics, parameters k, c, μ are assumed to be known. The asymptotic distribution of the estimate is presented. When compared with the maximal likelihood estimate, this estimate is easier to compute.

The estimation of the scale parameter based on order statistics in the special case $c = 1$ (known also as the Pareto distribution) has been studied by Kulldorf and Vännman [7] and Vännman [10]. They derived the best linear unbiased estimator provided that the parameter k is known and fulfills the condition $k > 2$, the case $k \leq 2$ is also covered, but by estimates based on a few selected order statistics, what may be disadvantageous. The asymptotic distribution of the estimates derived in [7] and [10] is unknown.

The basic idea used in [7] and [10] is to compute the means and covariances of the order statistics $X_n^{(1)} \leq X_n^{(2)} \leq \dots \leq X_n^{(n)}$ and compute the best linear unbiased estimator using Gauss-Markov theorem. The disadvantage of this method is, that only order statistics with finite second moments can be considered to obtain the estimate. E.g. in the case of Pareto distribution with $k < 2$ we can use just first $m < n + 1 - 2/k$ order statistics. It is not easy to compute the means and covariances of the order statistics from Burr type XII distribution, therefore we present another method of obtaining L-estimate of scale. The estimate is asymptotically normally distributed and asymptotically efficient.

2 Estimation of the scale parameter

We will use L-estimates in the form:

$$L_n = \sum_{i=1}^n c_{ni} X_n^{(i)}, \quad c_{ni} = \int_{(i-1)/n}^{i/n} J(u) du, \quad (3)$$

where $X_n^{(1)} \leq X_n^{(2)} \leq \dots \leq X_n^{(n)}$ are the order statistics and $J(u)$, $u \in \langle 0, 1 \rangle$ is a weights generating function. Under various set of conditions imposed on the distribution function of the random sample and the weights generating function, the asymptotic linearity of the estimate was proved by several authors [2, 6, 3]. The statement of the theorem proved in [3] is as follows:

Theorem 1. (I) *Let following assumption be fulfilled:*

(A 1) *The distribution function $F(t) = P(X \leq t)$ is continuous and strictly increasing on (d, D) , where $d = \inf\{t; F(t) > 0\}$, $D = \sup\{t; F(t) < 1\}$.*

(A 2) *The function $J : (0, 1) \rightarrow \mathcal{R}$ possesses the derivative J' on $(0, 1)$ and*

$$\nu = \int_0^1 J(u) F^{-1}(u) du$$

is a real number.

(A 3) *There exist real numbers $\gamma = \gamma_d > -2$, $K > 0$ such that for each $u \in (0, \frac{1}{2})$*

$$|J(u)| \leq K u^{1+\gamma}, \quad |J'(u)| \leq K u^\gamma.$$

There exist real numbers $\gamma = \gamma_D > -2$, $K > 0$ such that for each $u \in (\frac{1}{2}, 1)$

$$|J(u)| \leq K(1-u)^{1+\gamma}, \quad |J'(u)| \leq K(1-u)^\gamma.$$

(A 4) There exist real numbers $\kappa_d < \gamma_d + 1$, $\kappa_D < \gamma_D + 1$ such that the integrals

$$\int_0^{1/2} u^{\kappa_d} dF^{-1}(u), \quad \int_{1/2}^1 (1-u)^{\kappa_D} dF^{-1}(u)$$

are real numbers.

(A 5) For every real number x the integrals

$$H(x) = \int_x^{+\infty} |J(F(y))| dy, \quad \int_{-\infty}^{+\infty} H(x) dF(x) < +\infty, \quad \int_{-\infty}^{+\infty} J(F(y))F(y) dy$$

are real numbers and $F(x)H(x) \rightarrow 0$ as $|x| \rightarrow +\infty$.

Then for the L-estimate (3) the following representation holds

$$L_n = \nu + \frac{1}{n} \sum_{i=1}^n \psi(X_i) + \mathcal{O}_P\left(\frac{1}{n}\right), \tag{4}$$

where

$$\psi(x) = \int_{-\infty}^{+\infty} J(F(y))F(y) dy - \int_x^{+\infty} J(F(y)) dy.$$

(II) Put

$$L_n^* = \sum_{i=1}^n c_{ni} X_n^{(i)}, \quad c_{ni} = \frac{1}{n} J\left(\frac{i}{n+1}\right). \tag{5}$$

In addition to (A 1) - (A 5) suppose also that for some $t_d, t_D \in (d, D)$ and some positive real numbers β_d, β_D the inequalities

$$\sup\{|x|^{\beta_d} F(x); d < x \leq t_d\} < +\infty, \quad \sup\{|x|^{\beta_D} (1-F(x)); t_D \leq x < D\} < +\infty$$

hold. If

$$\delta_1 = 2 + \gamma_d - \frac{1}{\beta_d} > 0, \quad \delta_2 = 2 + \gamma_D - \frac{1}{\beta_D} > 0, \tag{6}$$

then

$$L_n^* = \mu + \frac{1}{n} \sum_{i=1}^n \psi(X_i) + R_n, \tag{7}$$

where

$$R_n = \begin{cases} \mathcal{O}_P\left(\frac{\log n}{n}\right), & \min\{\delta_1, \delta_2\} = 1, \\ \mathcal{O}_P\left(\frac{1}{n^{\delta^*}}\right) & \text{otherwise.} \end{cases}$$

Here (cf. (6))

$$\delta^* = \min\{1, \delta_1, \delta_2\}.$$

Since $\int_{\mathcal{R}} \psi(x)f(x)dx = 0$, by means of the central limit theorem from (4) follows, that $\sqrt{n}(L_n - \nu) \rightarrow N(0, V)$ in distribution, where $\nu = \int_0^1 J(u)F^{-1}(u) du$ and $V = \int_{\mathcal{R}} \psi^2(x)f(x)dx$.

Let X_1, \dots, X_n be a random sample from the Burr type XII distribution. The parameters μ, k, c are assumed to be known. To find the L-estimate of the scale parameter we will compute the weights generating function $J(u)$ in a manner described e.g in [5] p.77, which ensures the asymptotically efficiency of the estimate. Define

$$\phi(x) = \frac{1}{\mathcal{I}(\sigma)} \frac{\partial \ln(f(x, \mu, \sigma, k, c))}{\partial \sigma}$$

where

$$\mathcal{I}(\sigma) = E \left[\left(\frac{\partial \ln^2(f(x, \mu, \sigma, k, c))}{\partial \sigma} \right) \right]$$

is the Fisher information. For $c > 1/2, k > 0$ put

$$J(u) = \phi'(F^{-1}(u)) = \frac{(k+1)(k+2)}{k} (1-u)^{\frac{2}{k}} ((1-u)^{-\frac{1}{k}} - 1)^{\frac{c-1}{c}}, \quad u \in (0, 1)$$

and

$$\hat{\sigma}_n = \sum_{i=1}^n c_{ni} (X_n^{(i)} - \mu), \quad c_{ni} = \int_{(i-1)/n}^{i/n} J(u)du. \quad (8)$$

For $c > \frac{1}{2}, k > 0$ the functions $J(u), F(u)$ satisfy the assumption of the theorem 1 (I) (when compared with [6], [2], this work allows more general condition imposed on the parameters k, c in the Burr type XII distribution) with $\gamma_d = -\frac{1}{c}$, $\gamma_D = \frac{1}{k} + \frac{1}{kc} - 1$, $\kappa_d = -\frac{1}{c} + \Delta_d$, where $0 < \Delta_d < 1$, $\kappa_D = \frac{1}{kc} + \Delta_D$, where $0 < \Delta_D < \frac{1}{k}$, therefore the representation (4) holds with

$$\nu = \int J(u)F^{-1}(u)du = \sigma$$

$$\psi(x) = \frac{\sigma(2+k)}{c} \left[1 - \frac{(1+k)}{k} \left(1 + \left(\frac{x-\mu}{\sigma} \right)^c \right)^{-1} \right] \quad \text{for } x > \mu$$

and

$$\sqrt{n}(\hat{\sigma}_n - \sigma) \rightarrow N(0, V) \quad (9)$$

in distribution, where

$$V = \frac{1}{\mathcal{I}(\sigma)} = \frac{2+k}{kc^2} \sigma^2.$$

If it is tedious to compute the score c_{ni} defined in (8), we can use the approximation $\tilde{c}_{ni} = \frac{1}{n} J(\frac{i}{n+1})$. Since the conditions of theorem 1 (II) are fulfilled for $c > \frac{1}{2}, k > 0$ with $\beta_d = -c$, $\beta_D = ck$, the representation (4) holds also for estimate computed with score \tilde{c}_{ni} .

If μ is not known the estimate $\hat{\mu} = X_n^{(1)}$ can be used. The scale estimate takes the form $\hat{\sigma}_n = \sum_{i=1}^n c_{ni}(X_n^{(i)} - \hat{\mu})$. Since for $c \leq 2$ the relation $\{X_n^{(1)} - \mu\} = O_p(1/\sqrt{n})$ holds, the asymptotic normality (9) of the estimate remains unchanged for $\frac{1}{2} < c \leq 2$. This condition includes also the case $c = 1$, which was in connection with estimation the location and scale parameter studied in [7, 10]. In [10] the author derived the BLUE, based on order statistic, when μ, k of the Pareto distribution are known but only for $k > 2$. If $\frac{2}{n} < k \leq 2$, the estimate based on first m order statistics is derived, but under the condition $m < n + 1 - 2/k$, what may be disadvantageous, mostly if k is small. The estimate derived in [10] with $m = n - \lfloor \frac{2}{k} \rfloor$ (for which according to [10, p. 705] the highest efficiency is obtained) is given by:

$$\tilde{\sigma}_m = \frac{1}{T_m} \left((k+1) \sum_{i=1}^{m-1} B_i X_n^{(i)} + ((n-m+1)k-1) B_k X_n^{(k)} + (T_m + 2 - nk) \mu \right),$$

where

$$T_m = \frac{nk - 2 - ((n-m)k - 2) B_m}{k + 2} \tag{10}$$

$$B_i = \left(1 - \frac{2}{k(n-i+1)} \right) B_{i-1}, \quad B_0 = 1.$$

This estimate is unbiased, with variance $V(\tilde{\sigma}_m) = \frac{\sigma^2}{T_m}$ (cf. (10)). The estimate defined by (8) is

$$\hat{\sigma}_n = \sum_{i=1}^n c_{ni}(X_n^{(i)} - \mu), \tag{11}$$

where

$$c_{ni} = (k+1) \left[\left(1 - \frac{i-1}{n} \right)^{\frac{2}{k}+1} - \left(1 - \frac{i}{n} \right)^{\frac{2}{k}+1} \right].$$

This estimate is simpler to compute, $\sqrt{n}(\hat{\sigma}_n - \sigma) \rightarrow N(0, \frac{2+k}{k} \sigma^2)$ in distribution and it is asymptotically efficient. It is useful mostly if k is small, because then the estimate $\tilde{\sigma}_m$ is not defined, while $\hat{\sigma}_n$ is defined for all $k > 0$. In the following table are this two estimates compared on the basis of 5000 simulations of samples from the Pareto distribution with $\mu = 0, \sigma = 1$. We selected the values $k = 3, 1, 0.5, 0.1$ and the sample ranges $n = 20, 50, 100$. The 25% and 75% sample quantiles are computed for each sample of estimates.

n=20				
k	3	1	0.5	0.1
$(Q1(\hat{\sigma}_n), Q3(\hat{\sigma}_n))$	(0.84, 1.25)	(0.82, 1.42)	(0.8, 1.64)	(0.86, 3.17)
$(Q1(\tilde{\sigma}_m), Q3(\tilde{\sigma}_m))$	(0.79, 1.17)	(0.7, 1.2)	(0.6, 1.22)	*

n=50				
k	3	1	0.5	0.1
$(Q1(\hat{\sigma}_n), Q3(\hat{\sigma}_n))$	(0.9, 1.14)	(0.88, 1.21)	(0.86, 1.31)	(0.8, 1.96)
$(Q1(\tilde{\sigma}_m), Q3(\tilde{\sigma}_m))$	(0.87, 1.12)	(0.82, 1.14)	(0.77, 1.17)	(0.5, 1.24)
n=100				
k	3	1	0.5	0.1
$(Q1(\hat{\sigma}_n), Q3(\hat{\sigma}_n))$	(0.92, 1.09)	(0.9, 1.14)	(0.89, 1.2)	(0.8, 1.52)
$(Q1(\tilde{\sigma}_m), Q3(\tilde{\sigma}_m))$	(0.9, 1.08)	(0.87, 1.1)	(0.84, 1.14)	(0.63, 1.21)

* For $n = 20$, $k = 0.1$ is the estimate $\tilde{\sigma}_m$ not defined.

The simulation showed, that this two estimates seems to be comparable, but in case $k = 0.1$, $n = 20$ only the estimate given by (8) is defined. This estimate can be used also in the case $c \neq 1$. This can be useful mainly in the case, when the data come from unimodal distribution. The density function of Pareto distribution is strictly decreasing and cannot be used for such data.

References

- [1] Embrechts P., Schmidli H. (1994). *Modelling of extremal events in insurance and finance*. Mathematical Methods of Operations Research **39** 1–34.
- [2] Govindarajulu Z., Mason D.M. (1983). *A strong representation for linear combinations of order statistics with application to fixed-width confidence intervals for location and scale parameters*. Scan. J. Statist. **10**, 97–115.
- [3] Hönschová E., Rublík F. *On asymptotic linearity of L-estimates*. Submitted for publication.
- [4] Johnson N.L., Kotz S., Balakrishnan N. (1994). *Continuous Univariate Distributions*. Vol.1. John Wiley, New York.
- [5] Jurečková J.(2001). *Robustní statistické metody*. Karolinum, Praha.
- [6] Jurečková J., Sen P.K. (1996). *Robust Statistical Procedures*. Asymptotics and Interrelations. John Wiley & Sons, New York.
- [7] Kulldorf G., Vännman K. (1973). *Estimation of the location and scale parameters of a Pareto distribution by linear functions of order statistics*. Journal of the American Statistical Association, **68**, 218–227.
- [8] McDonald J.B. (1984). *Some generalized functions for the size distribution of income*. Econometrica **53**, 647–663.
- [9] Tadikamalla P.R. (1980). *A look at the burr and related distributions*. International Statistical Review **48**, 337–344.
- [10] Vannman K. (1976). *Estimators based on order statistics from a Pareto distribution*. Journal of the American Stat. Association, **71**, 704-708.

Acknowledgement: The research has been supported by the Grant VEGA 1/3016/06.

Address: E. Hönschová, FMFI UK, KAMS, Mlynská dolina, 842 48 Bratislava 4

E-mail: honschova@fmph.uniba.sk

KVANTILY V SEKVENČNÍ ANALÝZE BODU ZMĚNY

Ondřej Chochola

Klíčová slova: sekvenční analýza, bod změny, kvantily

Abstrakt: Příspěvek se věnuje možnému využití kvantilů v detekci bodu změny v sekvenčně pozorovaných datech. Teoretická část je rozšířením práce [3], která je doplněna výsledky simulační studie.

Abstract: The contribution is devoted to possible usage of quantiles in sequential change point analysis. The theoretical part is an extension of [3] accompanied with a simulation study.

1 Úvod

Práce vychází ze schématu navrženém v [2] a rozpracovaném v [3] na tzv. L_1 odhady parametru polohy a regresního parametru. Zde se zaměříme na testování změny kvantilu distribuce pozorování.

Zkoumání kvantilů bylo motivováno možnou aplikací na testování změny Value at Risk (VaR). VaR se často používá v ekonometrii jako míra nej-různějších rizik, především pak tržního rizika daného portfolia. VaR odpovídá ztrátě, kterou portfolio s danou pravděpodobností nepřekročí v daném časovém úseku. Z matematického hlediska je tedy kvantilem distribuce výnosů.

2 Model

Uvažujeme sekvenčně přicházející nezávislá data Y_i , $i = 1, 2, \dots$ s distribučními funkcemi F_i , $i = 1, 2, \dots$. Označme $t_i^\tau = F_i^{-1}(\tau)$ τ kvantil pozorování. Předpokládáme tzv. podmínku stability

(A) máme k dispozici historická (tréninková) data beze změny kvantilu pozorování o délce m , tj. $t_1^\tau = \dots = t_m^\tau =: t_0^\tau$, pocházející z distribuce F . Testujeme nulovou hypotézu, že se τ kvantil pozorování nemění:

$$H_0 : t_i^\tau = t_0^\tau, \quad 1 \leq i < \infty, \quad (1)$$

přesněji, že se distribuční funkce F_i v okolí t_0^τ shodují s F , proti alternativě, že se v neznámém bodě $m + k^*$ distribuční funkce pozorování změní tak, že se na okolí t_0^τ shoduje s F_* , tj. pro τ kvantil pozorování platí:

$$H_1 : \text{existuje } k^* \geq 1 \text{ takové, že } t_i^\tau = t_0^\tau, \quad 1 \leq i < m + k^*, \\ t_i^\tau = t_*^\tau, \quad m + k^* \leq i < \infty, \quad \text{kde } t_*^\tau = F_*^{-1}(\tau) \neq t_0^\tau. \quad (2)$$

Dále předpokládáme

(B) F je spojitá distribuční funkce, F'' existuje v okolí t_0^τ

a $F'(t_0^\tau) = f(t_0^\tau) > 0$,

(C) F_* je spojitá distribuční funkce.

3 Testová statistika a hlavní výsledky

Přirozené by bylo založit test na rozdílu výběrových kvantilů získaných z monitorovacího a tréninkového období. Přepočítávání výběrových kvantilů po každém novém pozorování je však výpočetně náročné, takže se v testové statistice využívá analogie L_1 reziduí. Definujeme

$$\tilde{\varepsilon}_i = \tau - I[Y_i < \tilde{t}_m^\tau],$$

kde $I[\dots]$ značí indikátor jevu a \tilde{t}_m^τ je výběrový τ kvantil historických pozorování získaný minimalizací

$$\min_a \sum_{i=1}^m \rho_\tau(Y_i - a) \text{ pro } \rho_\tau(v) = v(\tau - I[v < 0]).$$

Populační analogie je $\varepsilon_i = \tau - I[Y_i < t_0^\tau]$, pro které za H_0 platí: $E\varepsilon_i = 0$, $\text{var } \varepsilon_i = \tau(1 - \tau)$. Proto velké absolutní hodnoty testové statistiky

$$Q(m, k) = \frac{1}{\sqrt{\tau(1 - \tau)}} \sum_{i=m+1}^{m+k} \tilde{\varepsilon}_i$$

zřejmě vedou k zamítnutí nulové hypotézy.

Čas zastavení definujeme jako

$$\tau(m) = \inf\{k : |Q(m, k)|/g(m, k, \gamma) \geq c_m(\alpha)\}, \quad (3)$$

kde $g(m, k, \gamma) = \sqrt{m} \left(\frac{m+k}{m}\right) \left(\frac{k}{m+k}\right)^\gamma$, $\gamma \in [0, 1/2)$ je takzvaná hraniční funkce a $c_m(\alpha)$ je určeno tak, že jsou splněny následující podmínky

$$\lim_{m \rightarrow \infty} P(\tau(m) < \infty | H_0) \leq \alpha, \quad \lim_{m \rightarrow \infty} P(\tau(m) < \infty | H_1) = 1. \quad (4)$$

Skutečnou změnu tedy detekujeme s pravděpodobností jdoucí k 1, kdežto v nezměněném modelu zamítáme nulovou hypotézu pouze s pravděpodobností α .

K naplnění (4) využijeme asymptotické chování statistiky popsané pro obě uvažované hypotézy v následujících 2 větách (platí stejné vztahy jako pro statistiky zkoumané v [2] či [1]).

Věta 1. Necht' Y_1, Y_2, \dots jsou nezávislé náhodné veličiny splňující uvažovaný model s předpoklady (A),(B) a $\gamma \in [0, 1/2)$. Potom za platnosti nulové hypotézy (1) pro všechna $c > 0$

$$\lim_{m \rightarrow \infty} P\left(\sup_{1 \leq k < \infty} \frac{|Q(m, k)|}{g(m, k, \gamma)} \leq c\right) = P\left(\sup_{0 \leq t \leq 1} \frac{|W(t)|}{t^\gamma} \leq c\right),$$

kde $\{W(t), t \in [0, 1]\}$ je Wienerův proces.

Věta 2. Necht' Y_1, Y_2, \dots jsou nezávislé náhodné veličiny splňují uvažovaný model s předpoklady (A),(C) a $\gamma \in [0, 1/2)$. Potom za platnosti alternativní hypotézy (2)

$$\sup_{1 \leq k < \infty} \frac{|Q(m, k)|}{g(m, k, \gamma)} \xrightarrow{P} \infty, \quad m \rightarrow \infty.$$

Věta 1 nám umožňuje aproximovat kritické hodnoty pro testovací proceduru, zatímco Věta 2 zaručuje, že nastalá změna bude detekována s pravděpodobností jdoucí k 1.

4 Důkazy vět

Postupujeme analogicky jako v [3], blíže se tedy věnujeme pouze rozdílům způsobeným použitím kvantilů místo L_1 odhadu. Všechny konvergence jsou myšleny pro $m \rightarrow \infty$.

Předpoklady (A) umožňují použití Bahadurovy asymptotické reprezentace výběrového τ kvantilu \tilde{t}_m^τ (viz. [1]). Platí

$$\tilde{t}_m^\tau - t_0^\tau = \frac{\tau - \frac{1}{m} \sum_{i=1}^m I[Y_i < t_0^\tau]}{f(t_0^\tau)} + R_m \text{ sj.}, \quad (5)$$

kde $R_m = O(m^{-3/4}(\ln m)^{3/4})$ sj., tedy

$$\tilde{t}_m^\tau - t_0^\tau = \frac{\sum_{i=1}^m \varepsilon_i}{mf(u^\tau)} + R_m = O(m^{-1/2}(\ln \ln m)^{1/2}) \text{ sj.}, \quad (6)$$

neboť $\sum_{i=1}^m \varepsilon_i = O(\sqrt{m \ln \ln m})$ sj., jak plyne ze zákona iterovaného logaritmu.

Dále v důkazech uvažujeme m dostatečně velké tak, aby \tilde{t}_m^τ již sj. ležel v okolí t_0^τ , kde mají pozorování distribuční funkci shodnou s F resp. F_* . Jelikož jsou tréninková data Y_1, \dots, Y_m při začátku monitorování známa, můžeme veškeré výpočty provádět podmíněně na nich. V podmínce jsou tyto tréninková data reprezentována \tilde{t}_m^τ .

Důkaz Věty 1 je rozdělen do několika lemma, jenž mají stejné předpoklady jako věta. Definujeme pomocné náhodné veličiny Z_i , $i = m+1, m+2, \dots$ vztahem $Z_i = \tilde{\varepsilon}_i - \varepsilon_i = -I[Y_i < \tilde{t}_m^\tau] + I[Y_i < t_0^\tau]$. Tyto veličiny jsou zřejmě podmíněně nezávislé.

Lemma 1. Podmíněná střední hodnota $\sum_{i=m+1}^{m+k} Z_i$ je

$$E \left(\sum_{i=m+1}^{m+k} Z_i | \tilde{t}_m^\tau \right) = -\frac{k}{m} \sum_{j=1}^m \varepsilon_j (1 + O(m^{-3/4}(\ln m)^{3/4})) \text{ sj.}$$

Důkaz. Rozlišíme 2 případy v závislosti na $\text{sign}(\tilde{t}_m^\tau - t_0^\tau)$.

i) Pro $\tilde{t}_m^\tau - t_0^\tau < 0$ máme $Z_i = I[\tilde{t}_m^\tau \leq Y_i < t_0^\tau]$ a tedy

$$E(Z_i | \tilde{t}_m^\tau) = P(\tilde{t}_m^\tau \leq Y_i < t_0^\tau | \tilde{t}_m^\tau) = \int_{\tilde{t}_m^\tau}^{t_0^\tau} f(x) dx,$$

neboť dle (6) platí $\tilde{t}_m^\tau \rightarrow t_0^\tau$ sj. Protože f' existuje na okolí t_0^τ platí

$$\sup\{|f(t_0^\tau) - f(x)|, x \in [\tilde{t}_m^\tau, t_0^\tau]\} = o(1).$$

Proto $E(Z_i|\tilde{t}_m^\tau) = f(t_0^\tau)(t_0^\tau - \tilde{t}_m^\tau)(1 + o(1))$ sj.

ii) Pro $\tilde{t}_m^\tau - t_0^\tau \geq 0$ máme $Z_i = -I[t_0^\tau \leq Y_i < \tilde{t}_m^\tau]$ a analogicky jako v předchozím případě dostaneme $E(Z_i|\tilde{t}_m^\tau) = -f(t_0^\tau)(\tilde{t}_m^\tau - t_0^\tau)(1 + o(1))$ sj.

Tedy bez ohledu na znaménko $\tilde{t}_m^\tau - t_0^\tau$ platí

$$E(Z_i|\tilde{t}_m^\tau) = -f(t_0^\tau)(\tilde{t}_m^\tau - t_0^\tau)(1 + o(1)) \text{ sj.}$$

Odtud pro celou sumu dostáváme

$$E\left(\sum_{i=m+1}^{m+k} Z_i|\tilde{t}_m^\tau\right) = \sum_{i=m+1}^{m+k} E(Z_i|\tilde{t}_m^\tau) = -kf(t_0^\tau)(\tilde{t}_m^\tau - t_0^\tau)(1 + o(1)) \text{ sj.}$$

a použitím (6) máme

$$\begin{aligned} E\left(\sum_{i=m+1}^{m+k} Z_i|\tilde{t}_m^\tau\right) &= -kf(t_0^\tau)\left(\frac{\sum_{j=1}^m \varepsilon_j}{mf(t_0^\tau)} + R_m\right)(1 + o(1)) = \\ &= \left(-\frac{k}{m}\sum_{j=1}^m \varepsilon_j - kf(t_0^\tau)R_m\right)(1 + o(1)) = -\frac{k}{m}\sum_{j=1}^m \varepsilon_j(1 + O((\ln m/m)^{3/4})) \text{ sj.} \end{aligned}$$

□

Lemma 2. Pro podmíněný rozptyl $\sum_{i=m+1}^{m+k} Z_i$ platí

$$\frac{1}{k} \text{var}\left(\sum_{i=m+1}^{m+k} Z_i|\tilde{t}_m^\tau\right) = O(m^{-1/2}(\ln \ln m)^{1/2}) \text{ sj.}$$

stejněměrně v k .

Důkaz. Při stejném rozdělení v závislosti na $\text{sign}(\tilde{t}_m^\tau - t_0^\tau)$ jako v Lemma 1 platí.

i) $Z_i^2 = Z_i$ a tedy

$$\text{var}(Z_i|\tilde{t}_m^\tau) = E(Z_i^2|\tilde{t}_m^\tau) - [E(Z_i|\tilde{t}_m^\tau)]^2 = E(Z_i|\tilde{t}_m^\tau)(1 - E(Z_i|\tilde{t}_m^\tau)).$$

Z důkazu Lemma 1 víme, že

$$E(Z_i|\tilde{t}_m^\tau) = (\tilde{t}_m^\tau - t_0^\tau)O(1) = O(m^{-1/2}(\ln \ln m)^{1/2}) \text{ sj.},$$

a proto $\text{var}(Z_i|\tilde{t}_m^\tau) = O(m^{-1/2}(\ln \ln m)^{1/2})$ sj.

ii) $Z_i^2 = -Z_i$ a analogicky

$$\text{var}(Z_i|\tilde{t}_m^\tau) = E(Z_i|\tilde{t}_m^\tau)(1 + E(Z_i|\tilde{t}_m^\tau)) = O(m^{-1/2}(\ln \ln m)^{1/2}) \text{ sj.}$$

Podmíněná nezávislost Z_i dává tvrzení lemma.

□

Lemma 3. Platí

$$\sup_{1 \leq k < \infty} \frac{\left| \sum_{i=m+1}^{m+k} (Z_i - \mathbb{E}(Z_i | \tilde{t}_m^\tau)) \right|}{g(m, k, \gamma)} = o_p(1).$$

Důkaz. Viz. Lemma 5.3 v [3]. \square

Lemma 3 může být s pomocí Lemma 1 formulováno také následovně

$$\sup_{1 \leq k < \infty} \frac{\left| \sum_{i=m+1}^{m+k} \tilde{\varepsilon}_i - \sum_{i=m+1}^{m+k} \varepsilon_i + \frac{k}{m} \sum_{i=1}^m \varepsilon_i \right|}{g(m, k, \gamma)} = o_p(1),$$

kde již první sčítanec představuje (až na normování) testovou statistiku $Q(m, k)$. Zbytek důkazu je již zcela analogický postupu uvedeném v [2], tj. pomocí Komlós-Major-Tusnádýovy aproximace zbylé 2 sumy aproximujeme Wienerovými procesy, jež se v distribuci shodují s funkcí Wienerova procesu z Věty 1.

Důkaz Věty 2. Nenormovanou testovou statistiku $Q_1(m, k)$ rozdělíme na část před změnou distribuční funkce pozorování a po ní, kdy se distribuční funkce Y_i v okolí t_0^τ shoduje s F_*

$$Q_1(m, k) = Q_1^{(1)}(m, k) + Q_1^{(2)}(m, k) = \sum_{i=m+1}^{m+k^*-1} \tilde{\varepsilon}_i + \sum_{i=m+k^*}^{m+k} \tilde{\varepsilon}_i.$$

Z Věty 1 plyne, že první sčítanec dělený $g(m, k, \gamma)$ (rostoucí funkce v k) je omezen v pravděpodobnosti. K důkazu stačí tedy najít posloupnost $\{k_m\}$ takovou, že $k_m - k^* \rightarrow \infty$ a $Q_1^{(2)}(m, k_m) \xrightarrow{P} \infty$ rychleji než $g(m, k_m, \gamma)$. Statistiku

$$Q_1^{(2)} := Q_1^{(2)}(m, k_m) = \sum_{i=m+k^*}^{m+k_m} (\tau - I[Y_i < \tilde{t}_m^\tau])$$

aproximujeme podmíněnou střední hodnotou.

Pro $\varepsilon > 0$ označme $B_m^\tau(\varepsilon)$ množinu, na které $|\tilde{t}_m^\tau - t_0^\tau| < \varepsilon$. Z (6) máme $\tilde{t}_m^\tau \rightarrow t_0^\tau$ sj. a tedy $P(B_m^\tau(\varepsilon)) \rightarrow 1$. Necht' například $F_*(t_0^\tau) > F(t_0^\tau)$. Potom na $B_m^\tau(\varepsilon)$ pro vhodně zvolené malé ε a $i \geq m + k^*$ platí

$$P(Y_i < \tilde{t}_m^\tau | \tilde{t}_m^\tau) = F_*(\tilde{t}_m^\tau) > F_*(t_0^\tau - \varepsilon) > F(t_0^\tau) = \tau \text{ sj.}$$

a proto $\mathbb{E}(\tilde{\varepsilon}_i | \tilde{t}_m^\tau) = \tau - P(Y_i < \tilde{t}_m^\tau | \tilde{t}_m^\tau) := -c_1 < 0$ sj. Obdobně pro druhý možný vztah uvažovaných distribucí dostáváme $\mathbb{E}(\tilde{\varepsilon}_i | \tilde{t}_m^\tau) := c_2 > 0$ sj. Tedy bez ohledu typ změny

$$\mathbb{E} \left| (Q_1^{(2)} | \tilde{t}_m^\tau) \right| \geq (k_m - k^* + 1) c_3 \text{ sj.}, \quad (7)$$

kde $c_3 = \min(c_1, c_2)$.

Pro podmíněný rozptyl platí

$$\text{var}(\tilde{\varepsilon}_i | \tilde{t}_m^\tau) = \text{var}(I[Y_i < \tilde{t}_m^\tau] | \tilde{t}_m^\tau) = \text{EI}[Y_i < \tilde{t}_m^\tau] (1 - \text{EI}[Y_i < \tilde{t}_m^\tau]) < 1 \text{ sj.}$$

Použitím Čebyševovy nerovnosti

$$P\left(\left|Q_1^{(2)} - \text{E}(Q_1^{(2)} | \tilde{t}_m^\tau)\right| \geq A\right) \leq \frac{\text{var}(Q_1^{(2)} | \tilde{t}_m^\tau)}{A^2} < \frac{k_m - k^* + 1}{A^2},$$

pro libovolné $A > 0$. Dosazením $A = B(k_m - k^* + 1)^{1/2}$, $B > 0$ dostáváme

$$\left|Q_1^{(2)} - \text{E}(Q_1^{(2)} | \tilde{t}_m^\tau)\right| = O_p((k_m - k^*)^{1/2}). \quad (8)$$

Hraniční funkce může být pro $k > m$ odhadnuta

$$\begin{aligned} g(m, k, \gamma) &= \sqrt{m} \left(1 + \frac{k}{m}\right) \left(\frac{k}{m+k}\right)^\gamma = m^{-1/2} (m+k)^{1-\gamma} k^\gamma \leq \\ &\leq c_4 m^{-1/2} k^{1-\gamma+\gamma} = c_4 m^{-1/2} k, \end{aligned} \quad (9)$$

pro nějaké $c_4 > 0$.

Spojením (7), (8) a (9) dostáváme

$$\frac{|Q_1^{(2)}(m, k_m)|}{g(m, k_m, \gamma)} \geq \frac{(k_m - k^*) c_3 (1 + O_p((k_m - k^*)^{-1/2}))}{c_4 m^{-1/2} k_m} \geq c_5 m^{1/2},$$

pro nějaké $c_5 > 0$, což dokončuje důkaz. □

5 Simulační studie

V programu R byly provedeny simulace pro ohodnocení výkonnosti procedury s konečným tréninkovým obdobím při použití asymptotických kritických hodnot, tj. kritických hodnot pro uvedený funkcionál Wienerova procesu. Uvažovaly se vlivy

- délky historického období ($m = 50; 100; 500$)
- okamžiku změny ($k^* = 50; 100; 500$)
- doladovací konstanty ($\gamma = 0; 0,25; 0,45$)
- rozdělení chyb (normální a Laplaceovo)
- typu alternativy - změna střední hodnoty pozorování o δ a změna rozptylu pozorování

pro různé kvantily ($\tau = 0,1; 0,25; 0,5; 0,75; 0,9$). Vždy bylo použito 5000 simulací. To vše na 5% hladině spolehlivosti.

Jelikož se v testové statistice neobjevují přímo velikosti pozorování, není procedura citlivá na rozdělení chyb, jako to bývá u procedur založených na L_2 odhadech.

$\tau \backslash m$	50	100	500
0,5	3,9	4,1	3,7
0,25	4,2	4,3	3,9
0,1	9,7	6,8	4,7

Tabulka 1: Empirické velikosti testu v závislosti na délce tréninkového období a uvažovaném kvantilu (pro $\gamma = 0, 25$).

τ	$\delta \backslash m$	50	100	500
0,5	0,5	360	246	206
	2	176	156	151
	-2	173	156	151
0,25	0,5	1000	782	335
	2	322	228	199
	-2	135	129	128
0,75	0,5	249	207	187
	2	135	129	128
	-2	322	235	199

Tabulka 2: Mediány času zastavení pro různé velikosti změny střední hodnoty δ , $k^* = 100$, $\gamma = 0,25$.

0,1	0,25	0,5	0,75	0,9
168	376	1000	365	167

Tabulka 3: Mediány času zastavení při ztrojnásobení směrodatné odchylky pozorování, $m = k^* = 100$, $\gamma = 0,25$.

V Tabulce 1 jsou shrnuty výsledky za platnosti H_0 - procenta překročení asymptotické kritické hodnoty, tj. procenta chybných zastavení pro monitorovací období délky 10000. K překročení předepsané hladiny dochází tedy jen pro extrémní kvantily a zejména malé m .

Za alternativní hypotézy procedura vykazuje stejné rysy jako obdobné procedury, a to pro všechny kvantily (speciálně pro medián jsou výsledky shodné s [3]) To znamená:

- doladovací konstanta γ - pro k^* malé je nevhodnější γ blízké $1/2$, naopak pro pozdní změnu $\gamma = 0$. Konstanta $0,25$ je nevhodnější pro k^* srovnatelné s m
- zhoršení rychlosti detekce s nárůstem k^* - typický problém CUSUM procedur
- prodloužení historického období - zlepšení rychlosti detekce; výrazné pro malé změny ($\delta = 1/2 \cdot \text{sd}(Y_i)$)

- větší změny (nad $2 \cdot \text{sd}(Y_i)$) se již neprojevují zrychlením detekce - typické pro „L₁” procedury

Oproti jiným procedurám se však musí rozlišovat znaménko změny. Zatímco pro medián je zpoždění detekce pro změnu $+\delta$ resp. $-\delta$ prakticky stejné, u jiných kvantilů je situace rozdílná. Jak je vidět v Tabulce 2, pokud změna nastala „ve směru kvantilu”, tj. např. záporná změna pro 1. kvartil, dochází ke zrychlení detekce, naproti tomu při opačné změně dochází ke zpomalení. Pro extrémnější kvantily je situace ještě výraznější. Tento fakt je způsoben tvarem definice $\tilde{\varepsilon}_i$. V tabulce také můžeme pozorovat zrychlení detekce s růstem m u malé změny.

V Tabulce 3 uvažujeme situaci v modelu, kdy se distribuce chyb měnila z $N(0,1)$ na $N(0,9)$. Medián distribuce se nezměnil, nedetekuje se tedy pro něj žádná změna (1000 odpovídá max. monitorovacímu období).

Literatura

- [1] Bahadur R.R. (1966). *A note on quantiles in large samples*. Ann. Math. Statist., **37**, 1966.
- [2] Horváth L., Hušková M., Kokoszka P., Steinebach J. (2004). *Monitoring changes in linear models*. J. Stat. Plann. Inference, **126**, 225–251.
- [3] Koubková A. (2006). *Sequential change-point analysis*. PhD thesis, Faculty of Mathematics and Physics, Charles University Prague.

Poděkování: Práce byla částečně podporována grantem GAČR 201/09/J006. Dík též patří paní prof. Huškové za věcné připomínky a odbornou pomoc a ČSOB za umožnění účasti na konferenci.

Adresa: O. Chochola, MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8 – Karlín

E-mail: chochola@karlin.mff.cuni.cz

ŘEŠENÍ PROBLÉMŮ ANALÝZY OBRAZU POMOCÍ MINIMA TOTÁLNÍ VARIACE

Jiří Janáček

Klíčová slova: Totální variace, minimální řez.

Abstrakt: Práce je věnována řešení úloh analýzy obrazu minimalizací funkce složené z totální variace obrazu a L^p ztrátové funkce. Minimalizaci lze provést iterační metodou největšího spádu s výpočtem kroku založeným na nalezení minimálního řezu v grafu obrazových prvků. Jako příklad je uvedena regularizace zašuměného obrazu a registrace mikroskopických obrazů fyzických řezů.

Abstract: A solution of various problems in image analysis using concurrent minimization of total variation and L^p loss function is presented. The minimization is achieved by a steepest descend method using graph cut minimization in each step. Regularization of noisy image and registration of microscopic images of physical sections are demonstrated.

1 Úvod

Řadu problémů vznikajících při analýze obrazu a v počítačovém vidění lze formulovat jako optimalizační úlohy s vhodnými hodnotovými funkcemi. Jako příklad může sloužit filtrace obrazu, kde hledáme hladký obraz s malou vzdáleností od daného zašuměného obrazu. Časem byla vyzkoušena řada deterministických i stochastických metod na řešení podobných optimalizačních úloh, ale v obtížnějších případech trvaly stochastické metody neúměrně dlouho a deterministické metody končily v lokálním optimu. Průlom znamenalo nalezení odhadu s maximální a posteriori pravděpodobností v binárním Isingově modelu pomocí minimálního řezu ve vhodném grafu [2].

Ukázalo se poté, že podobnou metodu lze použít i pro další úlohy, například filtraci šedotónového obrazu, segmentaci nebo korespondenci mezi dvěma obrazy s využitím při registraci obrazů, ve stereoskopickém vidění a při výpočtu optického toku. Tento příspěvek má na úlohách filtrace obrazu a registrace mikroskopických snímků fyzických řezů demonstrovat zvlášť jednoduchou implementaci optimalizační metody založené na minimálním řezu. Uvedené formule pro čtvercové obrazy s horizontální a vertikální souvislostí mezi pixely lze snadno zobecnit na libovolný tvar a dimenzi obrazu a libovolnou sousednost mezi obrazovými prvky.

2 Řešení diskrétní $TV - L^p$ regularizace metodou největšího spádu

Mějme digitální obraz g a označme u hledaný regularizovaný (vyhlazený) obraz

$$g : Pix \rightarrow \{0 \dots 255\}, u : Pix \rightarrow \{0 \dots 255\},$$

kde Pix je mřížka obrazových prvků (pixelů)

$$Pix = \{\{i, j\}; i, j = 1 \dots N\}.$$

Označme dále Edg množinu vertikálních nebo horizontálních hran mezi sousedními obrazovými prvky

$$Edg = \{\{\{i, j\}, \{i+1, j\}\}; i = 1 \dots N-1, j = 1 \dots N\} \\ \cup \{\{\{i, j\}, \{i, j+1\}\}; i = 1 \dots N, j = 1 \dots N-1\}.$$

Jako kritérium pro volbu regularizovaného obrazu použijeme hodnotovou funkci

$$F_p(u) = \sum_{e \in Edg} |u_{e_1} - u_{e_2}| + \frac{1}{p\lambda} \sum_{x \in Pix} |u_x - g_x|^p$$

a u volíme jako argument minima F_p .

První člen funkce je diskrétní totální variace $TV(u)$ a penalizuje nehladké obrazy, druhý člen je násobek p -té mocniny L^p normy rozdílu obrazů, parametr λ řídí poměr mezi požadovanou hladkostí řešení a jeho blízkostí ke vstupnímu obrazu.

F_p je zřejmě konvexní pro $p \geq 1$ a striktně konvexní pro $p > 1$ a $F_p(u) \rightarrow \infty$ pro $u \rightarrow \infty$. Úloha tedy má řešení, pro $p > 1$ dokonce právě jedno.

2.1 Příklad:

Nechť $\gamma > 0$, S je množina o ploše A , I_S je její indikátorová funkce, $g(x) = \gamma I_S(x)$, pak $TV(g) = \gamma TV(I_S)$, $\|g\|_p^p = \gamma^p A$. Řešení minimalizačního problému je 0 pro $\frac{\gamma^p}{p\lambda} A = \frac{1}{p\lambda} \|g\|_p^p < TV(g) = \gamma TV(I_S)$ a g při opačné nerovnosti. Ekvivalentní nerovnost je $\gamma^{p-1} A < p\lambda TV(I_S)$. $TV(I_S)$ je přibližně úměrná obvodu objektu S . Pro $p = 1$ je řešení nezávislé na γ , takže $TV - L^1$ regularizace odstraňuje z obrazu objekty s velkým poměrem obvodu k ploše (např. malé objekty) bez ohledu na kontrast.

2.2 Minimalizace F_p metodou největšího spádu

Nechť $\delta \in \mathbb{Z}$, I je indikátorová funkce.

V každém kroku metody nalezneme $S \subseteq Pix$ tak, že $F_p(u + \delta I_S)$ je minimální a u nahradíme $u + \delta I_S$ [1].

F je konvexní a tedy postupné iterace $\delta = \pm 2^n$, prováděné pro každé $n = 7, 6, \dots, 0$ tak dlouho dokud funkce klesá, dokonvergují po konečném počtu kroků ke globálnímu minimu.

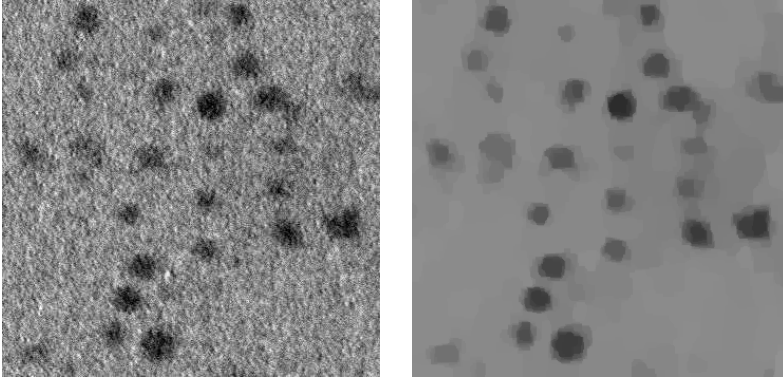
Vhodné $S \subseteq Pix$ v každém kroku najdeme pomocí minimálního řezu ve vhodném grafu. Vytvoříme ohodnocený orientovaný graf na množině vrcholů $Pix \cup \{source, sink\}$. Minimální hodnota $F_p(u + \delta I_S)$ odpovídá hodnotě nejmenšího řezu mezi *source* a *sink*, kde ohodnocení hran je

$$x \in Pix : h(source, x) = \frac{1}{p\lambda} |u_x + \delta - g_x|^p, \quad h(x, sink) = 0,$$

$$\{x, y\} \in Edg : h(x, y) = |u_x - u_y - \delta|.$$

K nalezení minimálního řezu můžeme použít větu o dualitě, podle níž se minimální řez mezi *source* a *sink* rovná maximálnímu toku mezi *source* a *sink*, a dostupné programy pro nalezení maximálního toku [3].

2.3 Detekce nanočástic



Snímek nanočástic stříbra (Silver proteinate, Fluka, Švýcarsko) před $TV - L^1$ regularizací a po ní. Snímky z transmisního elektronového mikroskopu laskavě poskytl Anatolij Filimoněnko z ÚMG AVČR. Filtrace zachovala i málo kontrastní objekty a výsledný obraz je vhodný pro další zpracování a detekci částic, např. adaptivním prahováním.

3 $TV - L^1$ korespondence metodou největšího spádu

Mějme digitální obrazy u a v a označme $\xi = (\xi^1, \xi^2)$ hledaný vektorový obraz korespondence přiřazující obrazovému prvku x obrazu u prvek $x + \xi_x$ obrazu v

$$u : Pix \rightarrow \mathbb{R}, \quad v : Pix \rightarrow \mathbb{R},$$

$$\xi : Pix \rightarrow \mathbb{Z}^2.$$

Jako kritérium pro volbu obrazu korespondence použijeme funkci

$$F(\xi) = \sum_{e \in Edg} |\xi_{e_1}^1 - \xi_{e_2}^1| + |\xi_{e_1}^2 - \xi_{e_2}^2| + \frac{1}{\lambda} \sum_{x \in Pix} |u_x - v_{x+\xi_x}|$$

a ξ volíme jako argument minima funkcionálu F .

3.1 Minimalizace metodou největšího spádu

Nechť $\delta \in \mathbb{Z}^2$, $\delta = (\delta^1, \delta^2)$, I je indikátorová funkce.

V každém kroku metody nalezneme $S \subseteq Pix$ tak, že $F(\xi + \delta I_S)$ je minimální a ξ nahradíme $\xi + \delta I_S$. Protože F nemusí být konvexní, konvergence ke globálnímu minimu není jistá. δ můžeme volit například tak, že se jeho směry vhodně mění a jeho velikost se v následných iteracích zmenšuje.

$S \subseteq Pix$ najdeme metodou minimálního řezu. Vytvoříme ohodnocený orientovaný graf na množině vrcholů $Pix \cup \{source, sink\}$. Minimální hodnota $F(u + \delta I_S)$ odpovídá hodnotě nejmenšího řezu mezi *source* a *sink*, kde ohodnocení hran je

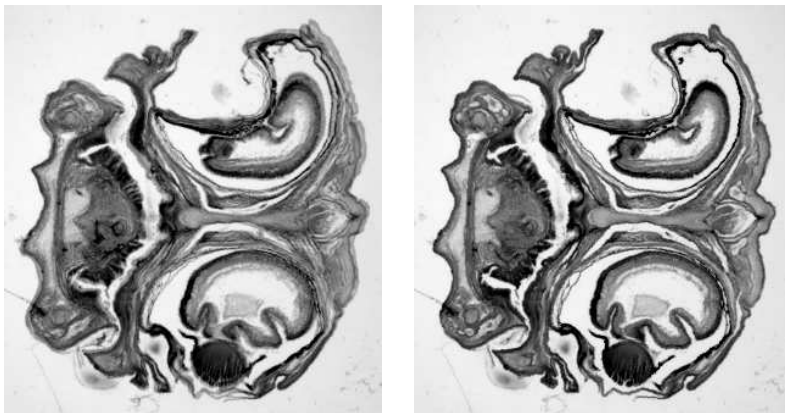
$$x \in Pix : h(source, x) = \frac{1}{\lambda} |u_x - v_{x+\xi_x+\delta}|, h(x, sink) = 0,$$

$$\{x, y\} \in Edg : h(x, y) = |\xi_x^1 - \xi_y^1 - \delta^1| + |\xi_x^2 - \xi_y^2 - \delta^2|.$$

K nalezení minimálního řezu můžeme opět použít větu o dualitě nejmenšího řezu a největšího toku [3].

3.2 Registrace snímků fyzických řezů

Pro rekonstrukci 3D objektu ze snímků paralelních fyzických řezů je obvykle nutné nalézt vzájemnou polohu řezů a kompenzovat případné deformace způsobené krájením.



Průměr dvou snímků sousedních fyzických řezů hlavou želvy před registrací a po ní. Snímky pořídila Barbora Tvarožková z PřF UK. Můžeme vidět úspěšnou elastickou registraci celého řezu včetně nespojitého přechodu mezi lebkou (vpravo) a čelistí (vlevo).

4 Závěr

V řešení optimalizačních úloh ve zpracování obrazu bylo v poslední době dosaženo znatelného pokroku díky kombinatorické optimalizaci založené na minimálním řezu v grafu. Na rozdíl od naivní optimalizace, při které bychom procházeli obraz postupně a v každém pixelu hledali optimální změnu, v demonstrované metodě hledáme k předem zvolenému kroku množinu, na níž změna o tento krok vyvolá největší pokles v penalizační funkci. Zvýšená výpočetní náročnost metody je vyvážena zvýšenou robustností k lokálním minimům.

Literatura

- [1] K. Murota (2003). *On steepest descent algorithms for discrete convex functions*. SIAM J. Optim. **14**, 699–707.
- [2] D.M. Greig, B.T. Porteous, A.H. Seheult (1989). *Exact maximum a posteriori estimation for binary images*. J. R. Statist. Soc. **B 51**, 271–279.
- [3] V. Kolmogorov, R. Zabih (2004.) *What energy functions can be minimized via graph cuts?* IEEE TPAMI **26**, 147–159.

Poděkování: Tato práce byla podporována granty AVČR č. A100110502 a AV0Z 50110509 a grantem GAČR 102/08/0691.

Adresa: J. Janáček, FgÚ AVČR, Vídeňská 1083, 142 20 Praha 4 – Krč

E-mail: janacek@biomed.cas.cz

ON THE IMPORTANCE OF ENTROPY

Martin Janžura

Keywords: Entropy, I-divergence, maximum likelihood, I-projection, large deviations, conditional limit theorem.

Abstract: The aim of the paper consists in demonstrating the relevance of the fundamental information-theoretic concepts, namely the entropy and the I-divergence, for both the statistical inference and the limit theorems of probability theory.

Abstrakt: Cílem je ukázat význam základních pojmů teorie informace, tj. entropie a I-divergence, jak pro statistické úlohy tak i pro limitní věty teorie pravděpodobnosti.

1 Introduction

Entropy, as a fundamental concept arising from statistical physics, was originally understood as a thermodynamical property of heat engines (see, e.g., [8] or [12]). Later, in the pioneering Shannon's paper [13] it was introduced as a crucial quantity of information theory. Its relevance as a measure of uncertainty (ignorance, information) was early recognized and widely exploited (see [3], [4], [10]). As the number of different names for the same concept indicates, the more general quantity of I-divergence (relative entropy, Kullback-Leibler number, information gain) became also widely used (see again [3], [4], or [11], [14], and the references therein).

The purpose of the present paper consists in showing even more fundamental importance of the concepts for the area of mathematical statistics and probability theory. We intend to illustrate how the quantities are inherent, especially, for the multidimensional joint distributions, and how they arise naturally in various situations.

Similar contents as here can be found, e.g., in [8]. For many topics of the present paper, in particular for the limit theorems, [3] or [4] are the basic references. For exponential distributions, closely related to the maximum entropy principle, see [2], or, more generally, [5]. For statistical problems see [11] or [14]. The presented results can be also generalized from the I.I.D. case to the random processes or fields, see [6], [7], [9], and [15].

2 Basic definitions and properties

Let us consider a finite state space $\mathcal{X} = \{x_1, \dots, x_M\}$.

By $\mathcal{P}(\mathcal{X})$ we denote the class of all probability measures on \mathcal{X} , and by $\mathcal{F}(\mathcal{X})$ the class of all real-valued functions on \mathcal{X} . In particular, by $R \in \mathcal{P}(\mathcal{X})$ we denote the *uniform* distribution, i.e., $R(x) = \frac{1}{M}$ for every $x \in \mathcal{X}$.

Let us recall the formulas for the *entropy* and the *I-divergence*, respectively, namely

$$H(P) = \int -\log P \, dP = \sum_{x \in \mathcal{X}} -\log P(x) P(x),$$

for $P \in \mathcal{P}(\mathcal{X})$, and

$$I(P|Q) = \int \log \frac{P}{Q} \, dP = \sum_{x \in \mathcal{X}} \log \frac{P(x)}{Q(x)} P(x)$$

for $P, Q \in \mathcal{P}(\mathcal{X})$, providing the terms are well defined. Otherwise we set $I(P|Q) = \infty$.

Proposition 1. For $P, Q \in \mathcal{P}(\mathcal{X})$ it holds

- i) $H(P) \in [0, \log M]$;
- ii) $H(P) = 0$ iff $P(x_i) = 1$ for some $i \in \{1, \dots, M\}$ and $P(x_j) = 0$ for every $j \neq i$;
- iii) $H(P) = \log M$ iff $P = R$.
- iv) $I(P|Q) \in [0, \infty]$;
- v) $I(P|Q) = 0$ iff $P = Q$;
- vi) $I(P|Q) = \infty$ iff $P \not\ll Q$.

Proof. $H(P) \geq 0$ follows from $-\log P(x) \geq 0$ for every $x \in \mathcal{X}$. Thus $H(P) = 0$ can occur only if $-\log P(x) \equiv 0$ which proves ii). $H(P) \leq \log M$ and iii) follow from iv) and v) with $Q = M$.

iv) and v) are due to Jensen's inequality, vi) is obvious. \square

Remark.

Due to the above properties of the entropy, which is minimal for non-random case and maximal for the uniform distribution, it can be understood as a measure of *uncertainty* contained in the probability distribution. Namely, if the entropy is zero then the output is sure, if it is maximal then all possible outputs are equally likely.

The *I-divergence*, on the other hand, can serve as a *distance*, being equal to zero for identical distributions, and maximal, equal to ∞ , if the first distribution is not supported on the same set as the second one.

3 Maximum entropy principle

As a rule, whenever we have no information about the distribution of some random phenomenon, we turn to the uniform distribution. It is justified by the fact that we have *no reason* for preferring any particular output. Thus, in the light of the above Remark, we opt for the distribution with maximum entropy.

Such approach can be extended (see, e.g., [10] as the standard reference) to the general *maximum entropy principle* (MAXENT). Suppose we have only a partial information about the distribution, namely $P \in \mathcal{E}$ where $\mathcal{E} \subset \mathcal{P}(\mathcal{X})$.

Then, applying the MAXENT, we seek for

$$\bar{P}_{\mathcal{E}} \in \operatorname{argmax}_{P \in \mathcal{E}} H(P)$$

or, more generally,

$$\bar{P}_{\mathcal{E}} \in \operatorname{argmin}_{P \in \mathcal{E}} I(P|Q)$$

where $Q \in \mathcal{P}(\mathcal{X})$ is some fixed reference probability measure.

Usually, the first definition, which, after all, agrees with the latter one for the uniform $Q = R$, is meant by the maximum entropy principle.

Example (maximum entropy with linear constraints). Let us consider a collection of statistics $\mathbf{f} = \{f_j\}_{j \in \mathcal{K}}$ with $|\mathcal{K}| < \infty$, where $f_j \in \mathcal{F}(\mathcal{X})$ for every $j \in \mathcal{K}$. Moreover, in order to guarantee the basic *regularity (identifiability) condition*, we assume the system $(1, \{f_j\}_{j \in \mathcal{K}})$ to be linearly independent.

Now, for a collection of constants $\mathbf{m} = \{m_j\}_{j \in \mathcal{K}}$ we denote

$$\mathcal{E} = \mathcal{M}(\mathbf{m}, \mathbf{f}) = \left\{ P \in \mathcal{P}; \int f_j \, dP = m_j \quad \text{for every } j \in \mathcal{K} \right\}.$$

Further, let us introduce the *exponential distribution* P^α given by

$$P^\alpha(x) = \exp \left\{ \sum_{j \in \mathcal{K}} \alpha_j f_j(x) - c(\alpha) \right\}$$

where $\alpha = (\alpha_j)_{j \in \mathcal{K}} \in R^{\mathcal{K}}$ is a parameter, and the appropriate normalizing constant is given by $c(\alpha) = \log \sum_{x \in \mathcal{X}} \exp \left\{ \sum_{j \in \mathcal{K}} \alpha_j f_j(x) \right\}$.

Then, we may deduce the following properties:

i) There is a one-to-one relation between the parameter α and the exponential distribution P^α . Namely, for $P^\alpha = P^\beta$ we have $\langle \alpha - \beta, \mathbf{f} \rangle = \text{const.}$, and the statement holds thanks to the identifiability condition above.

ii) Let $P^\alpha, P^\beta \in \mathcal{E}$. Then $\alpha = \beta$. We observe

$$0 \leq I(P^\alpha|P^\beta) + I(P^\beta|P^\alpha) = \langle \beta - \alpha, \int \mathbf{f} \, dP^\beta - \int \mathbf{f} \, dP^\alpha \rangle = 0.$$

Hence $P^\alpha = P^\beta$, and, due to i), we have $\alpha = \beta$.

iii) Let $P^\alpha \in \mathcal{E}$. Then $\bar{P}_{\mathcal{E}}$ is given uniquely, and $\bar{P}_{\mathcal{E}} = P^\alpha$.

As it is well-known, we have

$$0 \leq I(P|P^\alpha) = c(\alpha) - \langle \alpha, \mathbf{m} \rangle - H(P) = H(P^\alpha) - H(P)$$

where, by Proposition 1.v), the inequality turns into equality iff $P = P^\alpha$.

We may conclude that whenever there exists the *exponential representative* $P^{\bar{x}} \in \mathcal{E} = \mathcal{M}(\mathbf{m}, \mathbf{f})$ then it is given uniquely and satisfies the MAXENT. Thus, the exponential distributions maximize the entropy under the linear constraints, and, since the linear constraints are rather standard form of partial information (see, e.g., the moment conditions in mathematical statistics), the exponential families are well justified as probability models.

4 I.I.D. sequences

Let us consider a sequence $\bar{x}_n = (x_1, \dots, x_n)$ where $x_i \in \mathcal{X}$ for every $i = 1, \dots, n$. Let us denote by $N^{\bar{x}_n}(y) = \sum_{i=1}^n \delta(y, x_i)$ the *number of occurrences* of the state $y \in \mathcal{X}$ in the sequence \bar{x}_n . Consequently, we shall denote by $P^{\bar{x}_n} = \frac{1}{n} N^{\bar{x}_n}$ the *empirical distribution* induced by the sequence \bar{x}_n .

Now, we understand the sequence \bar{x}_n as a collection of data obtained from a sequence of I.I.D. random variables with a one-body marginal distribution $P \in \mathcal{P}(\mathcal{X})$. Then, for the joint distribution, we may easily observe

$$P_n(\bar{x}_n) = \prod_{i=1}^n P(x_i) = \prod_{y \in \mathcal{X}} P(y)^{N^{\bar{x}_n}(y)} = \exp\{-n[D(P^{\bar{x}_n}|P) + H(P^{\bar{x}_n})]\}.$$

As a result we may express the *log-likelihood* as

$$-\frac{1}{n} \log P_n(\bar{x}_n) = D(P^{\bar{x}_n}|P) + H(P^{\bar{x}_n}).$$

Thus, we may conclude that, first, the empirical distribution is a sufficient statistics for the joint distribution, and, moreover, the joint distribution depends on the empirical distribution just through the above quantities, namely the *entropy* and the *I-divergence*. Let us emphasize that such relation is not imposed or artificial, it is apparently natural and inherent for the joint probability distributions.

5 Statistical problems

5.1 Parameter estimation

In this section let us consider a *parametric family* of probability distributions $\{P^\theta\}_{\theta \in \Theta}$ where $P^\theta \in \mathcal{P}(\mathcal{X})$ for every $\theta \in \Theta$. For the sake of simplicity we shall assume $P^\theta > 0$ for every $\theta \in \Theta$ (we may, e.g., imagine the exponential family as introduced in Section 3).

Based on a data sequence $\bar{x}_n = (x_1, \dots, x_n)$, we may define the *maximum likelihood estimate* (MLE) standardly as: $\hat{\theta}^n = \arg \max_{\theta \in \Theta} \log P_n^\theta(\bar{x}_n)$.

But, due to the above formula for the log-likelihood, we have also

$$\hat{\theta}^n = \arg \min_{\theta \in \Theta} D(P^{\bar{x}_n}|P^\theta).$$

Thus, the MLE can be alternatively understood as the *Minimum I-divergence estimate* which provides us with an additional justification for the MLE: we seek for such value of the parameter that makes the theoretical distribution as close (in the sense of I-divergence) to the empirical one as possible. Conversely, if we stay within the framework of “minimum distance estimation” then the distance measured by the I-divergence is privileged since it yields the maximum likelihood estimation with all its favourable properties.

Remark(Consistency). The consistency of the MLE is in general rather well-known. Nevertheless, it can be simultaneously derived from the minimum I-divergence approach. Let us give a sketch of the proof:

Suppose $\theta_0 \in \Theta$ is the true parameter. Then, obviously,

$$\theta_0 = \arg \min_{\theta \in \Theta} D(P^{\theta_0} | P^\theta).$$

Denote $L_n(\theta) = D(P^{\bar{x}_n} | P^\theta)$ for every $n = 1, \dots$, and $L_0(\theta) = D(P^{\theta_0} | P^\theta)$. Then, due to the law of large numbers, we obtain $L_n(\theta) \xrightarrow{n \rightarrow \infty} L_0(\theta)$ a.s.[P] point-wise for every $\theta \in \Theta$. But, in order to prove

$$\hat{\theta}^n = \arg \min_{\theta \in \Theta} L_n(\theta) \xrightarrow{n \rightarrow \infty} \arg \min_{\theta \in \Theta} L_0 = \theta_0$$

we need the above convergence uniform at least on every compact set. The latter is satisfied if all the functions L_n are convex, which is the case, e.g., of the exponential families.

5.2 Testing hypotheses

Similarly as for the estimates, the likelihood ratio test may be understood as tests based on the I-divergence. In particular, for $\Theta = R^K$, let us consider the test of a simple hypothesis $H_0 : \theta = \theta_0$ against the alternative $H_1 : \theta \in \Theta \setminus \{\theta_0\}$. Then

$$2nD(P^{\hat{\theta}^n} | P^{\theta_0}) \xrightarrow{n \rightarrow \infty} \chi_K^2 \quad \text{in distribution} \quad [P^{\theta_0}].$$

More generally, for a subspace or a hyperplane $\Theta_0 \subset \Theta$, we may consider the affine hypothesis $H_0 : \theta \in \Theta_0$ against the alternative $H_1 : \theta \in \Theta \setminus \Theta_0$. Then

$$2nD(P^{\hat{\theta}^n} | P^{\hat{\theta}_0^n}) \xrightarrow{n \rightarrow \infty} \chi_{K-\dim(\Theta_0)}^2 \quad \text{in distribution} \quad [P^\theta]$$

for every $\theta \in \Theta_0$, where

$$\hat{\theta}_0^n = \arg \min_{\theta \in \Theta_0} D(P^{\bar{x}_n} | P^\theta).$$

For more detailed treatment see, e.g., [9] or [14].

6 I.I.D. sequences and types

From Section 3 we know $P_n(\bar{x}_n) = \exp\{-n[D(P^{\bar{x}_n}|P) + H(P^{\bar{x}_n})]\}$ providing x_1, \dots, x_n are drawn I.I.D. according to $P \in \mathcal{P}(\mathcal{X})$.

Now, let $\bar{x}_n \in T_n(Q) = \{\bar{x}_n : P^{\bar{x}_n} = Q\}$ for some $Q \in \mathcal{P}$, we say \bar{x}_n is in the *type class* of Q , then

$$P_n(\bar{x}_n) = \exp\{-n[D(Q|P) + H(Q)]\}.$$

In particular, for $\bar{x}_n \in T_n(P)$ we have $P_n(\bar{x}_n) = \exp\{-nH(P)\}$.

At the same time, by combinatorial arguments we can obtain $|T_n(P)| \doteq \exp\{nH(P)\}$ or, more precisely

$$\exp\{nH(P) + o(n)\} \leq |T_n(P)| \leq \exp\{nH(P)\}$$

(see, e.g. [3]), and, therefore

$$P_n[T_n(P)] \doteq 1$$

while, in general,

$$P_n[T_n(Q)] \doteq \exp\{-nD(Q|P)\}.$$

Thus, we may conclude that the joint distribution P_n is approximately supported on the set $T_n(P) \subset \mathcal{X}^n$, that contains approximately $e^{nH(P)}$ configurations. And each of the configurations has the equal probability $e^{-nH(P)}$. In such a way the joint distribution P_n is essentially determined by the entropy $H(P)$.

As a corollary we obtain the *law of large numbers*. Namely, for $g : \mathcal{X} \rightarrow R$ we obtain

$$\frac{1}{n} \sum_{i=1}^n g(x_i) = \int g dP^{\bar{x}_n} \doteq \int g dP.$$

7 Limit theorems

From the preceding section we know that every joint distribution P_n is concentrated on the configurations from its own type class $T_n(P)$. Nevertheless, we are still interested in the behaviour of P_n outside its type class. Such behaviour is again characterized by the entropy and related notions.

Definition(I-projection). For $\mathcal{E} \subset \mathcal{P}(\mathcal{X})$ we denote $D(\mathcal{E}|P) = \inf_{Q \in \mathcal{E}} D(Q|P)$. If there exists $P_* = \arg \min_{Q \in \mathcal{E}} D(Q|P)$ we call it *I-projection*.

Due to compactness of $\mathcal{P}(X)$ and continuity of $D(\bullet|P)$ we observe that P_* is attained if $\mathcal{E} \subset \mathcal{P}(\mathcal{X})$ is closed.

The proofs of the following results can be found. e.g., in [3] and [4], or, in the most general form, also in [7].

Theorem 1(Large deviations). If $\mathcal{E} = \text{cl}(\text{int}\mathcal{E})$ then

$$-\frac{1}{n} \log P_n(P^{\bar{x}_n} \in \mathcal{E}) \xrightarrow{n \rightarrow \infty} D(P_*|P).$$

□

The above theorem has many interesting consequences, as an example let us introduce the following result.

Corollary(Test power). When testing a simple hypothesis $\mathcal{H}_0 : P = P^0$ against a simple alternative $\mathcal{H}_1 : P = P^1$ with a level (first kind error) equal to $\alpha \in (0, 1)$ we obtain for the second kind error

$$-\frac{1}{n} \log \beta_n \xrightarrow{n \rightarrow \infty} D(P^0|P^1).$$

□

The theorem on large deviations proclaims that, whenever the distance $D(\mathcal{E}|P)$ is positive, the probability $P_n(P^{\bar{x}_n} \in \mathcal{E})$ is very small, it tends to zero exponentially fast. The following important and nice theorem shows what will happen if we still “enforce” the distribution to concentrate on such a “small” set.

Theorem 2 (Conditional limit theorem). Let \mathcal{E} be a closed convex subset of $\mathcal{P}(X)$ and $P \notin \mathcal{E}$. Then

$$P_n(\bullet|P^{\bar{x}_n} \in \mathcal{E}) \xRightarrow{n \rightarrow \infty} P_*^\infty \quad (\text{weakly}).$$

□

The conditional limit theorem proofs, e.g., one of the fundamental results of statistical physics.

Remark (the second law of thermodynamics). For $P = R$ we have $P_* = \arg \max_{Q \in \mathcal{E}} H(Q)$. That means the conditional distribution attains (at infinity) the maximum entropy.

Usually we have the set \mathcal{E} given by a linear constraint (see Section 3), i.e. $\mathcal{E} = \{Q \in \mathcal{P}(\mathcal{X}); \int E dQ \leq c\}$. Then we obtain the exponential distribution $P_* \propto \exp\{\alpha E\} \cdot P$ as the I-projection.

Remark (importance sampling). The above conditional limit theorem can be also used for approximating the conditional distribution, e.g., for the simulation method known as “importance sampling”.

Namely, for estimating some “small” probability $P(A)$ we can use $\widehat{P(A)} = \frac{1}{k} \sum_{i=1}^k I_A(y_i) \frac{P(y_i)}{Q(y_i)}$ where y_1, \dots, y_k are drawn I.I.D. with some $Q \in \mathcal{P}(\mathcal{X})$.

From the minimum variance point of view, the optimal choice is $Q = P(\bullet|A)$, and, whenever we have $A = A_n = \{\bar{x}_n; P^{\bar{x}_n} \in \mathcal{E}\}$, we may use the approximation $P_n(\bullet|A_n) \doteq P_*^n$ (see [1]).

References

- [1] Antoch J. (2006). *O simulaci řídkých jevů*. In: Robust 2006 (Eds.: J. Antoch, G. Dohnal), JČMF, Praha.
- [2] Barndorff-Nielsen O.E. (1978). *Information and Exponential Families in Statistical Theory*. Wiley, New York.
- [3] Cover T.M., Thomas J.A. (1991). *Elements of Information Theory*. Wiley, New York.
- [4] Csiszár I., Körner J. (1981). *Information Theory*. Akademiai Kiado, Budapest.
- [5] Csiszár I., Matúš F. (2006). *Generalize maximum likelihood estimates for exponential families*. Probability Theory and Related Fields **141**, 213–246.
- [6] Föllmer H. (1973). *On entropy and information gain in random fields*, Z. Wahrs. verw. Geb. **26**, 207–217.
- [7] Georgii H.-O. (1993). *Large deviations and maximum entropy principle for interacting random fields on Zd*. Ann. Prob. **21**, 1845–1875.
- [8] Georgii H.-O. (2003). *Probabilistic aspects of entropy*. In: Entropy (Eds.: A. Greven, G. Keller, G. Warnecke), Princeton University Press, 37–54.
- [9] Janžura M. (1997). *Asymptotic results in parameter estimation for Gibbs random fields*. Kybernetika **33**, 2, 133–159.
- [10] Jaynes E.T. (1982). *On the rationale of the maximum entropy methods*. Proc. IEEE **70**, 939–952.
- [11] Liese F., Vajda I. (2006). *On divergences and informations in statistics and information theory*, IEEE Transactions on Information Theory **52**, 10, 4394–4412.
- [12] Sethna J.P. (2006). *Statistical Mechanics: Entropy, Order Parameters, and Complexity*. Oxford University Press.
- [13] Shannon C.E. (1948). *A mathematical theory of communication*. Bell System Techn. J. **27**, 379–423, 623–657.
- [14] Vajda I. (1989). *Theory of Statistical Inference and Information*. Kluwer Academic Publishers, Dordrecht.
- [15] Winkler G. (1995). *Image Analysis, Random Fields and Dynamic Monte Carlo Methods*. Springer-Verlag, Berlin.

Acknowledgement: Supported by grant GA ČR No.201/06/1323 and Research Center DAR (MŠMT ČR Project No. 1M0572).

Address: M. Janžura, Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 4, Prague 8

E-mail: janžura@utia.cas.cz

TESTOVÁNÍ NORMALITY VE VÍCEROZMĚRNÉM PŘÍPADĚ

Daniela Jarušková

Klíčová slova: Normalita, empirický proces, šikmost, špičatost, Kolmogorovova-Smirnovova statistika, velké odchylky.

Abstrakt: Jsou uvažovány tři testy (Pearsonův, Koziolův a Kolmogorovův-Smirnovův) pro testování normality ve vícerozměrném případě.

Abstract: Three tests (Pearson's, Koziol's and Kolmogorov-Smirnov's tests) are considered for testing normality in a multivariate case.

1 Úvod

Testování normality ve vícerozměrném případě je komplikovaný problém, protože narušení normality se může dít mnoha různými způsoby. V řeči matematické statistiky to znamená, že máme sice jen jednu nulovou hypotézu, která tvrdí, že data jsou výběrem z normálního rozdělení, ale můžeme uvažovat nepřeborné množství alternativ.

O testování normality existuje řada článků a knih. Čtenářům, kteří by chtěli do problematiky nahlédnout hlouběji, doporučujeme monografii [2]. V našem příspěvku se budeme věnovat jen třem testům, které jsme pro jednoduchost nazvali Pearsonův, Koziolův a Kolmogorovův-Smirnovův.

2 Transformace

Studovaná data jsou p -rozměrné vektory $\mathbf{X}_1 = (X_{11}, \dots, X_{1p})^T, \dots, \mathbf{X}_n = (X_{n1}, \dots, X_{np})^T$. Nulová hypotéza tvrdí, že vektory $\mathbf{X}_1, \dots, \mathbf{X}_n$ můžeme považovat za výběr z p -rozměrného normálního rozdělení $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Pokud bychom $\boldsymbol{\mu}$ a $\boldsymbol{\Sigma}$ znali, bylo by možné založit naše rozhodnutí o normalitě na vektorech $\mathbf{Y}_1, \dots, \mathbf{Y}_n$, kde

$$\mathbf{Y}_i = \boldsymbol{\Sigma}^{-1/2}(\mathbf{X}_i - \boldsymbol{\mu}).$$

Poznámka. Vektory $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ mají nezávislé komponenty. Statistiky, které jsou počítány z různých komponent, jsou pak přirozeně nezávislé. Většina praktických statistiků začíná své šetření studiem normality marginálních rozdělení složek vektorů $\mathbf{X}_1, \dots, \mathbf{X}_n$. Pro závěrečné rozhodnutí mohou pak použít např. Bonferoniho nerovnost, to znamená uvažovat v jednotlivých testech místo hladiny významnosti α hladinu α/p .

Pokud $\boldsymbol{\mu}$ a $\boldsymbol{\Sigma}$ neznáme, můžeme je nahradit maximálně věrohodnými odhady, tj. vektorem průměrů $\bar{\mathbf{X}} = (\bar{X}_1, \dots, \bar{X}_p)^T$ a výběrovou kovarianční

maticí $\mathbf{S} = (s_{ij})_{i,j=1}^p$, kde $s_{ij} = (1/n) \sum_{k=1}^n (X_{ki} - \bar{X}_i)(X_{kj} - \bar{X}_j)$, a naše rozhodnutí založit na vektorech $\mathbf{Z}_1, \dots, \mathbf{Z}_n$, kde

$$\mathbf{Z}_i = \mathbf{S}^{-1/2}(\mathbf{X}_i - \bar{\mathbf{X}}).$$

Pro $j = 1, \dots, p$ platí $(1/n) \sum_{k=1}^n Z_{kj} = 0$ a $(1/n) \sum_{k=1}^n Z_{kj}^2 = 1$.

3 Pearsonův test

Jedním z nejpoužívanějších testů pro testování normality v jednorozměrném případě je D'Agostinův test založený na standardizované šikmosti a špičatosti. Pro výběr X_1, \dots, X_n je standardizovaná šikmost C_3 a špičatost C_4 definovaná následovně:

$$C_3 = \frac{1}{n} \sum_{k=1}^n \left(\frac{X_k - \bar{X}}{s} \right)^3 \sqrt{\frac{n}{6}}, \quad C_4 = \left(\frac{1}{n} \sum_{k=1}^n \left(\frac{X_k - \bar{X}}{s} \right)^4 - 3 \right) \sqrt{\frac{n}{24}}.$$

Za platnosti nulové hypotézy, která tvrdí, že X_1, \dots, X_n je náhodným výběrem z normálního rozdělení $N(\mu, \sigma^2)$, jsou standardizovaná šikmost a špičatost asymptoticky nezávislé a mají standardní normální rozdělení. Testová statistika $\chi^2 = C_3^2 + C_4^2$ má za platnosti nulové hypotézy asymptoticky χ^2 -rozdělení o 2 stupních volnosti.

Ve vícerozměrném případě lze pro testování použít statistiku

$$\chi_{PE}^2 = \sum_{i=1}^p \left(C_3^2(i) + C_4^2(i) \right),$$

kde $C_3(i)$ je standardizovaná šikmost a $C_4(i)$ standardizovaná špičatost pro i -tou komponentu vektorů $\mathbf{Z}_1, \dots, \mathbf{Z}_n$:

$$C_3(i) = \frac{\sum_{k=1}^n Z_{ki}^3}{n} \sqrt{\frac{n}{6}}, \quad C_4(i) = \left(\frac{\sum_{k=1}^n Z_{ki}^4}{n} - 3 \right) \sqrt{\frac{n}{24}}, \quad i = 1, \dots, p.$$

Jak uvidíme později, má statistika χ_{PE}^2 za platnosti nulové hypotézy asymptoticky χ^2 rozdělení o $2p$ stupních volnosti. Statistika byla původně navržena proti alternativě, že data pocházejí z rozdělení patřícího do určité Pearsonovy třídy. V následujícím textu uvidíme, proti jakým typům alternativ může být test zaměřen, a také se seznámíme s důkazem o asymptotickém chování veličiny χ_{PE}^2 .

4 Koziolův test (jednorozměrný případ)

Abychom lépe pochopili Koziolův test, začněme opět s jednorozměrným případem. Nejprve předpokládejme, že střední hodnotu μ a rozptyl σ^2 pozorovaných (nezávislých, stejně rozdělených) veličin X_1, \dots, X_n známe. Označme $f_Y(y)$ hustotu a $F_Y(y)$ distribuční funkci veličin Y_1, \dots, Y_n , kde $Y_i =$

$(X_i - \mu)/\sigma$, a $\phi(y)$ hustotu, respektive $\Phi(y)$ distribuční funkci, standardního normálního rozdělení. Testujeme nulovou hypotézu H_0 proti alternativě A :

$$\begin{aligned} H_0 : f_Y(y) &= \phi(y), \\ A : f_Y(y) &= \phi(y)(1 + \theta_3 H_3(y) + \theta_4 H_4(y)), \end{aligned} \quad (1)$$

kde θ_3 a θ_4 jsou reálná čísla a funkce $H_3(y)$ a $H_4(y)$ jsou Hermitovy polynomy 3. a 4. stupně. Hermitovy polynomy $\{H_i(y), i = 1, 2, \dots\}$ mají zajímavé vlastnosti vzhledem ke standardnímu normálnímu rozdělení. Nechť Y je veličina řídicí se standardním normálním rozdělením, pak pro $i, j = 1, 2, \dots$

$$\begin{aligned} \mathbb{E} H_i(Y) &= \int_{-\infty}^{\infty} H_i(y) \phi(y) dy = 0, \quad \mathbb{E} (H_i(Y))^2 = \int_{-\infty}^{\infty} (H_i(y))^2 \phi(y) dy = i!, \\ \mathbb{E} [H_i(Y) H_j(Y)] &= \int_{-\infty}^{\infty} H_i(y) H_j(y) \phi(y) dy = 0 \quad \text{pro } i \neq j. \end{aligned}$$

Pro zajímavost uvedme tvar Hermitových polynomů až do 4. stupně:

$$H_0(y) = 1, H_1(y) = y, H_2(y) = y^2 - 1, H_3(y) = y^3 - 3y, H_4(y) = y^4 - 6y^2 + 3.$$

Poznámka. Pro $\theta_3 \neq 0$ nebo $\theta_4 \neq 0$ funkce $\phi(y)(1 + \theta_3 H_3(y) + \theta_4 H_4(y))$ není hustota, protože může nabývat záporných hodnot. Rovnosti (1) je třeba rozumět tak, že platí jen na okolí bodu 0.

Přirozené testové statistiky jsou, vzhledem ke zvolené alternativě a definici Hermitových polynomů, statistiky

$$U_3 = \frac{1}{\sqrt{6n}} \sum_{i=1}^n H_3(Y_i), \quad U_4 = \frac{1}{\sqrt{24n}} \sum_{i=1}^n H_4(Y_i).$$

Jestliže neznáme střední hodnotu μ a rozptyl σ^2 , použijeme dvojici:

$$\hat{U}_3 = \frac{1}{\sqrt{6n}} \sum_{i=1}^n H_3(Z_i), \quad \hat{U}_4 = \frac{1}{\sqrt{24n}} \sum_{i=1}^n H_4(Z_i),$$

kde $Z_i = (X_i - \bar{X})/s$, $i = 1, \dots, n$. Vzhledem k tomu, že $\sum Z_i = 0$ a $(1/n) \sum Z_i^2 = 1$, není \hat{U}_3 a \hat{U}_4 nic jiného než standardizovaná šikmost a standardizovaná špičatost. Statistika $\chi_{KO}^2 = \hat{U}_3^2 + \hat{U}_4^2$ je pak testovou statistikou D'Agostinova testu. Nyní se pojdme podívat, proč má tato statistika asymptoticky χ^2 rozdělení o 2 stupních volnosti.

5 Konvergence empirického procesu

Nechť $G_n(z)$ je empirická distribuční funkce spočtená z transformovaných veličin Z_1, \dots, Z_n . Definujme empirický proces

$$D_n(t) = \sqrt{n} (G_n(\Phi^{-1}(t)) - t), \quad t \in [0, 1].$$

Jedna z nejzákladnějších vět asymptotické statistiky tvrdí (viz [3] a [7]):

$$D_n(t) \xrightarrow{\mathcal{D}(D[0,1])} D(t), \quad (2)$$

kde $\{D(t), t \in [0, 1]\}$ je gaussovský proces s nulovou střední hodnotou a kovarianční funkcí $r_D(t, s)$:

$$r_D(t, s) = \min(t, s) - ts - \phi(\Phi^{-1}(t))\phi(\Phi^{-1}(s)) - \frac{1}{2}\Phi^{-1}(t)\phi(\Phi^{-1}(t))\Phi^{-1}(s)\phi(\Phi^{-1}(s)). \quad (3)$$

Připomeňme, že funkce $r_B(t, s) = \min(t, s) - ts$ je kovarianční funkce Brownova můstku. Pokud bychom znali střední hodnotu μ i rozptyl σ^2 a uvažovali empirickou distribuční funkci $F_n(y)$ standardizovaných veličin $Y_i = (X_i - \mu)/\sigma$, $i = 1, \dots, n$, pak by proces $\{D_n^*(t) = \sqrt{n}(F_n(\Phi^{-1}(t)) - t), t \in [0, 1]\}$, konvergoval právě k Brownovu můstku.

Poznámka. Konvergence v (2) je konvergence v distribuci na prostoru po částech spojitých funkcí definovaných na intervalu $[0, 1]$, které jsou v bodech nespojitosti spojitě zprava a mají limitu zleva, přičemž uvažujeme takzvanou Skorochodovu metriku, viz [1]. Jestliže nějaký náhodný proces $\{X_n(t), t \in [0, 1]\}$ s trajektoriemi v prostoru $D([0, 1])$ konverguje v distribuci k limitnímu procesu $\{X(t), t \in [0, 1]\}$, pak pro libovolný spojitý funkcionál f definovaný na prostoru $D(0, 1)$ platí, že $f(X_n(t))$ konverguje v distribuci na R_1 k $f(X(t))$. Takovým funkcionálem f , který k funkci $g \in D([0, 1])$ přiřazuje reálné číslo $f(g)$ je nejčastěji $\max_{0 \leq t \leq 1} g(t)$ nebo $\int_0^1 h(t) dg(t)$, kde h je spojitá funkce na $[0, 1]$.

Velmi často je kvůli zápisu šikvnější pracovat s empirickým procesem $\{E_n(x), x \in R^1\}$ definovaným:

$$E_n(x) = \sqrt{n}(G_n(x) - \Phi(x)), \quad x \in R^1,$$

a procesem $E(x) = D(\Phi(x))$, $x \in R^1$. Proces $\{E(x), x \in R^1\}$ je opět gaussovský s nulovou střední hodnotou a kovarianční funkcí:

$$K(x, y) = \Phi(\min(x, y)) - \Phi(x)\Phi(y) - \phi(x)\phi(y) - \frac{1}{2}x\phi(x)y\phi(y). \quad (4)$$

Budeme říkat, že $E_n(x) \rightarrow E(x)$ právě tehdy, když $D_n(t) \xrightarrow{\mathcal{D}(D[0,1])} D(t)$.

Zřejmě $\hat{U}_3 = \int \frac{1}{\sqrt{6}}H_3(x) dE_n(x)$ a $\hat{U}_4 = \int \frac{1}{\sqrt{24}}H_4(x) dE_n(x)$ a díky konvergenci (2) platí:

$$\begin{aligned} \hat{U}_3 &= \int \frac{1}{\sqrt{6}}H_3(x) dE_n(x) \xrightarrow{\mathcal{D}} \int \frac{1}{\sqrt{6}}H_3(x) dE(x), \\ \hat{U}_4 &= \int \frac{1}{\sqrt{24}}H_4(x) dE_n(x) \xrightarrow{\mathcal{D}} \int \frac{1}{\sqrt{24}}H_4(x) dE(x). \end{aligned}$$

Vzhledem k tomu, že limitní veličiny jsou integrály vzhledem ke gaussovskému procesu s nulovou střední hodnotou, pak je zřejmé, že vektor

$$\left(\int (1/\sqrt{6}) H_3(x) dE(x), \int (1/\sqrt{24}) H_4(x) dE(x) \right)$$

je normálně rozdělený s nulovou střední hodnotou. Nyní pouze stačí ukázat, že:

$$\mathbb{E} \left(\int (1/\sqrt{6}) H_3(x) dE(x) \right)^2 = 1, \quad \mathbb{E} \left(\int (1/\sqrt{24}) H_4(x) dE(x) \right)^2 = 1, \quad (5)$$

$$\mathbb{E} \left(\int H_3(x) dE(x) \int H_4(x) dE(x) \right) = 0, \quad (6)$$

neboť z (6) pak plyne nezávislost. Zřejmě pro libovolné funkce $f(x)$ a $g(x)$ (pro které integrály existují) platí

$$\mathbb{E} \left(\int f(x) dE(x) \right) \left(\int g(x) dE(x) \right) = \int \int f(x)g(y) dK(x, y)$$

a dále pro $i, j \geq 3$ platí

$$\int \int H_i(x)H_j(y) dK(x, y) = \int H_i(x)^2 \phi(x) dx = i! \quad (7)$$

$$\int \int H_i(x)H_j(y) dK(x, y) = \int H_i(x)H_j(x) \phi(x) dx = 0. \quad (8)$$

Důkaz vztahů (7) a (8) pouze nastíníme. Poměrně snadno je možno ukázat, že integrály vzhledem k diferencovatelné části jádra se rovnají nule, a tedy $\int \int H_i(x)H_j(y) dK(x, y) = \int \int H_i(x)H_j(y) d\Phi(\min(x, y))$. Rovnosti (5) a (6) plynou z toho, že $\int \int f(x)g(y) d\Phi(\min(x, y)) = \int f(x)g(x) \phi(x) dx$.

Obdobný typ konvergence jako (2) lze dokázat i pro více dimenzí. Limitní p -rozměrný proces $\{E(\mathbf{x}), \mathbf{x} \in R^p\}$ je opět gaussovský s nulovou střední hodnotou a kovarianční funkcí

$$K(\mathbf{x}, \mathbf{y}) = \Phi(\mathbf{x} \wedge \mathbf{y}) - \Phi(\mathbf{x})\Phi(\mathbf{y}) - \int_{\mathbf{z} < \mathbf{x}} \mathbf{z}^T d\Phi(\mathbf{z}) \cdot \int_{\mathbf{z} < \mathbf{y}} \mathbf{z} d\Phi(\mathbf{z}) \quad (9)$$

$$- \frac{1}{2} \text{tr} \left(\int_{\mathbf{z} < \mathbf{x}} (\mathbf{z}\mathbf{z}^T - \mathbf{I}) d\Phi(\mathbf{z}) \cdot \int_{\mathbf{z} < \mathbf{y}} (\mathbf{z}\mathbf{z}^T - \mathbf{I}) d\Phi(\mathbf{z}) \right).$$

Proces $\{D(t_1, \dots, t_p), (t_1, \dots, t_p) \in [0, 1]^p\}$ vznikne časovou transformací $D(t_1, \dots, t_p) = E(\Phi^{-1}(t_1), \dots, \Phi^{-1}(t_p))$, a analogicky získáme i jeho kovarianční funkci.

Pro $p = 2$ má kovarianční funkce (9) tvar:

$$K((x_1, x_2), (y_1, y_2)) = \Phi(\min(x_1, y_1))\Phi(\min(x_2, y_2)) - \Phi(x_1)\Phi(x_2)\Phi(y_1)\Phi(y_2)$$

$$- \phi(x_1)\phi(y_1)\Phi(x_2)\Phi(y_2) - \phi(x_2)\phi(y_2)\Phi(x_1)\Phi(y_1)$$

$$- \frac{1}{2}x_1y_1\phi(x_1)\phi(y_1)\Phi(x_2)\Phi(y_2) - \frac{1}{2}x_2y_2\phi(x_2)\phi(y_2)\Phi(x_1)\Phi(y_1)$$

$$- \phi(x_1)\phi(x_2)\phi(y_1)\phi(y_2).$$

6 Koziolův test (vícerozměrný případ)

Hermitovy polynomy více proměnných vznikají jako součiny Hermitových polynomů jedné proměnné. Tak například Hermitovy polynomy 3. stupně mají tvar: xyz , $\frac{1}{\sqrt{2}}(x^2-1)y$, $\frac{1}{\sqrt{6}}(x^3-3x)$. Poznamenejme, že všechny Hermitovy polynomy 3. stupně proměnných x_1, \dots, x_p dostaneme záměnnou proměnných, a tedy všechny Hermitovy polynomy 3. stupně jsou:

$H_3^{\{i,j,k\}}(x_1, \dots, x_p) = x_i x_j x_k$, $i < j < k$, $H_3^{\{i,i,j\}}(x_1, \dots, x_p) = \frac{1}{\sqrt{2}}(x_i^2 - 1)x_j$, $i \neq j$, $H_3^{\{i,i,i\}}(x_1, \dots, x_p) = \frac{1}{\sqrt{6}}(x_i^3 - 3x_i)$; $i, j, k = 1, \dots, p$. Jejich počet je $\binom{p+2}{3}$. Obdobně všechny Hermitovy polynomy 4. stupně mají tvar:

$$xyzuz, \frac{1}{\sqrt{2}}(x^2-1)yz, \frac{1}{2}(x^2-1)(y^2-1), \frac{1}{\sqrt{6}}(x^3-3x)y, \frac{1}{\sqrt{24}}(x^4-6x^2+3)$$

a jejich počet je $\binom{p+3}{4}$.

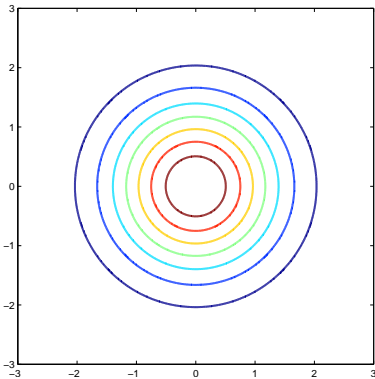
V Koziolově testu se opět testuje nulová hypotéza H_0 proti alternativě A :

$$H_0 : f_{\mathbf{Y}}(y_1, \dots, y_p) = \phi(y_1) \dots \phi(y_p)$$

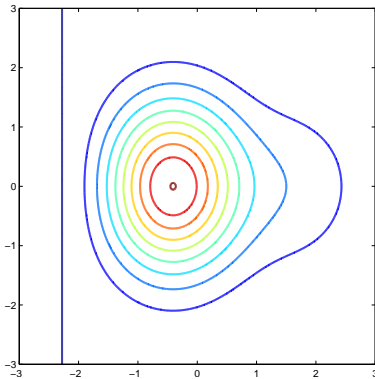
$$A : f_{\mathbf{Y}}(y_1, \dots, y_p) = \phi(y_1) \dots \phi(y_p) (1 + \theta_3^{\{1,1,1\}} H_3^{\{1,1,1\}}(y_1, \dots, y_p) + \dots + \theta_3^{\{1,1,2\}} H_3^{\{1,1,2\}}(y_1, \dots, y_p) + \dots).$$

V závorce $(1 + \theta_3^{\{1,1,1\}} H_3^{\{1,1,1\}}(y_1, \dots, y_p) + \dots)$ se vyskytují všechny Hermitovy polynomy 3. a 4. stupně.

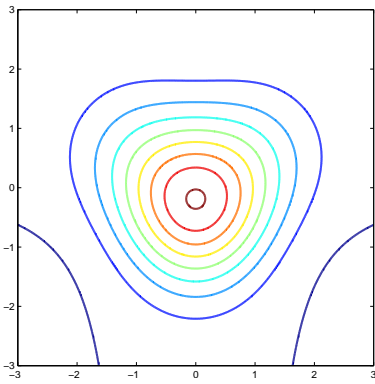
Lepší představu o tom, jak vypadají funkce vyskytující se v alternativě pro případ dvou proměnných, umožňují následující obrázky. Obrázek 1 představuje vrstevnice funkce $f(x, y) = \phi(x)\phi(y)$, obrázek 2 vrstevnice funkce $f(x, y) = \phi(x)\phi(y)(1 + 0.2(x^3 - 3x))$, obrázek 3 vrstevnice funkce $f(x, y) = \phi(x)\phi(y)(1 + 0.2(x^2 - 1)y)$, obrázek 4 vrstevnice funkce $f(x, y) = \phi(x)\phi(y)(1 + 0.2(x^3 - 3x)y)$, obrázek 5 vrstevnice funkce $f(x, y) = \phi(x)\phi(y)(1 + 0.1(x^2 - 1)(y^2 - 1))$, obrázek 6 vrstevnice funkce $f(x, y) = \phi(x)\phi(y)(1 - 0.1(x^2 - 1)(y^2 - 1))$.



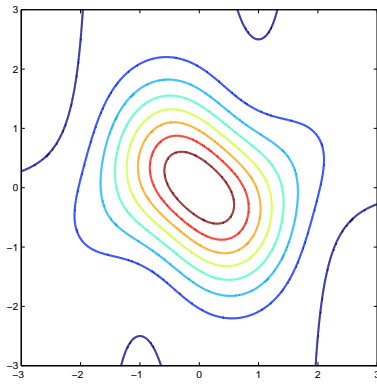
Obrázek 1:



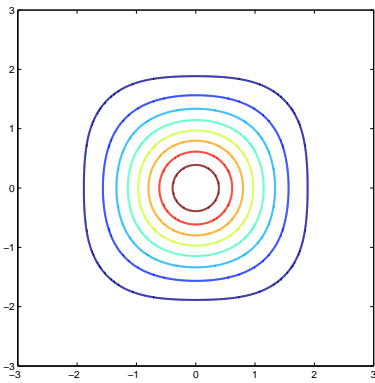
Obrázek 2:



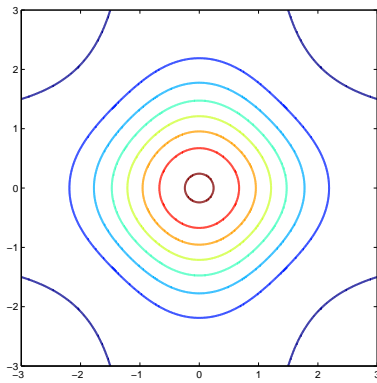
Obrázek 3:



Obrázek 4:



Obrázek 5:



Obrázek 6:

Analogicky jako v jednorozměrném případě můžeme pomocí konvergence k limitnímu vícerozměrnému gaussianovskému procesu ukázat, že testová statistika $\chi_{KO}^2 = \hat{U}_3^2 + \hat{U}_4^2$, kde

$$\begin{aligned} \hat{U}_3^2 &= \sum_{j < k < l} \frac{1}{n} \left(\sum_i Z_{ij} Z_{ik} Z_{il} \right)^2 + \sum_{j \neq k} \frac{1}{2n} \left(\sum_i Z_{ij}^2 Z_{ik} \right)^2 + \sum_j \frac{1}{6n} \left(\sum_i Z_{ij}^3 \right)^2, \\ \hat{U}_4^2 &= \sum_{j < k < l < m} \frac{1}{n} \left(\sum_i Z_{ij} Z_{ik} Z_{il} Z_{im} \right)^2 + \sum_{j \neq k, l; k < l} \frac{1}{2n} \left(\sum_i Z_{ij}^2 Z_{ik} Z_{il} \right)^2 \\ &\quad + \sum_{j < k} \frac{1}{4n} \left(\sum_i (Z_{ij}^2 Z_{ik}^2 - 1) \right)^2 + \sum_{j \neq k} \frac{1}{6n} \left(\sum_i Z_{ij}^3 Z_{ik} \right)^2 \\ &\quad + \sum_j \frac{1}{24n} \left(\sum_i (Z_{ij}^4 - 3) \right)^2, \end{aligned}$$

má za platnosti nulové hypotézy asymptoticky χ^2 rozdělení o $\binom{p+2}{3} + \binom{p+3}{4}$

stupni volnosti. Každý sčítanec v součtu $\widehat{U}_3^2 + \widehat{U}_4^2$ má χ^2 rozdělení o 1 stupni volnosti. Jestliže zamítneme nulovou hypotézu, můžeme z rozboru velikosti jednotlivých členů poznat, jakým způsobem byla normalita porušena.

Poznámka. Pochopitelně můžeme uvažovat i alternativy, které neobsahují všechny Hermitovy polynomy 3. a 4. stupně. Budeme-li např. uvažovat jen Hermitovy polynomy $H_3^{\{j,j,j\}}(y_1, \dots, y_p)$, $H_4^{\{j,j,j,j\}}(y_1, \dots, y_p)$, $j = 1, \dots, p$, bude testová statistika shodná se statistikou χ_{PE}^2 Pearsonova testu, přičemž její asymptotické rozdělení bude χ^2 o $2p$ stupních volnosti.

7 Kolmogorovův-Smirnovův test (jednorozměrný případ)

Nejobecnější test dobré shody pro normální rozdělení má samozřejmě tvar:

$$\begin{aligned} H_0 : F_Y(y) &= \Phi(y), \\ A : F_Y(y) &\neq \Phi(y). \end{aligned}$$

Testová statistika KS Kolmogorovova-Smirnovova testu má tvar:

$$KS = \max_{t \in [0,1]} |D_n(t)| = \max_{x \in R^1} |E_n(x)|. \quad (10)$$

Z konvergence (2) vyplývá, že za platnosti H_0

$$\max_t |D_n(t)| \xrightarrow{\mathcal{D}} \max_t |D(t)|,$$

kde $\{D(t)\}$ je gaussovský proces s nulovou střední hodnotou a kovarianční funkcí (3). Rozdělení veličiny $\max_t |D(t)|$ je však velmi složité. Vyjdeme-li z předpokladu, že nás zajímají kritické hodnoty testu pro malé hodnoty hladiny významnosti α , pak je možno získat jejich přiblížení z teorie velkých odchylek, viz [4]:

$$P\left(\max_{t \in [0,1]} |D(t)| > u\right) = 2\sqrt{\frac{2\pi}{\pi-2}} \exp\left(-\frac{2\pi}{\pi-2} u^2\right) (1 + o(1)). \quad (11)$$

Tabulka 1 porovnává 5% a 1% kvantily statistiky KS získané z aproximace (11), pomocí simulací (50 000 vygenerovaných výběrů o rozsahu $n=300$ a 10 000 vygenerovaných výběrů o rozsahu $n=500$) a kvantily maxima absolutní hodnoty Brownova můstku $\max_t |B(t)|$. Na přání Járy Antocha zde znovu zdůrazňeme, že Kolmogorovova-Smirnovova testová statistika konverguje k $\max_t |B(t)|$ pouze pokud jsou známy parametry normálního rozdělení a empirická distribuční funkce je počítána z hodnot $\{Y_i\}$. Zároveň připomeňme, že

$$P\left(\max_{t \in [0,1]} |B(t)| > u\right) = 2 \sum_{k=1}^{\infty} (-1)^{k-1} e^{-2k^2 u^2}.$$

Tabulka 1 ukazuje jasně, že použití kvantilů maxima absolutní hodnoty Browanova můstku je nevhodné. Dříve běžné chybě se dnes vyhýbají i tvůrci statistických softwarů. Ve statistickém toolboxu Matlabu se v takzvaném Lillieforsově testu, v kterém se testuje normalita pomocí Komogorovovy-Smirnovovy statistiky, používá kritická hodnota 1.103 pro hladinu významnosti $\alpha = 0.01$ a hodnota 0.886 pro hladinu $\alpha = 0.05$. Domníváme se, že tyto hodnoty byly získané simulací.

	aprox. (11)	sim. ($n = 300$)	sim. ($n = 500$)	$\max B(t) $
5% kvantil	0.909	0.899	0.905	1.35
1% kvantil	1.057	1.052	1.054	1.64

Tabulka 1: 5% a 1% kvantily Kolmogorovovy-Smirnovovy statistiky KS získané pomocí (11) a pomocí simulací ve srovnání s kvantily maxima absolutní hodnoty Brownova můstku.

8 Kolmogorovův-Smirnovův test (vícerozměrný případ)

Testujeme opět nulovou hypotézu H_0 proti alternativě A :

$$H_0 : F_Y(\mathbf{y}) = \Phi(\mathbf{y}),$$

$$A : F_Y(\mathbf{y}) \neq \Phi(\mathbf{y}),$$

kde $F_Y(\mathbf{y})$ je distribuční funkce vektorů $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ a $\Phi(\mathbf{y}) = \Phi(y_1) \cdots \Phi(y_n)$.

Testová statistika má tvar:

$$KS = \max_{\mathbf{z} \in R^p} \sqrt{n} |G_n(\mathbf{z}) - \Phi(\mathbf{z})|, \quad (12)$$

kde $G_n(\mathbf{z})$ je p -rozměrná empirická distribuční funkce veličin

$$\mathbf{Z}_1 = \mathbf{S}^{-1/2}(\mathbf{X}_1 - \bar{\mathbf{X}}), \dots, \mathbf{Z}_n = \mathbf{S}^{-1/2}(\mathbf{X}_n - \bar{\mathbf{X}}).$$

Platí $KS \xrightarrow{\mathcal{D}} \max_{(t_1, \dots, t_p) \in [0,1]^p} |D(t_1, \dots, t_p)|$.

Poznámka. Výpočet empirické distribuční funkce více proměnných, a tedy i Kolmogorovovy-Smirnovovy statistiky, je velmi časově a paměťově náročný. Například pro $p = 2$, kde data tvoří množinu $M = \{(u_1, v_1), \dots, (u_n, v_n)\}$, je třeba pro všechny dvojice (u_i, v_j) , $i, j = 1, \dots, n$, spočítat počet dvojic z množiny M takových, že první souřadnice je menší než u_i a druhá menší než v_j a dále počet dvojic z M , kde první souřadnice je menší rovna u_i a druhá menší rovna v_j . Je-li např. $n = 100$, je třeba spočítat $2 \cdot 100^2$ hodnot.

Pro výpočet kritických hodnot statistiky (12) lze použít aproximaci, která platí pro velké hodnoty n , navrženou opět Fatalovem (1993). Pro $p = 2$ má tato aproximace tvar:

$$P\left(\sup_{(t_1, t_2) \in (0,1) \times (0,1)} |D(t_1, t_2)| > u\right) = 309.4 u \exp(-4.939u^2) (1 + o(1)). \quad (13)$$

Tabulka 2 porovnává 5% a 1% kvantily statistiky KS získané z aproximace (13), pomocí simulací (10 000 vygenerovaných výběrů o rozsahu $n=200$) s kvantily maxima absolutní hodnoty Brownova můstku $\max_{t_1, t_2} |B(t_1, t_2)|$, které byly také získány metodou velkých odchylek.

	aprox. (13)	sim. ($n = 200$)	$\max B(t_1, t_2) $
5% kvantil	1.352	1.292	1.695
1% kvantil	1.474	1.544	1.957

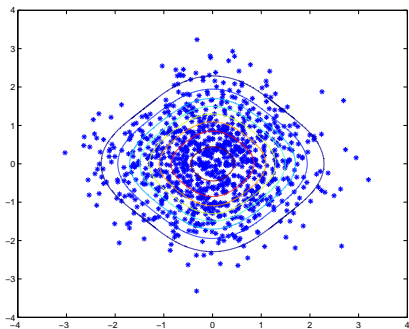
Tabulka 2: 5% a 1% kvantily Kolmogorovovy-Smirnovovy statistiky KS získané aproximací (13) a pomocí simulací ve srovnání s kvantily $\max_{t_1, t_2} |B(t_1, t_2)|$.

9 Příklad

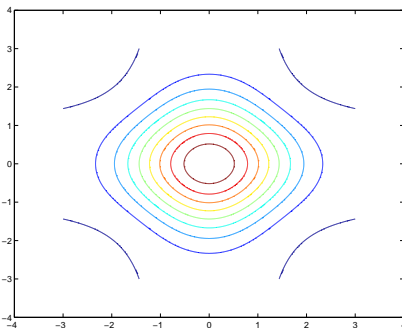
Obrázek 7 představuje rozptylový graf odpovídající 1000 bodům $\{(y_{i,1}, y_{i,2})\}$ vygenerovaným z rozdělení s hustotou

$$f(y_1, y_2) = \frac{1}{2\pi} \left(\left(\sqrt{2}e^{-y_1^2/2} - e^{-y_1^2} \right) e^{-y_2^2} + \left(\sqrt{2}e^{-y_2^2/2} - e^{-y_2^2} \right) e^{-y_1^2} \right) \quad (14)$$

spolu s jejími vrstevnicemi. Rozdělení není dvourozměrné normální, ale má nekorelované standardní normální marginály, viz [6, strana 211]. Obrázek 8 představuje vrstevnice funkce $\phi(y_1)\phi(y_2)(1 - 0.117(y_1^2 - 1)(y_2^2 - 1))$. (Srovnej s Tabulkou 3.)



Obrázek 7:



Obrázek 8:

Tabulka 3 udává velikosti jednotlivých členů v Koziolově statistice $\chi_{KO}^2 = \widehat{U}_3 + \widehat{U}_4$. Na hladině $\alpha = 0.05$ se nulová hypotéza H_0 zamítá ($\chi_{KO}^2 = 17.25$) díky nenulovosti parametru $\theta_4^{\{1,1,2,2\}}$ u Hermitova polynomu $H_4^{\{1,1,2,2\}}$. Pearsonova statistika $\chi_{PE}^2 = 1.73$ nulovou hypotézu nezamítá. Kolmogorovova-Smirnovova statistika, která nabývá hodnoty $KS = 0.745$, také H_0 nezamítá.

	stat.	st.vol.	p-hodnota
$\frac{1}{2n} (\sum_i Z_{i1}^2 Z_{i2})^2 + \frac{1}{2n} (\sum_i Z_{i2}^2 Z_{i1})^2$	1.37	2	0.50
$\frac{1}{6n} (\sum_i Z_{i1}^3)^2 + \frac{1}{6n} (\sum_i Z_{i2}^3)^2$	0.93	2	0.63
$\frac{1}{4n} (\sum_i (Z_{i1}^2 Z_{i2}^2 - 1))^2$	13.68	1	$2.16 \cdot 10^{-4}$
$\frac{1}{6n} (\sum_i Z_{i1}^3 Z_{i2})^2 + \frac{1}{6n} (\sum_i Z_{i2}^3 Z_{i1})^2$	0.47	2	0.79
$\frac{1}{24n} (\sum_i (Z_{i1}^4 - 3))^2 + \frac{1}{24n} (\sum_i (Z_{i2}^4 - 3))^2$	0.80	2	0.33
\widehat{U}_3	2.30	4	0.68
\widehat{U}_4	14.95	5	0.01
χ_{KO}^2	17.25	9	0.04

Tabulka 3: Rozbor jednotlivých členů v Koziolově statistice.

Literatura

- [1] Billingsley P. (1968). *Convergence of Probability Measures*. John Wiley & Sons, New York - London - Sydney - Toronto.
- [2] Thode H.C. (2002). *Testing Normality*. Marcel Dekker, New York - Basel.
- [3] Durbin J. (1973). *Weak convergence of the sample distribution function when parameters are estimated*. Annals of Stats. **1**, 279–290.
- [4] Fatalov V.R. (1993). *Asymptotics of large deviation probabilities for Gaussian fields: Applications*. Izvetiya Natsionalnoi Akademii Nauk Armenii. Matematika **28**, 21–44.
- [5] Koziol J.A. (1987). *An alternative formulation of Neyman's smooth goodness of fit tests under a composite alternatives*. Metrika **34**, 17–24.
- [6] Rényi A. (1972). *Teorie pravděpodobnosti* (český překlad). Academia Praha.
- [7] Tjurin Ju. N. (1973). *Proverka gipotezy o normalnosti vyborky bolshovo objema*. Teor. veroj. i ee primen. **18**, 651–655.

Poděkování: Tato práce byla podporována grantem MSM684077002.

Adresa: D. Jarušková, FSv ČVUT, kat. mat., Thákurova 7, 166 29 Praha 6

E-mail: jarus@mat.fsv.cvut.cz

LÈVY PROCESSES AND BALAYAGE SPACE

Andrea Karlová

Keywords: Lévy processes, Hunt processes, balayage space, harmonic kernels, hyperharmonic functions.

Abstract: Bliedtner, Hansen [2] state the theorem which gives the characterization for the cones of excessive functions for sub-Markov processes, where all excessive functions are lower semicontinuous and positive hyperharmonic functions for a family of harmonic kernels on X . We formulate the consequence of this rather general theorem for Lévy processes.

Abstrakt: Bliedtner, Hansen [2] uvádějí obecný theorem, který umožňuje charakterizovat kužel excesivních funkcí pro sub-markovské procesy tak, že všechny tyto excesivní funkce jsou zdola polospojité a kladné hyperharmonické funkce pro rodinu harmonických jader na lokálně kompaktním prostoru X . V tomto článku vysvětlíme výše uvedené pojmy, vyslovíme zmíněnou větu a formulujeme důsledek pro Lévyho procesy.

1 Introduction

The classical potential theory provides a very close connection between theory of probability and mathematical calculus. A very known is a powerfull Feymann-Kac theorem which very formally speaking allows us to interpret the solution of the partial differential equation of a specific form as an expected value of the functional of Brownian motion stopped at a certain time. The link between these two branches of mathematics is however much wider. We can look at potential theory from four different angles via Hunt processes, sub-Markov semigroups, families of harmonic kernels and balayage space. This is stated in Theorem IV.8.1 in Bliedtner, Hansen [2]).

Our motivation comes from the interest in the family of functions which solves generalized Dirichlet problem. Consider for example a heat equation. The functions which are harmonic with respect to the heat equation are a special example of the general class of harmonic kernels. The harmonic functions are also closely related with the class of Brownian semigroups. Our intention is to study Lévy processes. The later covers large class of stochastic processes which however still owns very nice and handable properties.

The article is organized as follows. We firstly introduce notation and give definitions of balayage space, sub-Markov semigroups, excessive functions, hyperharmonic functions, family of harmonic kernels and Hunt processes in the section Preliminaries. The reader familiar with these terms may prefer to skip this part. In the second section we state a rather general theorem from Bliedtner J., Hansen W. [2]. In the last section we recall the definition of Lévy

process and use the theorem from the second part to formulate a consequence of it for Lèvy processes as a special case of Hunt processes.

The introduced theory is of great importance not only for the theoretical studies. It has great applicability to optimal stopping problems and stochastic control problems and consequently is of interest e.g. in the pricing of the American contingent claims. For an introduction to the latter topics and Dirichlet problem the reader may appreciate chapters 9-12 from Oksendal [3].

2 Preliminaries

We consider a locally compact space X with a countable base and denote \mathcal{U} a base for the topology on X consisting of relatively compact open sets in X . We further denote by $\mathcal{C}(X)$ space of all continuous real functions on X and $\mathcal{C}_0(X)$ space of all continuous real functions vanishing in infinity. We also fix the probability space (Ω, \mathcal{F}, P) on which live considered stochastic processes.

In the following we explain what is a balayage space. Let us consider \mathcal{W} a convex cone of positive lower semicontinuous¹ numerical functions on X . For $v \in \mathcal{W}$ we denote \hat{v} the lower regularization of v , i.e. $\hat{v}(x) := \liminf_{y \rightarrow x} v(y)$, $x \in X$.

The coarsest² topology which is at least as fine as the initial topology and for which all the functions from \mathcal{W} are continuous will be called (\mathcal{W}) -fine topology.

Observe that functions from a convex cone \mathcal{W} are missing the continuity property in the initial topology \mathcal{U} of locally relatively compact open sets and thus we take new topology such that we add some sets to topology \mathcal{U} in which the functions from \mathcal{W} will be continuous.

By \hat{v}^f we will denote the lower regularization of the function v in the \mathcal{W} -fine topology.

Definition 1. (X, \mathcal{W}) is a balayage space, if the following properties are satisfied:

1. \mathcal{W} is σ -stable, i.e. for every increasing sequence (v_n) of functions from \mathcal{W} the supremum $\sup v_n$ is in \mathcal{W}
2. $\widehat{\inf \mathcal{V}}^f \in \mathcal{W}$ for every subset \mathcal{V} of \mathcal{W} .
3. if $u, v', v'' \in \mathcal{W}$ such that $u \leq v' + v''$, then there exists $u', u'' \in \mathcal{W}$ such that $u = u' + u''$ and $u' \leq v', u'' \leq v''$.
4. \mathcal{W} is linearly separating, i.e. $\forall x, y \in X, x \neq y$ and $\lambda \geq 0$ exists $v \in \mathcal{W}$, s.t. $v(x) \neq \lambda v(y)$

¹Real function f defined on the topological space is lower semicontinuous if for every real α is the set $\{x; f(x) > \alpha\}$ open.

²Consider two topologies Θ_1, Θ_2 on X , s.t. $\Theta_1 \subseteq \Theta_2$ then we say that Θ_2 is finer than Θ_1 and Θ_1 is coarser than Θ_2 .

and there exist positive $u_0, v_0 \in \mathcal{W} \cap \mathcal{C}(X)$ s.t. $\frac{u_0}{v_0} \in \mathcal{C}_0(X)$
and $v = \sup\{u \in \mathcal{W} \cap \mathcal{C}(X) : u \leq v\}$ for every $v \in \mathcal{W}$

As an example consider X a discrete space (at most countable) and \mathcal{W} the set of all positive numerical functions on X , then (X, \mathcal{W}) is a balayage space.

Just for completeness recall what we understand by the sub-Markov (resp. Markov) semigroup. We also explain what are excessive functions.

Definition 2. A family $\mathbb{P} = (P_t)_{t>0}$ of kernels³ on X is a semigroup, if $P_{s+t} = P_s P_t$ for $s, t > 0$.

It is a sub-Markov (resp. Markov), if for every $t > 0$ $P_t \mathbb{I} \leq \mathbb{I}$ (resp. $P_t \mathbb{I} = \mathbb{I}$).

Further note, that having define a semigroup of Markov kernels we will call a positive borel measurable function $f \in \mathcal{B}^+(X)$ excessive with respect to the Markov semigroup \mathbb{P} (\mathbb{P} -excessive) if $\sup_{t>0} P_t f = f$. The set of all \mathbb{P} -excessive functions we denote by $E_{\mathbb{P}} := \{f \in \mathcal{B}^+(X) : \sup_{t>0} P_t f = f\}$.

Now, we describe what are families of harmonic kernels and give definitions of hyperharmonic and superharmonic functions.

Consider a family of kernels $(H_U)_{U \in \mathcal{U}}$ on X . We call $(H_U)_{U \in \mathcal{U}}$ a family of sweeping kernels (relative to U) if for every set U from the base \mathcal{U} of X , we have a kernel H_U on X for which holds that $H_U(x, U) = 0$ for every $x \in U$ and for every $x \in U^C$: $H_U(x, \cdot) = \delta_x$ where δ_x denotes dirac mass.

Now for every open subset V of X we denote $\mathcal{U}(V)$ the set of all open subsets $W \subseteq V$ such that closure \bar{W} is a compact set in V . Let $\mathcal{H}^*(V)$ denote the set of all positive hyperharmonic functions on V , that is the set of all positive borel measurable functions v on X , such that v is lower semicontinuous on V and $-\infty < H_U v \leq v$ for all $U \subset \mathcal{U}(V)$. We also denote by $\mathcal{S}^*(V)$ the set of all positive superharmonic functions on V , i.e. the set of all positive hyperharmonic functions w on V for which $H_U w|_U \in \mathcal{C}(U)$ for all $U \subset \mathcal{U}(V)$.

We call $(H_U)_{U \in \mathcal{U}}$ a family of harmonic kernels if the following axioms are fulfilled:

1. $\forall x \in X, \lim_{U \downarrow x} H_U \mathbb{I}(x) = \mathbb{I}$
2. $H_V H_U = H_U$ for all $U, V \in \mathcal{U}$ and $\bar{V} \subset U$
3. For all $U \in \mathcal{U}$ and $f \in \mathcal{B}(X)$ which are bounded on ∂U the function $H_U f$ is continuous on U .
4. for $U \in \mathcal{U}$ and every $x \in U$ there exists a hyperharmonic function w s.t. $w(x) < \infty$ and $\lim_{\mathcal{A}} = \infty$ for every non-regular ultrafilter⁴ \mathcal{A} on U

³A kernel K on X is a mapping $K : X \times \mathcal{B}(X) \rightarrow \mathbb{R}^+$ s.t. $x \rightarrow K(x, B)$ is Borel measurable for every $B \in \mathcal{B}(X)$ and $B \rightarrow K(x, B)$ is a measure on $(X, \mathcal{B}(X))$ for every $x \in X$

⁴For non-empty set U a ultrafilter \mathcal{A} is non-empty system of subsets of U s.t. (i): $0 \notin \mathcal{A}$, (ii): $A_1, A_2 \in \mathcal{A}$ then $A_1 \cap A_2 \in \mathcal{A}$, (iii): $A_1 \in \mathcal{A}$ and $A_1 \subseteq A_2$ then $A_2 \in \mathcal{A}$, and finally

5. the space of positive hyperharmonic functions $\mathcal{H}^*(X)$ is linearly separating and there exists a strictly positive function $s_0 \in \mathcal{S}^*(X) \cap \mathcal{C}(X)$.

For better understanding of the definition reader can verify (or read e.g. in Bliedtner, Hansen [2]) that heat kernel in the classical Dirichlet problem is a harmonic kernel.

Finally we need to define Hunt process. As we will see from the definition it is a rather general stochastic process. However it still posses a Markov property.

Definition 3. A markov process $Z = (Z(t), t \geq 0)$ is called a Hunt process if the following conditions are satisfied:

1. the trajectories of $Z(t)$ are right-continuous on $t \geq 0$ and have limits from left for $0 \leq t \leq T$ for $T \in \mathbb{R}^+$
2. $(Z(t), t \geq 0)$ has the strong markov property w.r.t to the augmented natural filtration $(\mathcal{G}_t^Z, t \geq 0)$, i.e. given any \mathcal{G}_t^Z -adapted stopping time τ and $\forall s \geq 0, B \in \mathcal{B}(\mathbb{R}^d)$ the following holds:

$$P(Z(\tau + s) \in B | \mathcal{G}_t^Z) = P(Z(s) \in B | Z(\tau))$$

3. $(Z(t), t \geq 0)$ is quasi-left continuous, i.e. whenever (τ_n) is sequence of \mathcal{G}_t^Z -adapted stopping times and τ is \mathcal{G}_t^Z -adapted stopping time, s.t. $\tau_n \nearrow \tau$ then

$$\lim_{n \rightarrow \infty} Z(\tau_n) = Z(\tau) \text{ a.s. on } [\tau < \infty]$$

We will further denote by E_Z the set of all functions excessive with respect to Markov semigroup \mathbb{P} of Hunt process Z .

3 Theorem

The following theorem provides us with a general result which allows us to describe the cone \mathcal{W} of excessive functions for sub-Markov processes, where all excessive functions are lower semicontinuous and positive hyperharmonic functions for a family of harmonic kernels on X . Further (X, \mathcal{W}) is a balayge space.

Theorem 1. Let $\mathcal{P} \in \mathcal{C}^+(X)$ be a function cone and $\mathcal{W} = S(\mathcal{P}) := \{\sup f_n : (f_n) \in \mathcal{P} \text{ increasing}\}$ s.t. $\mathbb{I} \in \mathcal{W}$. Then the following statements are equivalent:

(iv): $A_1 \in \mathcal{A}$ or $A_1^c \in \mathcal{A}$. Considering an ultrafilter \mathcal{A} on U converging to point z from interior of U , then we say that ultrafilter \mathcal{A} is regular w.r.t to the family of kernels H_U if $\lim_{x, \mathcal{A}} H_U(z, \cdot) = \delta_z$ (δ_z is dirac mass at z). Every ultrafilter which is not regular is called non-regular.

1. (X, \mathcal{W}) is a balayage space
2. there exists a family $(H_U)_{U \in \mathcal{U}}$ of harmonic kernels on X such that $\mathcal{H}^*(X) = \mathcal{W}$.
3. \mathcal{W} is min-stable, i.e for every $f, g \in \mathcal{W}$ the infimum $\inf(f, g) \in \mathcal{W}$ and there exists a sub-Markov semigroup $\mathbb{P} = (P_t)_{t>0}$ on X s.t. $E_{\mathbb{P}} = \mathcal{W}$
4. there exists a Hunt process $Z = (Z(t), t \geq 0)$ with state space (X, \mathcal{X}) s.t. $E_Z = \mathcal{W}$.

Proof: see Blidtner J., Hansen W. (1986) [2], Theorem IV.8.1, p.168.

4 Application to Lèvy processes

Lèvy processes are an example of Markov processes with more restrictive conditions. They are also closely related to convolution semigroups of measures such that the distribution of the Lèvy process forms a convolution semigroup. To have nice insight about their behaving, we can also describe them as processes which can in general consist from the process with finite variation, a Brownian motion part and jump process which has countable many jumps. Each of these component can be missing so for example Brownian motion or Poisson process or mixture of these is Lèvy process.

Definition 4. A stochastic process $Z = (Z(t), t \geq 0)$ on X is a *Lèvy process* if the following conditions holds:

1. $Z(0) = 0$ a.s.
2. $(Z(t), t \geq 0)$ has independent and stationary increments
3. (stochastic continuity)
 $\forall \varepsilon > 0$ and all $h > 0$ $\lim_{h \rightarrow 0} \mathbb{P}(|Z(t+h) - Z(h)| > \varepsilon) = 0$
4. (cadlag property of trajectories)
 $(Z(t), t \geq 0)$ is right continuous in $t \geq 0$ and has left limits in $t > 0$.

The following proposition which we can formulate due to Theorem 1 gives us properties for the set of all excessive functions with respect to considered Lèvy process such that the set is closed under infimum and supremum and can be characterized via hyperharmonic functions.

Proposition 1. For a function cone $\mathcal{P} \in \mathcal{C}^+(X)$ and $\mathcal{W} = S(\mathcal{P})$ s.t. $\mathbb{1} \in \mathcal{W}$ and Lèvy process $Z = (Z(t), t \geq 0)$ with the state space (X, \mathcal{X}) , s.t. $E_Z = \mathcal{W}$ there exists a family $(H_U)_{U \in \mathcal{U}}$ of harmonic kernels on X , s.t. $\mathcal{H}^*(X) = \mathcal{W}$, \mathcal{W} is min-stable and σ -stable and there exists a sub-Markov semigroup $\mathbb{P} = (P_t)_{t>0}$ on X , s.t. $E_{\mathbb{P}} = \mathcal{W}$

Proof: The fact that Lèvy process is Hunt process follows from definitions of these processes. Note, that Lèvy process is a Markov process (see e.g. Sato K. [4]) and every Lèvy process is a Feller process and every Feller process is a Hunt process (Applebaum D. [1]). The rest is a consequence of Theorem 1.

References

- [1] Applebaum D. (2004). *Lévy processes and stochastic calculus*. Cambridge University Press.
- [2] Bliedtner J., Hansen W. (1986). *Potential theory - an analytic and probabilistic approach to balayage*. Springer-Verlang, Berlin.
- [3] Oksendal B.K. (2000). *Stochastic differential equations: an introduction with applications, 5th Edt.*. Springer-Verlang, Berlin Heidelberg New York
- [4] Sato K. (1999). *Lévy processes and infinitely divisible distributions*. Cambridge University Press.

Acknowledgement: I would like to thank my colleagues at KMPS for creating a great working environment, to my supervisor doc. Petr Volf for his generous support and to Michaela Prokešová for a fruitfull debate on the form of this text.

Address: A. Karlová, MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8 – Karlín;
ČSOB, Value and Risk Management, Radlická 333, 150 57 Praha 5 – Radlice

E-mail: karlova@karlin.mff.cuni.cz, akarlova@csob.cz

INTERVALOVÝ ODHAD PARAMETRU p BINOMICKÉHO ROZDĚLENÍ: CO JE (RELATIVNĚ) NOVÉHO?

Jan Klaschka

Klíčová slova: Binomické rozdělení, znáhodněné konfidenční intervaly, fuzzy konfidenční intervaly

Abstrakt: Práce navazuje na článek z ROBUSTu 2006, kde byly vyloženy různé typy konfidenčních intervalů pro parametr p binomického rozdělení. Pravděpodobnost pokrytí skutečného p konfidenčním intervalem nemůže být na celém intervalu $[0, 1]$ rovna předepsané nominální hladině spolehlivosti $1 - \alpha$. Odtud vyplývají dilemata (musí být pravděpodobnost pokrytí vždy alespoň $1 - \alpha$, nebo je lepší aproximovat $1 - \alpha$ z obou stran?) a množství návrhů, jak konfidenční interval konstruovat. Nyní bude výklad rozšířen o zobecnění „obyčejných“ konfidenčních intervalů, která dosažení konstantní pravděpodobnosti pokrytí umožňují: znáhodněné konfidenční intervaly (známé už několik desetiletí) a relativní novinku – tzv. fuzzy konfidenční intervaly.

Abstract: The present work follows up the ROBUST 2006 paper where various types of confidence intervals for binomial parameter p have been exposed. The coverage probability cannot equal the nominal confidence level $1 - \alpha$ in the whole domain $[0, 1]$. This leads to dilemmas (is the coverage of at least $1 - \alpha$ a must, or is it better to approximate $1 - \alpha$ from both sides?), and to multiplicity of proposals of confidence interval types. The present work extends the scope of the previous paper by such generalizations of “ordinary” confidence intervals that enable a constant coverage, namely by the randomized confidence intervals (introduced several decades ago), and by the relatively new idea of the fuzzy confidence intervals.

1 Úvod

S intervalovým odhadem pravděpodobnosti jevu na základě počtu jeho výskytů v určitém počtu nezávislých pokusů se v nejrůznějších aplikacích setkáváme velmi často. Aktuálně jej potřebuji při studiu možností detekce stavů hrozících mikrosnánkem ze záznamu EEG, ale stejně tak se bez něj neobejdu jako statistický konzultant v lékařském výzkumu a narazíme na něj v mnoha situacích v biologii, ekonomii, technice a v řadě dalších disciplin.

Dalo by se očekávat, že problém tak frekventovaný a jednoduchý bude dávno mít uspokojivé řešení. Dávno ano – Laplace publikoval vzorec, který se dodnes užívá (tolik, že se mu říká *standardní*), v r. 1812 [9]. S uspokojivostí už je to horší: Interval kdysi navržený Laplacem má dnes značně pochroumanou reputaci a existuje řada novějších návrhů typů konfidenčních intervalů,

z nichž ale žádný „nevítězí“ tak jednoznačně, aby se nenašly důvody pro další vylepšení. Z údivu nad tím, že se tak starou úlohou zabývají dobří autoři ještě po r. 2000, vznikla má přednáška na ROBUSTu 2006 a článek [8]. Jestliže se dnes k tématu znovu vracím, je to proto, že jsem krátce po ROBUSTu 2006 zjistil, že kolem konfidenčních intervalů pro parametr binomického rozdělení je živo ještě více, než jsem tušil.

Tou v nadpisu avizovanou relativně novou myšlenkou jsou tzv. *fuzzy konfidenční intervaly*. Než se k nim dostaneme, shrnu v sekci 2 některá fakta o „obyčejných“ konfidenčních intervalech a v sekci 3 se zmíním o znárodných konfidenčních intervalech. Fuzzy konfidenčním intervalům pak bude věnována sekce 4.

2 Malé opakování

V dalším budeme X značit počet „úspěchů“ (výskytů nějakého jevu) v n nezávislých pokusech, když pravděpodobnost „úspěchu“ v jednotlivém pokusu je $p \in [0, 1]$. X tedy má binomické rozdělení s parametry n a p , tj.

$$P_p(X = i) = \binom{n}{i} p^i (1-p)^{n-i}, \quad i = 0, 1, \dots, n.$$

Budeme se zabývat oboustrannými konfidenčními intervaly $CI = [p_D, p_H]$ pro p na hladině spolehlivosti $1 - \alpha$ při pevném známém n .

V elementárních statistických učebních textech stejně jako v aplikacích se nejčastěji setkáme s tzv. *standardním* konfidenčním intervalem

$$CI^{st} = \hat{p} \pm z_{\alpha/2} \sqrt{\hat{p}(1-\hat{p})/n}, \quad (1)$$

kde $\hat{p} = X/n$ a $z_{\alpha/2}$ je $(1 - \alpha/2)$ -kvantil $N(0, 1)$. Tento interval publikoval Laplace již v r. 1812.

Vzorec (1) je založen na aproximaci binomického rozdělení normálním rozdělením. Pro malé n nebo p blízké krajním hodnotám 0 a 1 aproximace není kvalitní, a proto se užívání standardního intervalu doporučuje jen tehdy, jsou-li splněny určité podmínky, např. $\min(n\hat{p}, n(1-\hat{p})) \geq 5$. Pokud splněny nejsou, saháme v praxi většinou k „exaktnímu“ Clopper-Pearsonovu intervalu (publikovanému v r. 1934 [6]). Jeho meze p_D a p_H jsou řešením rovnic

$$\sum_{i=X}^n \binom{n}{i} (p_D)^i (1-p_D)^{n-i} = \alpha/2, \quad \sum_{i=0}^X \binom{n}{i} (p_H)^i (1-p_H)^{n-i} = \alpha/2 \quad (2)$$

s výjimkami $p_D = 0$ pro $X = 0$ a $p_H = 1$ pro $X = n$.

Standardní a Clopper-Pearsonův konfidenční interval se mnohdy chápou jako dvě řešení jedné a té samé úlohy – první jednodušší, přibližné a platné jen v určitých mezích, druhé složitější, ale přesné a univerzální. To však není tak úplně pravda. Každý z daných dvou typů konfidenčního intervalu můžeme brát jako reprezentanta jiného přístupu, tedy vlastně jako řešení jiné úlohy.

Exaktní přístup požaduje, aby pro funkci pokrytí $cover(p) = P_p(p \in CI)$ v celém definičním oboru platilo $cover(p) \geq 1 - \alpha$. (Např. C. Mehta v diskusi k článku [5] argumentuje, že zaručit něco na 95 % znamená alespoň na 95 %, ne na přibližně 95 %.) Clopper-Pearsonův konfidenční interval dané kritérium evidentně splňuje, na rozdíl od standardního intervalu.

Příznivci *přibližných konfidenčních intervalů* (výmluvný je název Agrestiho a Coullova článku [1]: *Approximate is better than “exact” for interval estimation of binomial proportions*) považují naproti tomu za dostatečné, když funkce $cover(p)$ nominální hladinu spolehlivosti $1 - \alpha$ dobře aproximuje z kterékoli strany; velké kladné odchylky od $1 - \alpha$ pokládají za stejně nežádoucí jako odchylky záporné. Poukazují na to, že striktní omezením $cover(p) \geq 1 - \alpha$ vede k tomu, že funkce $cover(p)$ téměř v celém či v úplně celém definičním oboru fakticky probíhá vysoko nad $1 - \alpha$. Důsledkem jsou pak příliš dlouhé intervaly, zbytečně rozsáhlé experimenty atp.

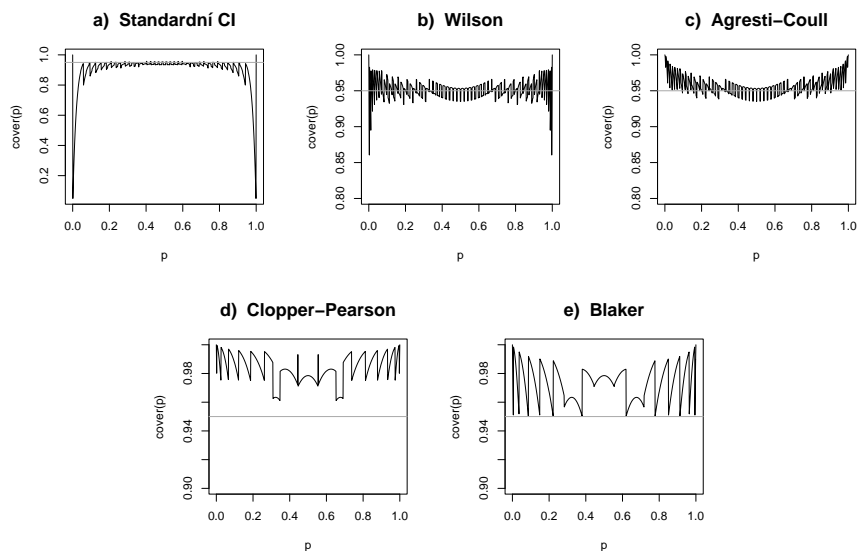
Aby nebyla mýlka: Pokud někdo dává přednost přibližným konfidenčním intervalům, neznamená to ještě zdaleka, že má v oblíbě standardní interval (1). Ví se, že funkce pokrytí $cover^{st}(p) = P_p(p \in CI^{st})$ aproximuje konstantu $1 - \alpha$ špatně (i v mezích, kde literatura metodu doporučuje); na $1 - \alpha$ „se dívá zdola“ příliš často, takže se sotva dá mluvit o kolísání kolem $1 - \alpha$ – viz příklad na obr. 1a. V práci [5], která srovnává řadu různých typů přibližných konfidenčních intervalů pro p , byl standardní interval podroben zdrcující kritice, uzavřené „doporučením“: Neužívat a nečit!

Existuje řada lepších alternativ ke standardnímu intervalu. K těm přijímaným jako nejlepší patří mj. tyto:

- Wilsonův interval je podobně jako standardní interval založen na normální aproximaci binomického rozdělení, ale místo Waldova testu hypotézy $p = p_0$ s testovou statistikou $(\hat{p} - p_0)/\sqrt{\hat{p}(1 - \hat{p})/n}$, invertuje skórový test se statistikou $(\hat{p} - p_0)/\sqrt{p_0(1 - p_0)/n}$. (Vzorec pro meze Wilsonova intervalu viz např. [8].) Tato drobná změna má dramatické důsledky – srovnajme obr. 1a a 1b.
- Agresti-Coullův interval byl navržen jako jednodušší aproximace Wilsonova intervalu. Počítá se jako standardní interval, ale s tím rozdílem, že počet „úspěchů“ i „neúspěchů“ se oproti skutečně pozorovaným počtům zvýší o 2. Příklad funkce pokrytí viz obr. 1c.
- Jeffreysův interval je odvozen Bayesovsky – meze jsou kvantily aposteriorního rozdělení p při apriorním rozdělení $Beta(1/2, 1/2)$.

Více podrobností o výše uvedených i dalších přibližných konfidenčních intervalech pro p (mid-P, arcsinovém, logitovém aj.) lze najít v [5].

Ani Clopper-Pearsonův interval není v rámci exaktního přístupu jedinou a nejlepší možností. Lze mu vytknout, že je příliš konzervativní – funkce pokrytí je typicky v celém definičním oboru ostře (a o dost) větší než nominální hladina spolehlivosti, viz obr. 1d. V článku [8] cituji sérii prací datovaných



Obrázek 1: Pravděpodobnost pokrytí p při $n = 50$ a $1 - \alpha = 0,95$. a) Standardní interval, b) Wilsonův interval, c) Agresti-Coullův interval, d) Clopper-Pearsonův interval, e) Blakerův interval.

léty 1954 až 2000, které přinášejí různé návrhy méně konzervativních konfidenčních intervalů zachovávajících omezení $cover(p) \geq 1 - \alpha$. (Vesmět u nich na rozdíl od Clopper-Pearsonova intervalu smí být jedna z pravděpodobností $P_p(p < p_D)$ a $P_p(p_H < p)$ větší než $\alpha/2$ – „hlídá se“ pouze jejich součet.) Na obr. 1e je pro srovnání s Clopper-Pearsonovým intervalem graf funkce pokrytí Blakerova intervalu (viz [3]) pro stejné n .

Literatura o konfidenčních intervalech pro p by patrně byla chudší a v praxi by nám ubylo nejasností, co vlastně používat, kdybychom uměli intervaly spolehlivosti konstruovat tak, že by se funkce pokrytí $cover(p)$ všude rovnala $1 - \alpha$. Bohužel to neumíme, a není to jen tím, že bychom konfidenční intervaly konstruovali nešikovně. Funkci pokrytí můžeme vyjádřit ve tvaru

$$cover^{CI}(p) = \sum_{i=0}^n I[CI(i)] \cdot P_p(X = i), \quad (3)$$

kde $CI(i)$ je konfidenční interval při $X = i$ a $I[A]$ značí charakteristickou funkci množiny A . Dle (3) je funkce pokrytí po částech polynom stupně n a je nespojitá v krajních bodech intervalů $CI(i)$, kde má skoky o $P_p(X = i)$ – rovnat se v celém definičním oboru konstantě $1 - \alpha$ tedy nemůže. (Všimněme si, že (3) platí pro diskrétní rozdělení obecně. Nespojitá, a tudíž nekonstantní funkce pokrytí tedy zdaleka není jen specifikem binomického rozdělení.)

3 Známenné konfidenční intervaly

Navzdory tomu, co bylo řečeno na konci předcházející sekce, funkce pokrytí vlastně konstantní být může – sáhne-li k randomizaci. Meze známenného konfidenčního intervalu jsou určeny nejen empirickými daty (tj. počtem „úspěchů“ X), ale také realizací náhodné veličiny U (nezávislé na X).

Příkladem známenného konfidenčního intervalu je Stevensův interval publikovaný v r. 1950 (viz [12]). Jeho meze p_D, p_H jsou řešením rovnic

$$\sum_{i=X+1}^n \binom{n}{i} (p_D)^i (1-p_D)^{n-i} + (1-U) \binom{n}{X} (p_D)^X (1-p_D)^{n-X} = \alpha/2, \quad (4)$$

$$\sum_{i=0}^{X-1} \binom{n}{i} (p_H)^i (1-p_H)^{n-i} + U \binom{n}{X} (p_H)^X (1-p_H)^{n-X} = \alpha/2, \quad (5)$$

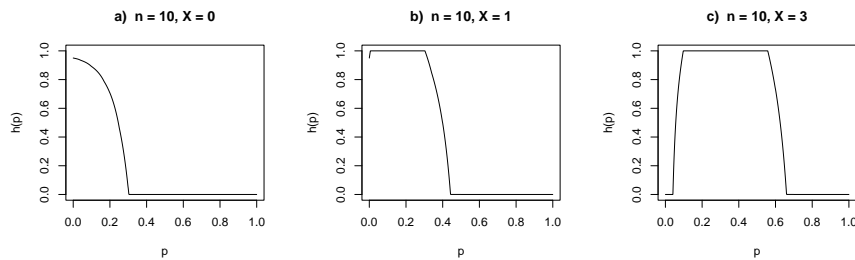
kde U je náhodná veličina s rovnoměrným rozdělením na $[0, 1]$. Přesněji řečeno, p_D je řešením (4) (resp. p_H je řešením (5)), pokud takové řešení je v $[0, 1]$, jinak $p_D = 0$ (resp. $p_H = 1$).

Optimální procedura mezi známennými testy hypotézy $p = p_0$ proti oboustranné alternativě je stejnoměrně nejsilnější nestranný známenný test (UMP_U test; viz [10]). Jeho invertováním získáme Neymanovsky nejkratší nestranný známenný konfidenční interval. („Nejkratší“ je v tom smyslu, že ve třídě nestranných známenných konfidenčních intervalů CI , pro které $\text{cover}^{CI}(p) \geq 1 - \alpha$ na $[0, 1]$, pro $p' \neq p$ minimalizuje $P_p(p' \in CI)$.) Tabulky mezi takového známenného konfidenčního intervalu – říkejme mu známenný UMP_U interval – byly publikovány v r. 1960 ([4]).

Známenný UMP_U interval by mohl znamenat konec všech nejasností, jaký typ konfidenčního intervalu pro p používat – kdyby ovšem praxe byla s to takový „dar“ od teorie přijmout. Kdo pracuje pro uživatele a váží si jejich důvěry, nebude asi snadno sbírat odvahu nabízet jim metodu, která konfidenční interval odpovídající jedněm a týmž datům (tj. n a X) umístí tu více vlevo, tu více vpravo podle toho, jaké náhodné číslo počítač právě vygeneroval.¹

Známenné konfidenční intervaly byly v době, kdy představovaly novinku, považovány za nadějně a očekávalo se, že by se v praxi mohly uchytit. Tak např. E. S. Pearson [11] se domníval, že statistici by mohli přijmout myšlenku *známenní po experimentu* zrovna tak, jako se dříve postupně szili s randomizací experimentu jako takového. V dnešní době však, jak konstatují např. Agresti a Gottard [2], má známenná inferenze význam jen v teorii a o jejím reálném používání prakticky nikdo vážně neuvažuje.

¹U známenných testů se dodatečný náhodný pokus alespoň používá jen v „hraničních“ případech, ale meze typického známenného konfidenčního intervalu závisí na náhodných vstupech vždy, při jakýchkoli empirických datech.



Obrázek 2: Fuzzy UMPU konfidenční interval pro p při $1 - \alpha = 0,95$, $n = 10$ a $X = 0$ (a), 1 (b), 3 (c).

4 Fuzzy konfidenční intervaly

V r. 2005 přišli Geyer a Meeden [7] s novým nápadem, jak dosáhnout u intervalových odhadů parametrů diskrétních rozdělání konstantního pokrytí.

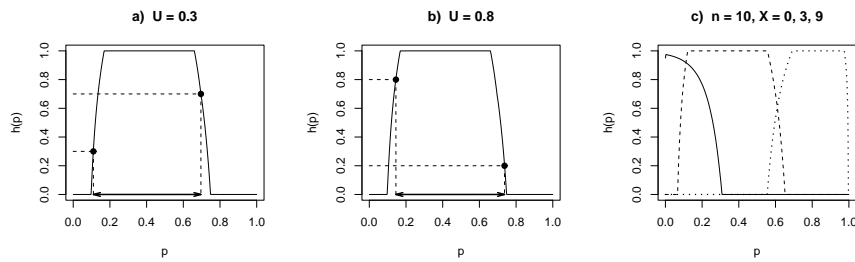
Mějme rodinu rozdělání s parametrem $\theta \in \Theta$ a necht' $\phi(x, \alpha, \theta_0)$ je kritická funkce znáhodněného testu hypotézy $\theta = \theta_0$ na hladině α , tj. pravděpodobnost, že na základě dat x bude hypotéza $\theta = \theta_0$ na hladině α zamítnuta. Je-li test neznáhodněný, tj. ϕ nabývá pouze hodnot 0 a 1, je $1 - \phi(x, \alpha, \theta)$ pro pevné x a α charakteristická funkce konfidenčního intervalu pro θ na hladině spolehlivosti $1 - \alpha$. Obecně (v případě znáhodněného testu) je $h(\theta) = 1 - \phi(x, \alpha, \theta)$ funkce s hodnotami v $[0, 1]$. Geyer a Meeden navrhují brát ji jako fuzzy množinu, nazývat ji *fuzzy konfidenční interval* a interpretovat její hodnoty jako stupeň příslušnosti ke konfidenčnímu intervalu.

Dle zažité terminologie teorie fuzzy množin definují autoři [7] *nosič* (support) a *jádro* (core) konfidenčního intervalu jako množiny $\{\theta; h(\theta) > 0\}$ a $\{\theta; h(\theta) = 1\}$. Do konfidenčního intervalu „úplně patří“ prvky jádra, „částečně patří“ ostatní prvky nosiče a „vůbec nepatří“ prvky Θ mimo nosič.

Pokud hodnoty funkce $h(\theta)$ neinterpretujeme pravděpodobnostně (což bychom snadno mohli, ale Geyer a Meeden se tomu brání), nemá pro fuzzy konfidenční intervaly smysl definice funkce pokrytí jakožto *pravděpodobnosti* pokrytí. Definici ale můžeme rozšířit a položit $cover(\theta) = E_\theta[1 - \phi(X, \alpha, \theta)]$. Takto definovaná funkce pokrytí je pak evidentně na celém Θ rovna $1 - \alpha$, pokud test s kritickou funkcí ϕ je exaktní, tj. když pro všechna α a θ platí $E_\theta\phi(X, \alpha, \theta) = \alpha$.

Vraťme se ale od obecných pojmů k binomickému rozdělení. Geyer a Meeden odvodili fuzzy konfidenční interval pro p odpovídající znáhodněnému UMPU testu a naprogramovali jej jako balík *ump* v R. Na obr. 2 jsou příklady vypočtené tímto softwarem. Právě takto – ve formě grafu funkce $h(\theta)$ – by se podle Geyera a Meedena měly fuzzy konfidenční intervaly prezentovat.

Článek [7] lze zájemci o více podrobností doporučit i s celou diskusí, v níž se velká chvála střídá i s dosti příkrou kritikou.



Obrázek 3: Alternativní fuzzy konfidenční interval podle [2]. a–b) Vztah mezi Stevensovým znáhodněným konfidenčním intervalem a funkcí $h(p)$, c) průběh $h(p)$ pro $1 - \alpha = 0,95$, $n = 10$, $X = 0$ (—), 3 (---), 9 (···).

S asi nejzajímavějším diskusním příspěvkem k článku [7] přišli Agresti a Gottard [2], kteří navrhují alternativní typ fuzzy konfidenčního intervalu pro p . Základem jejich konstrukce je Stevensův znáhodněný konfidenční interval, určený rovnicemi (4) a (5). Konstrukci funkce $h(p)$ popisují způsobem naznačeným na obr. 3a a 3b: Pro X různé od 0 a n meze Stevensova intervalu rostou s rostoucí hodnotou veličiny U . (Speciální případy $X = 0$ a $X = n$ vyžadují pečlivější formulaci, kterou si dovoluji podobně jako autoři [2] pomínout.) Jsou-li pro $U = u$ meze Stevensova intervalu p_D a p_H , je $h(p_D) = u$ a $h(p_H) = 1 - u$. (Tím je funkce $h(p)$ definovaná ve dvou disjunktních intervalech I_D, I_H . Vlevo od I_D a napravo od I_H je $h(p) = 0$, mezi I_D a I_H pak $h(p) = 1$.) Příklady fuzzy konfidenčních intervalů dle [2] jsou na obr. 3c.

Alternativní fuzzy konfidenční interval lze chápat jako „obálku“ na Stevensovy znáhodněné intervaly – z funkce $h(p)$ lze odečíst, jak by znáhodněný interval vypadal pro všechny možné hodnoty náhodného vstupu U . Samotný fuzzy konfidenční interval přitom, což je z hlediska praktické použitelnosti klad, na žádných náhodných vstupech nezávisí a je určen pouze empirickými daty. Poznamenejme, že obdobný vztah, jaký existuje mezi fuzzy intervalem dle [2] a Stevensovým znáhodněným intervalem, existuje také mezi fuzzy UMPU intervalem a znáhodněným UMPU intervalem.

Fuzzy konfidenční interval dle [2] má úzký vztah ke dvěma známým typům „klasických“ konfidenčních intervalů. Jeho nosičem je Clopper-Pearsonův interval. (Nosič fuzzy UMPU intervalu často Clopper-Pearsonův interval obsahuje jako vlastní podmnožinu, což autorům [2] zavání přehnanou konzervativností.) Kdybychom chtěli fuzzy konfidenční interval daného typu aproximovat „klasickým“ intervalem, mohla by k tomu posloužit množina $\{p; h(p) \geq 1/2\}$, což ale není nic jiného než tzv. mid-P interval (přibližný konfidenční interval, který dostaneme, když v rovnicích (4), (5) nahradíme U konstantou $1/2$.) Agresti a Gottard to vyzdvihují jako novou motivaci pro mid-P interval.

Geyer a Meeden předložili v [7] fuzzy konfidenční intervaly velmi sebevědomě, jako ideu, která se musí v praxi prosadit, protože představuje to

pravé řešení starých problémů. Agresti a Gottard [2] ovšem nabádají k větší zdrženlivosti v očekáváních: Znáhodněné konfidenční intervaly také kdysi vstupovaly na scénu s nadějí, že se prosadí, a kde dnes jsou? I kdyby ale fuzzy konfidenční intervaly zůstaly jen kapitolkou v teorii, měli bychom jim být vděční i za skromnější „zisky“, jako je např. již zmiňovaná nová motivace pro mid-P interval.

Sám bych také byl, co se týká šance, že se fuzzy konfidenčních intervaly uchytí v praxi, poměrně skeptický. Ze zkušenosti vím, že pro mnoho uživatelů statistiky je i „obyčejný“ interval spolehlivosti příliš komplikovaný pojem.

Literatura

- [1] Agresti A., Coull A. (1998). *Approximate is better than “exact” for interval estimation of binomial proportions*. Am. Statistician **52**, 119–126.
- [2] Agresti A., Gottard A. (2005). *Comment: Randomized confidence intervals and the mid-P approach*. Statistical Science **20**, 367–371.
- [3] Blaker H. (2000). *Confidence curves and improved exact confidence intervals for discrete distributions*. Canadian J. of Statistics **28**, 783–798.
- [4] Blyth C.R., Hutchinson D.W. (1960). *Table of Neyman-shortest unbiased confidence intervals for the binomial parameter*. Biometrika **47**, 381–391.
- [5] Brown L.D., Cai T.T., DasGupta A. (2001). *Interval estimation for a binomial proportion*. Statistical Science **16**, 101–133.
- [6] Clopper C.J., Pearson E.S. (1934). *The use of confidence or fiducial limits illustrated in the case of the binomial*. Biometrika **26**, 404–413.
- [7] Geyer C.J., Meeden G.D. (2005). *Fuzzy and randomized confidence intervals and P-values*. Statistical Science **20**, 358–366.
- [8] Klaschka J. (2006). *O intervalových odhadech pravděpodobností, zvláště malých*. In: ROBUST 2006. Sborník prací 14. zimní školy JČMF, J. Antoch & G. Dohnal (eds.), Praha, JČMF, 119–126.
- [9] Laplace P.S. (1812). *Théorie Analytique des Probabilités*. Courcier, Paris.
- [10] Lehmann E.L. (1959). *Testing statistical hypotheses*. Wiley, New York.
- [11] Pearson E.S. (1950). *On questions raised by the combination of tests based on discontinuous distributions*. Biometrika **37**, 383–398.
- [12] Stevens W.L. (1950). *Fiducial limits of the parameter of a discontinuous distribution*. Biometrika **37**, 117–129.

Poděkování: Tato práce byla podporována grantem ME 949 MŠMT ČR a výzkumným záměrem AVOZ10300504.

Adresa: J. Klaschka, Ústav informatiky AV ČR, Pod Vodárenskou věží 2, 18207 Praha 8

E-mail: klaschka@cs.cas.cz

LOGISTICKÁ REGRESE V SUBKOHORTNÍCH STUDIÍCH

Petr Klášterecký, Michal Kulich

Klíčová slova: Subkohortní studie, model proporcionálních šancí, logistická regrese.

Abstrakt: V subkohortní studii pozorujeme u sledovaných jedinců zprava cenzorovanou dobu do výskytu určité události. Pro analýzu však máme k dispozici pouze jedince, u nichž událost nastala, a výběr z kontrolní skupiny. Cílem analýzy je obvykle odhadnout parametry vhodného regresního modelu. Subkohortní odhady používané v současnosti nevykazují dobré chování v situacích, kdy má sledovaná událost velmi malou pravděpodobnost. V příspěvku proto navrheme nový odhad s lepšími vlastnostmi, založený na zkombinování odhadů získaných opakovaně logistickou regresí v časech události.

Abstract: In a case-cohort study, right censored times to some event are recorded on the study subjects, however only the cases are followed together with a sample from the control group. The goal of the statistical analysis usually is to estimate the parameters of a suitable regression model. All currently known case-cohort estimators suffer from serious performance problems in situations with rare events. We therefore aim to develop a new estimator with better properties, based on combining logistic regression estimators obtained at failure times.

1 Subkohortní studie

V mnoha statistických aplikacích, zejména pak v medicíně a epidemiologii, nás zajímá doba do výskytu určité události (doba do propuknutí nemoci, doba do vyléčení pacienta atd.) a vztah této doby k dalším faktorům. Ke statistickému zhodnocení podobných dat slouží metody analýzy přežívání, nejčastěji pak Coxův regresní model proporcionálních rizik. Tyto základní metody předpokládají, že analyzovaná data tvoří prostý náhodný výběr z cílové populace, který po určitou dobu sledujeme a zaznamenáváme události a cenzorování.

Pojem subkohortní studie (angl. case-cohort study) zavedl R. L. Prentice v článku [6], přičemž hlavní motivací byla snaha o úsporu nákladů u rozsáhlých epidemiologických studií. V klasické kohortní studii je totiž nutné sledovat všechny jedince, přestože nejpodstatnější informaci (pro všechny používané modely) nesou ti, u nichž se vyskytla pozorovaná událost (případy, cases). Subkohortní studie se omezují na sledování případů a pouze malého vzorku z kontrolní skupiny. Čím vzácnější je ve sledované populaci výskyt dané události, tím větší je v klasické kohortní studii podíl kontrolní skupiny a tím větší také je úspora plynoucí z využití studie subkohortní.

Při analýze dat pocházejících ze subkohortní studie je cílem odhadnout parametry regresního modelu stejně jako v klasické analýze přežívání. Nemáme však k dispozici pozorování kontrolních jedinců, kteří nebyli vybráni do subkohorty. Standardní postupy je tak nutno upravit a tato úprava obvykle spočívá v zavedení vah do odhadovacích rovnic. Nepozorovaným kontrolním jedincům se přiřadí váha nulová, ostatním pak váha zohledňující pravděpodobnost výběru do subkohorty. Různé volby váhových funkcí vedou k různým subkohortním odhadům (stručný přehled např. v [4]). Kromě Coxovy regrese bylo pro subkohortní studie postupně v literatuře publikováno zobecnění odhadovacích rovnic mnoha dalších modelů, ale základní myšlenka vážení jednotlivých příspěvků zůstává stále stejná.

2 Problémy s řídkými jevy

V praxi je poměrně časté vybírat kontrolní skupinu přibližně stejně početnou jako je skupina případů. Události s malou pravděpodobností generující malý počet případů tak vedou k požadavkům na velmi malé subkohorty. To má za následek, že pravděpodobnost výběru i -tého jedince do kontrolní skupiny, označme ji p_i , je také velmi malá. Právě tato pravděpodobnost však vstupuje do vah, které se přidávají do odhadovacích rovnic, a hraje tak podstatnou roli při určování asymptotických vlastností subkohortních odhadů.

Jedním z předpokladů pro platnost asymptotických výsledků je totiž požadavek, aby tato pravděpodobnost byla tzv. odražená od nuly, tj. $p_i \geq \epsilon > 0$. Důkazy asymptotické normality a konzistence subkohortních odhadů se opírají o aproximaci subkohortní skórové funkce součtem skórové funkce pro úplná data a zbytkového členu. Výsledky dokázané již dříve pro skórové funkce při úplných datech potom dávají požadované vlastnosti i pro subkohortní odhady, ale bez splnění předpokladu na kladné pravděpodobnosti výběru nelze ukázat konvergenci zbytkového členu k nule a celá konstrukce tudíž nevede k cíli (podrobněji viz např. [2], kap. 3).

V příspěvku na minulém Robustu [3] jsme ukázali, jak špatně mohou dopadnout subkohortní odhady při modelování velmi řídkých jevů. Simulovali jsme situaci s 1% podílem případů a při vybrání srovnatelně velké nebo menší kontrolní skupiny bylo patrné vychýlení odhadů, menší než předepsaná pravděpodobnost pokrytí u konfidenčních intervalů a špatná konvergence k teoretickému asymptotickému rozdělení odhadů. Uvedené problémy byly tím patrnější, čím menší byla pravděpodobnost výběru do subkohorty, tj. podíl velikostí subkohorty a původní populace. Při výběru velké subkohorty se odhady chovají podle očekávání dobře, ale na druhou stranu subkohortní studie s velkou subkohortou již ztrácí na své úspornosti.

3 Model proporcionálních šancí

Všechny dosud známé subkohortní odhady využívají principu vážení a týkají se jich tak všechny uvedené problémy. Proto bylo naším cílem vyvinout

odhad, který by se choval lépe. Označme T dobu do události, $S(t)$ funkci přežití T (doplňek distribuční funkce), $\delta(t)$ indikátor udávající, zda se do času t vyskytla sledovaná událost, a symbolem $\mathbf{Z} = (Z_1, \dots, Z_p)$ vektor regresních proměnných. Z mnoha modelů používaných v analýze přežívání pro naše účely nejlépe vyhovuje model proporcionálních šancí, který lze definovat rovnicí¹

$$\text{logit}(\text{P}[\delta(t) = 1 | \mathbf{Z} = \mathbf{z}]) = \alpha(t) + \beta' \mathbf{z}. \quad (1)$$

Jedná se o semiparametrický model, jehož parametry lze interpretovat jako referenční logaritmus šance (log odds) na výskyt události v čase t – funkce $\alpha(t)$, respektive logaritmičké poměry šancí (log odds ratios) na výskyt události – parametry β_1, \dots, β_p . Model předpokládá, že tyto poměry šancí zůstávají v čase neměnné (podobně jako u Coxova modelu zůstává konstantní relativní riziko).

Všimněme si jisté podobnosti s modelem logistické regrese. Kdybychom subkohortní studii zastavili v nějakém pevně daném čase t_1 a podívali se na naměřená data, mohli bychom vyjádřit vztah mezi $\delta(t_1)$ a regresory logistickým modelem

$$\text{logit}(\text{P}[\delta(t_1) = 1 | \mathbf{Z} = \mathbf{z}]) = \alpha(t_1) + \beta' \mathbf{z} = \alpha_1 + \beta' \mathbf{z}. \quad (2)$$

Poměry šancí β zavedené v (1) se v čase nemění a jsou proto v obou modelech (1) i (2) totožné, navíc logistická regrese umožňuje tyto poměry šancí odhadovat bez nutnosti vážit jednotlivé příspěvky do skórové funkce. Pokud tedy nepotřebujeme odhadovat absolutní člen, zdánlivě bychom mohli jednoduchou logistickou regresí, provedenou například na konci studie, nahradit komplikované procedury vyvinuté speciálně pro model proporcionálních šancí. Taková analýza je však značně neefektivní, protože nevyužívá všechny dostupné informace – konkrétně zahazuje časy a pořadí jednotlivých událostí.

4 Alternativní odhad

Jako jedno z možných řešení jak zachytit časový faktor se nabízí provést analýzu pomocí logistické regrese opakovaně a získané odhady poté zkombinovat do odhadu nového. Předpokládejme prozatím, že jediné přípustné cenzorování je cenzorování konstantou – časem τ (například konec studie). Předpokládejme dále, že skutečné parametry modelu proporcionálních šancí jsou rovny $\alpha_0(t)$ a β_0 a označme jejich odhady získané pomocí logistické regrese v předem daných časech t_1, \dots, t_K jako $\hat{\alpha}(t_1), \hat{\beta}_1, \dots, \hat{\alpha}(t_K), \hat{\beta}_K$. Tyto odhady přitom počítáme z dat dostupných v konkrétním čase analýzy. Ke kontrolní skupině tak postupně přidáváme případy, nebo měníme status pozorování, která byla vybrána do kontrolní skupiny, ale před časem τ se u nich vyskytla událost.

¹V literatuře se model proporcionálních šancí obvykle definuje pomocí funkce přežití rovnicí $-\text{logit}(S(t | \mathbf{Z} = \mathbf{z})) = G(t) + \beta' \mathbf{z}$. Jelikož $S(t) = \text{P}[T > t] = 1 - \text{P}[\delta(t) = 1]$ a záporné znaménko zamění v logaritmu čitatele a jmenovatele, jsou oba zápisy ekvivalentní.

Označme symbolem \mathbf{I} jednotkovou matici a definujme pro libovolné matice konstant $\mathbf{W}_1, \dots, \mathbf{W}_K$ typu $p \times p$ takové, že $\sum_{k=1}^K \mathbf{W}_k = \mathbf{I}_{p \times p}$, třídu kombinovaných logistických odhadů vztahem

$$\tilde{\beta}_{\mathbf{W}} = \sum_{k=1}^K \mathbf{W}_k \hat{\beta}_k. \quad (3)$$

Tvrzení: Každý odhad $\tilde{\beta}_{\mathbf{W}}$ z třídy (3) je konzistentní a asymptoticky normální, tj.

$$\sqrt{n}(\tilde{\beta}_{\mathbf{W}} - \beta_0) \xrightarrow{\mathcal{D}} N_p(\mathbf{0}, \Sigma_{\mathbf{W}}), \quad (4)$$

kde $\Sigma_{\mathbf{W}}$ závisí na váhových maticích $\mathbf{W} = \{\mathbf{W}_1, \dots, \mathbf{W}_K\}$.

Důkaz: Konzistence je zřejmá díky skutečnosti, že každý z logistických odhadů $\hat{\beta}_k$ konzistentně odhaduje β_0 a $\sum_{k=1}^K \mathbf{W}_k = \mathbf{I}_{p \times p}$. Asymptotická normalita pak plyne ze sdružené asymptotické normality vektoru $(\hat{\beta}_1, \dots, \hat{\beta}_K)'$. Skutečnost, že vektor odhadů má asymptoticky sdružené normální rozdělení lze ukázat přes sdružené rozdělení skórových funkcí, které je také asymptoticky normální. Podrobněji viz [2], kap. 4. ■

Ze všech možných odhadů bychom samozřejmě chtěli zvolit ten, který bude podle nějakých kritérií nejlepší. Jedním z takových kritérií může být rozptyl odhadu, což vede k úloze najít kombinaci \mathbf{W}^{opt} , která tento rozptyl minimalizuje. Váhové matice ale závisí na datech skrze varianční matice jednotlivých logistických odhadů. To znamená, že by bylo nutné odhadnout pro všechna $k = 1, \dots, K$ dalších p^2 parametrů (všechny hodnoty matice $\mathbf{W}_k^{\text{opt}}$). Přitom již pro $K = 2$ je výpočet numericky nestabilní a výsledky nespolehlivé (viz [4]). Jako řešení tohoto problému navrhuje omezit se v (3) na diagonální matice $\mathbf{W}_{k(p \times p)} = \text{diag}\{\mathbf{w}_k\} = \text{diag}\{(w_{1k}, \dots, w_{pk})\}$, $k = 1, \dots, K$.

Omezením na diagonální matice se zbavíme $Kp(p-1)$ variančních parametrů, stále ovšem pracujeme s vysoce korelovanými odhady získanými z jednotlivých logistických analýz. Často tak dochází ke kompenzaci vlivu jednoho odhadu jiným (je přiřazena podobně velká váha s opačným znaménkem). Tomu zabráníme dalším omezením na váhové matice – kromě diagonální struktury budeme navíc požadovat, aby lineární kombinace generované v (3) byly konvexní. Výsledný odhad $\tilde{\beta}$ je pak možno počítat po složkách jako

$$\tilde{\beta}_j = \sum_{k=1}^K w_{jk} \hat{\beta}_{jk}, \quad \sum_{k=1}^K w_{jk} = 1, \quad 0 \leq w_{jk} \leq 1 \quad (5)$$

a rozptyl každé ze složek kombinovaného odhadu $\tilde{\beta}$ je dán vztahem

$$\text{Var} \tilde{\beta}_j = \sum_{k=1}^K w_{jk}^2 \text{Var} \hat{\beta}_{jk} + 2 \sum_{k=1}^K \sum_{l=k+1}^K w_{jk} w_{jl} \text{Cov}(\hat{\beta}_{jk}, \hat{\beta}_{jl}). \quad (6)$$

Při použití asymptotických rozptylů a kovariancí pak tvoří prvky (6) diagonálu asymptotické varianční matice $\Sigma_{\mathbf{W}}$. Jako optimální kombinovaný

logistický odhad $\tilde{\beta}^{\text{opt}}$ označíme odhad, který tyto diagonální prvky $\Sigma_{\mathcal{W}}$ minimalizuje. Tuto optimalizační úlohu lze v případě známých (asymptotických) variančních matic $\text{Var}\hat{\beta}_k$ a meziodhadových kovariancí $\text{Cov}(\hat{\beta}_k, \hat{\beta}_{k'})$, $k, k' = 1, \dots, K$ řešit pomocí standardních numerických algoritmů, například kvazi newtonovským L-BFGS-B algoritmem s omezujícími podmínkami (viz [1]). Skutečné hodnoty variančních a kovariančních matic opět neznáme, nicméně lze ukázat, že je můžeme nahradit jejich odhady při zachování platnosti všech výsledků.

Odhady asymptotických rozptylů $\hat{\beta}_{jk}$ pro prospektivní studie (pozorujeme výskyt události při daných regresorech) známe z klasické teorie pro maximálně věrohodné odhady parametrů v logistické regresi. Při analýze retrospektivních dat máme k dispozici výběr z rozdělení regresorů při dané odezvě (událost nastala/nenastala) a právě marginální hustota regresorů obecně vstupuje do analýzy jako rušivý parametr. Prentice a Pyke [7] ukázali, že v modelu logistické regrese lze tento rušivý parametr eliminovat a retrospektivní data analyzovat pomocí formálně stejného logistického modelu jako data prospektivní. Odhady poměrů šancí a jejich asymptotická varianční matice přitom zůstávají v platnosti, mění se pouze odhad absolutního členu v modelu a souvisejících nebo odvozených charakteristik.

Detailnější rozbor lze kromě původního článku [7] nalézt v [2], kde je také uveden výpočet kovariancí mezi jednotlivými odhady. Jeho odvození je založeno na výpočtu kovariance $\text{Cov}(n^{-1/2}\mathbf{U}(s), n^{-1/2}\mathbf{U}(t))$ mezi normovanými logistickými skórovými funkcemi vyhodnocenými ve dvou časech t a s . Postupným dosazováním t_1, \dots, t_K za t a s následně dostaneme všech K^2 bloků kovarianční matice sdruženého rozdělení vektoru logistických skóre ze všech jednotlivých analýz. Pro výpočet $\text{Cov}(\hat{\beta}_{jk}, \hat{\beta}_{jl})$ pak nahradíme každé $\hat{\beta}_k$ aproximací založenou na Taylorově rozvoji skórových funkcí kolem skutečné hodnoty parametru β_0 a použijeme zmíněné meziodhadové kovariance normovaných skórových funkcí.

5 Časy jednotlivých analýz

Podívejme se nyní blíže na možnosti volby časů, ve kterých studii jakoby zastavujeme a provádíme analýzu logistickou regresí. Datový soubor, který analyzujeme, se mezi dvěma danými časy může změnit pouze v případě, že nastane sledovaná událost. Pak buď přibude jedno nebo více pozorování, a/nebo některá pozorování přejdou z kontrolní skupiny mezi případy. Nabízí se proto myšlenka opustit pevně volené časy zastavení t_1, \dots, t_K a provádět logistickou analýzu po každé pozorované události. Vkládání dalších odhadů mezi dva časy událostí je neefektivní, neboť by se počítaly regresní odhady z naprosto stejných dat, naopak vynechání některého z časů událostí vede ke ztrátě informace a hrubší aproximaci.

Ukážeme, že také věrohodnost pro model proporcionálních šancí se mění pouze v časech událostí. Murphy a kol. v [5] odvodili, že věrohodnostní funkci v modelu proporcionálních šancí lze zapsat ve tvaru

$$\begin{aligned}
L_{\text{PO}}(H, \boldsymbol{\beta}) &= \prod_{i=1}^n \left[\frac{\exp(\mathbf{Z}'_i \boldsymbol{\beta}) h(y_i)}{(1 + H(y_i) \exp(\mathbf{Z}'_i \boldsymbol{\beta}))^2} \right]^{\delta_i} \left[\frac{1}{1 + H(y_i) \exp(\mathbf{Z}'_i \boldsymbol{\beta})} \right]^{1-\delta_i} \\
&= \prod_{i=1}^n \left[\frac{\exp(\mathbf{Z}'_i \boldsymbol{\beta}) h(y_i)}{1 + H(y_i) \exp(\mathbf{Z}'_i \boldsymbol{\beta})} \right]^{\delta_i} \left[\frac{1}{1 + H(y_i) \exp(\mathbf{Z}'_i \boldsymbol{\beta})} \right], \quad (7)
\end{aligned}$$

kde y_i značí cenzorovanou dobu do události pozorování s indexem i , $i = 1, \dots, n$, $H(t) = \exp\{\alpha(t)\}$ a $h(t) = \partial/\partial t\{H(t)\}$. Zápis (7) je sice kompaktní, ale poněkud maskuje vývoj závislosti věrohodnostní funkce na čase. Označme symboly f_1, \dots, f_K indexy případů ve výběru a jako t_{f_1}, \dots, t_{f_K} pak jejich doby do události. Jednoduchou úpravou (7) můžeme získat tvar

$$\begin{aligned}
L_{\text{PO}}(t) &= 1 \text{ pro } t < t_{f_1}, \\
L_{\text{PO}}(t_{f_1}) &= 1 \times \frac{h(t_{f_1}) \exp(\mathbf{Z}'_{f_1} \boldsymbol{\beta})}{1 + H(t_{f_1}) \exp(\mathbf{Z}'_{f_1} \boldsymbol{\beta})} \times \prod_{i=f_1}^n \frac{1}{1 + H(t_{f_1}) \exp(\mathbf{Z}'_i \boldsymbol{\beta})}, \\
L_{\text{PO}}(t) &= L_{\text{PO}}(t_{f_1}) \text{ pro } t \in (t_{f_1}, t_{f_2}), \\
L_{\text{PO}}(t_{f_2}) &= 1 \times \frac{h(t_{f_1}) \exp(\mathbf{Z}'_{f_1} \boldsymbol{\beta})}{1 + H(t_{f_1}) \exp(\mathbf{Z}'_{f_1} \boldsymbol{\beta})} \times \prod_{i=f_1}^{f_2-1} \frac{1}{1 + H(t_{f_1}) \exp(\mathbf{Z}'_i \boldsymbol{\beta})} \\
&\quad \times \frac{h(t_{f_2}) \exp(\mathbf{Z}'_{f_2} \boldsymbol{\beta})}{1 + H(t_{f_2}) \exp(\mathbf{Z}'_{f_2} \boldsymbol{\beta})} \times \prod_{i=f_2}^n \frac{1}{1 + H(t_{f_2}) \exp(\mathbf{Z}'_i \boldsymbol{\beta})}, \\
L_{\text{PO}}(t) &= L_{\text{PO}}(t_{f_2}) \text{ pro } t \in (t_{f_2}, t_{f_3}) \text{ atd.}
\end{aligned}$$

Kdybychom stejným způsobem pokračovali dále, dostaneme vyjádření

$$L_{\text{PO}}(t_K) = \prod_{k=1}^K \left[\frac{h(t_{f_k}) \exp(\mathbf{Z}'_{f_k} \boldsymbol{\beta})}{1 + H(t_{f_k}) \exp(\mathbf{Z}'_{f_k} \boldsymbol{\beta})} \prod_{i=f_k}^{f_{k+1}-1} \frac{1}{1 + H(t_{f_k}) \exp(\mathbf{Z}'_i \boldsymbol{\beta})} \right], \quad (8)$$

kde jsme pro usnadnění zápisu položili $f_{K+1} = n + 1$. Ihned vidíme, že věrohodnostní funkce (8) se mění právě v časech událostí a jinak zůstává konstantní. Přejít od pevně volených časů jednotlivých analýz k náhodným časům událostí vyžaduje jistou opatrnost při přejímání teoretických výsledků. Lze však ukázat, že i v tomto případě je výsledný kombinovaný odhad konzistentní a asymptoticky normální.

Zavedením náhodných časů jednotlivých analýz zároveň také připouštíme situace, kdy bychom odhadovali parametry z dat s několika málo událostmi (odhady v prvních logistických modelech). Takové odhady je nutno vyloučit a pro výsledný kombinovaný odhad uvažovat až odhady $\hat{\boldsymbol{\beta}}_k$ pro $k > k_0$. Volba parametru k_0 je subjektivní, v simulačních studiích vykazovala dobré výsledky hodnota $k_0 = 30$. Při volbě $k_0 = K - 1$ dostáváme jako speciální případ jedinou logistickou regresi provedenou na konci studie. Podrobněji je volba tohoto parametru diskutována v práci [2].

6 Závěrem

V příspěvku jsme navrhli nový odhad regresních parametrů pro model proporcionálních šancí, založený na zkombinování odhadů získaných opakovaně, v časech událostí, pomocí logistické regrese. Souhrnně můžeme jeho výpočet popsat následující procedurou:

1. Zvolíme parametr k_0 .
2. Odhadneme poměry šancí pomocí logistické regrese po každé události, získáme tak odhady $\hat{\beta}_{k_0+1}, \dots, \hat{\beta}_K$.
3. Odhadneme asymptotické kovarianční matice odhadů $\hat{\beta}_k$ a kovariance mezi jednotlivými odhady $\text{Cov}(\hat{\beta}_k, \hat{\beta}_{k'})$.
4. Odhadneme optimální diagonální váhové matice a spočteme kombinovaný odhad $\tilde{\beta}^{\text{opt}}$.
5. Odhadneme Σ_W^{opt} , asymptotickou varianční matici odhadu $\tilde{\beta}^{\text{opt}}$.

Přestože hlavní motivací bylo selhávání odhadů používaných v současné době v subkohortních studiích, je uvedený odhad použitelný obecně i pro úplná data. Pro posouzení jeho kvality jsme provedli rozsáhlou simulační studii, v níž jsme jeho chování porovnali s chováním ostatních běžně používaných odhadů na úplných i subkohortních datech. Zatímco při práci s úplnými daty fungoval kombinovaný odhad podobně dobře jako eficientní maximálně věrohodný odhad navržený v [5], na subkohortních datech vykazoval pouze zanedbatelné vychýlení a až o 50 % nižší střední čtvercovou chybu než standardní subkohortní odhady. Podrobný popis a výsledky studie lze prozatím nalézt v [2], časopisecky budou publikovány později.

Pro odvození odhadu byl z technických důvodů zvolen model proporcionálních šancí, který lze chápat jako přímé zobecnění modelu logistické regrese. Má stejně interpretovatelné parametry, podobné odhadovací rovnice (viz [2]) a standardní subkohortní odhady parametrů u něho při malých pravděpodobnostech výběru selhávají podobně jako u ostatních modelů. Při analýze velmi řídkých jevů však lze výsledný kombinovaný odhad bez problémů použít i pro odhad parametrů v Coxově modelu proporcionálních rizik. Pokud je totiž obecně pravděpodobnost výskytu události v populaci velmi malá, lze relativní riziko odhadované Coxovým modelem dobře aproximovat právě poměrem šancí.

Samotný vývoj kombinovaného odhadu je stále ještě spíše na počátku, i zde jsme při odvození odhadu například pro jednoduchost předpokládali cenzorování konstantou (koncem studie) a neexistenci shod v časech událostí. Zatímco případné shody na výpočtu ani vlastnostech odhadu nic nemění, vliv obecnějších typů cenzorování bude třeba ještě podrobit dalšímu výzkumu. První výsledky jsou prozatím k dispozici pro rovnoměrné cenzorování nezávislé na době do události a budou publikovány v další práci.

Literatura

- [1] Byrd R.H., Lu P., Nocedal J., Zhu C. (1995). *A limited memory algorithm for bound constrained optimization*. SIAM J. Scientific Computing, **16**, 1190–1208.
- [2] Klášterecký P. *Odhady parametrů v subkohortních studiích*. Doktorská dizertační práce, MFF UK Praha.
- [3] Klášterecký P., Kulich M. (2006). *A note on parameter estimation in regression models for case-cohort data*. Sborník Robust'06, JČMF, Praha, 127–134.
- [4] Kulich M., Lin D.Y. (2004). *Improving the efficiency of relative-risk estimation in case-cohort studies*. JASA **99**, 832–844.
- [5] Murphy S.A., Rossini A.J., van der Vaart A.W. (1997). *Maximum likelihood estimation in the proportional odds model*. JASA, **92**, 968–976.
- [6] Prentice R.L. (1986). *A case-cohort design for epidemiologic cohort studies and disease prevention trials*. Biometrika **73**, 1–11.
- [7] Prentice R.L., Pyke R. (1979). *Logistic disease incidence models and case-control studies*. Biometrika **66**, 403–411.

Poděkování: Tato práce vznikla za podpory výzk. záměru MSM 0021620839.

Adresa: P. Klášterecký, M. Kulich, MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8 – Karlín

E-mail: klaster@karlin.mff.cuni.cz

R BALÍČEK PRO ODHAD JEDNO I VÍCE-ROZMĚRNÝCH ROZDĚLENÍ POMOCÍ NORMÁLNÍCH SMĚSÍ, PŘÍPADNĚ S INTERVALOVĚ CENZOROVANÝMI DATY

Arnošt Komárek

Klíčová slova: Monte Carlo Markovův řetězec, odhad hustoty, software, transdimenzionální simulace.

Abstrakt: V článku je představen R balíček `mixAK`, jenž je určen k odhadu neznámé jedno i vícerozměrné hustoty, přičemž data mohou být cenzorována. Odhad je proveden metodou MCMC na základě modelu normální směsi.

Abstract: The present paper introduces R package `mixAK` which is designed to estimate uni- and multivariate density. At the same time, the data are allowed to be censored. The estimation proceeds using the MCMC method and the normal mixture model.

1 Úvod

Na mnoha místech bylo ukázáno, že normální směsi jsou vhodným semiparametrickým nástrojem pro odhad neznámého spojitého rozdělení. Kromě odhadu rozdělení lze směsí využít též k diskriminaci či shlukování (viz např. McLachlan a Basford [6]). Jako jedni z prvních navrhli Diebolt a Robert [2] využití MCMC metody a Gibbsova algoritmu k bayesovskému odhadu normální směsi s předem daným počtem komponent. Následně Richardsonová a Green [8] využili transdimenzionální MCMC (RJ-MCMC) k současnému odhadu počtu složek jednorozměrné normální směsi i jejích parametrů. K výběru počtu složek směsi lze však využít též některého z ukazatelů vhodnosti a složitosti modelu. V kontextu bayesovské statistiky je devianční informační kritérium (DIC) jedním z často používaným ukazatelem tohoto typu, viz Spiegelhalter a kol. [11], Celeux a kol. [1].

V současné době je možné pomocí MCMC odhadovat normální směsi s předem daným počtem složek programovým balíkem WinBUGS (Lunn a kol. [4]), který se stal víceméně standardem pro aplikované problémy bayesovské statistiky. Nicméně výpočet DIC v kontextu směsí lze ve WinBUGSu provést pouze „ručně“ a s jistou dávkou programovacího úsilí. Pro případ současného odhadu počtu složek směsi s ostatními parametry směsi lze nalézt návrhy na implementaci obecně použitelného transdimenzionálního MCMC ve WinBUGSu v práci Lunn a kol. [5]. Avšak v případě směsí neposkytuje jejich přístup plnou flexibilitu jako např. algoritmus popsany Richardsonovou a Greenem [8]. V neposlední řadě, mnohý uživatel je od používání WinBUGSu odrazen faktem, že v současné době ho lze spouštět pouze pod operačním

systémem jednoho konkrétního výrobce. Na druhou stranu v současnosti je balík R (R Development Core Team [7]) jedním (nejen v akademické sféře) z nejrozšířenějších a nejprogresivněji se rozvíjejících statistických programových balíčků. Smyslem práce popsané v tomto článku bylo rozšířit R o možnosti bayesovského odhadu normálních směsí pomocí (transdimenzionálního) MCMC s automatickým výpočtem DIC pro volbu počtu složek směsi v případě, že není odhadován spolu s ostatními parametry směsi. S ohledem na autorovy potřeby a širší uplatnění umožňuje pro tyto účely vytvořený R balíček `mixAK` též odhad směsí v případě, kdy data jsou zprava, zleva či intervalově cenzorována. Pro zajištění rozumné výpočetní rychlosti při MCMC výpočtu jsou hlavní procedury balíčku `mixAK` implementovány v C++, přičemž R volá pouze jejich zkompileovanou variantu. Vzhledem k omezenému prostoru ve sborníku ROBUST 2008 představí tento článek pouze základní principy a krátkou ilustraci. Detaily lze nalézt v práci Komárek [3]. Balíček `mixAK` lze zdarma získat na základě e-mailové žádosti od autora článku.

2 Model

2.1 Model pro neznámou hustotu

Necht' $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ jsou p -dimenzionální i.i.d. náhodné vektory s hustotou $g_y(\mathbf{y})$. Za účelem zahrnutí zprava, zleva a intervalově cenzorovaných pozorování do modelu budeme předpokládat, že pozorujeme $[l_1, \mathbf{u}_1], \dots, [l_n, \mathbf{u}_n]$, kde

$$[l_i, \mathbf{u}_i] = \begin{pmatrix} [l_{i,1}, u_{i,1}] \\ \vdots \\ [l_{i,p}, u_{i,p}] \end{pmatrix} \quad (i = 1, \dots, n). \quad (1)$$

Ve výrazu (1) jsou $-\infty \leq l_{i,j} < \infty$ dolní meze pozorovaných intervalů, $-\infty < u_{i,j} \leq \infty$ horní meze pozorovaných intervalů ($l_{i,j} \leq u_{i,j}$, $i = 1, \dots, n$, $j = 1, \dots, p$) a $[]$ je dle kontextu otevřený, uzavřený či zpola otevřený interval. Jinými slovy, $y_{i,j} = l_{i,j} = u_{i,j}$ pro necenzorované pozorování, $-\infty = l_{i,j} < u_{i,j} < \infty$ značí pozorování cenzorované zleva, $-\infty < l_{i,j} < u_{i,j} = \infty$ označuje zprava cenzorované pozorování a $-\infty < l_{i,j} < u_{i,j} < \infty$ označuje intervalově cenzorované pozorování. Předpokládáme, že cenzorování vedoucí k pozorovaným intervalům je neinformativní. Hustotu $g_y(\mathbf{y})$ vyjádříme následující posunutou a škálovanou normální směsí:

$$g_y(\mathbf{y}) = |\mathbf{S}|^{-1} \sum_{k=1}^K w_k \varphi(\mathbf{S}^{-1}(\mathbf{y} - \mathbf{m}) \mid \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \quad (2)$$

kde $\varphi(\cdot \mid \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ je hustota normálního rozdělení $\mathcal{N}_p(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$, $\mathbf{w} = (w_1, \dots, w_K)'$, $0 \leq w_k \leq 1$, $\sum_{k=1}^K w_k = 1$, je vektor neznámých vah, $\boldsymbol{\mu} = \{\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_K\}$ jsou neznámé střední hodnoty složek směsi a $\boldsymbol{\Sigma} = \{\boldsymbol{\Sigma}_1, \dots, \boldsymbol{\Sigma}_K\}$

jsou neznámé varianční matice složek směsi. Dále, $\mathbf{m} = (m_1, \dots, m_p)'$ je daný vektor posunutí a $\mathbf{S} = \text{diag}(s_1, \dots, s_p)$ je daná škálovací matice. Zahnutí vektoru posunutí \mathbf{m} a škálovací matice \mathbf{S} do modelu bylo provedeno za účelem umožnit zlepšení vlastností MCMC algoritmu, zejména v situacích, kdy jsou jednotlivé složky náhodného vektoru \mathbf{Y} měřeny s výrazně odlišnými měřítky. Konečně necht' $\mathbf{Q} = \{\mathbf{Q}_1, \dots, \mathbf{Q}_K\}$ jsou inverze variančních matic složek směsi, tj., $\mathbf{Q}_k = \boldsymbol{\Sigma}_k^{-1}$ ($k = 1, \dots, K$). Po zbytek článku, necht' $\mathbf{y}^* = \mathbf{S}^{-1}(\mathbf{y} - \mathbf{m})$ je posunutá a škálovaná hodnota vektoru \mathbf{y} a necht'

$$g_{\mathbf{y}^*}(\mathbf{y}^*) = \sum_{k=1}^K w_k \varphi(\mathbf{y}^* | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \quad (3)$$

je hustota náhodného vektoru $\mathbf{Y}^* = \mathbf{S}^{-1}(\mathbf{Y} - \mathbf{m})$. Závěrem necht' $\mathbf{l}_i^* = \mathbf{S}^{-1}(\mathbf{l}_i - \mathbf{m})$ a $\mathbf{u}_i^* = \mathbf{S}^{-1}(\mathbf{u}_i - \mathbf{m})$ ($i = 1, \dots, n$) jsou posunutá a škálovaná meze pozorovaných intervalů.

2.2 Bayesovská specifikace

Balíček `mixAK` využívá bayesovskou metodu a MCMC k odhadu neznámých parametrů. Je proto nutné specifikovat apriorní rozdělení. Necht' $f(\cdot)$ a $f(\cdot | \cdot)$ označují hustotu a podmíněnou hustotu a dále necht'

$$\boldsymbol{\theta} = (w_1, \dots, w_K, \boldsymbol{\mu}'_1, \dots, \boldsymbol{\mu}'_K, \text{vec}(\mathbf{Q}_1)', \dots, \text{vec}(\mathbf{Q}_K)', \boldsymbol{\gamma}')' \quad (4)$$

je vektor neznámých parametrů v situaci, kdy počet komponent směsi je dán. Hyperparametr $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_p)'$ je v modelu zahrnut za účelem možnosti specifikovat neinformativně apriorní rozdělení variančních matic složek směsi. Necht' dále [data] označuje množinu pozorovaných mezí intervalů $\{\mathbf{l}_i, \mathbf{u}_i : i = 1, \dots, n\}$. Dle výrazu (2) je věrohodnost založená na *pozorovaných datech* rovna

$$\begin{aligned} L_{\boldsymbol{\theta}}(\boldsymbol{\theta}, K) &= f([\text{data}] | \boldsymbol{\theta}, K) \\ &= |\mathbf{S}|^{-n} \prod_{i=1}^n \left\{ \int_{\mathbf{l}_{i,1}}^{\mathbf{u}_{i,1}} \cdots \int_{\mathbf{l}_{i,p}}^{\mathbf{u}_{i,p}} \sum_{k=1}^K w_k \varphi(\mathbf{S}^{-1}(\mathbf{y}_i - \mathbf{m}) | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) dy_{i,p} \cdots dy_{i,1} \right\}. \end{aligned} \quad (5)$$

V zápise věrohodnosti je použita konvence $\int_l^u f(y) dy = \int_l^u f(y) dy$, je-li $l < u$ a $\int_l^u f(y) dy = f(l) = f(u)$ je-li $l = u$ (necenzorované pozorování).

Apriorní rozdělení je specifikováno hierarchicky. V případě daného počtu komponent směsi je dáno vztahem

$$f(\boldsymbol{\theta} | K) = f(\mathbf{w} | K) \times f(\boldsymbol{\mu}, \mathbf{Q} | \boldsymbol{\gamma}, K) \times f(\boldsymbol{\gamma}). \quad (6)$$

Pro počet složek směsi K odhadovaný spolu s ostatními parametry je apriorní rozdělení dáno výrazem

$$f(\boldsymbol{\theta}, K) = f(\boldsymbol{\theta} | K) \times f(K). \quad (7)$$

Aposteriorní rozdělení $f(\boldsymbol{\theta} \mid [\text{data}], K)$ (K dáno), respektive $f(\boldsymbol{\theta}, K \mid [\text{data}])$ (K odhadováno spolu s ostatními parametry směsi) je poté dle Bayesova vzorce proporcionální součinu $L_\theta(\boldsymbol{\theta}, K) \times f(\boldsymbol{\theta} \mid K)$, respektive $L_\theta(\boldsymbol{\theta}, K) \times f(\boldsymbol{\theta}, K)$, tj., např.,

$$f(\boldsymbol{\theta}, K \mid [\text{data}]) \propto L_\theta(\boldsymbol{\theta}, K) \times f(\boldsymbol{\theta}, K). \quad (8)$$

Možnosti konkrétních voleb činitelů ve výrazech (6) a (7) lze nalézt v práci Komárek [3].

3 Aposteriorní inference

3.1 MCMC metoda

Aposteriorní inference je založena na metodě MCMC, jejíž použití dále přiblížíme pro případ, kdy je počet složek směsi, K , odhadován současně se zbylými parametry směsi. Pro případ, kdy je K dáno předem je ve všech rozděleních nutno přesunout K do podmínky. Pro použití MCMC je výhodné zavést skryté parametry, jež označíme $\boldsymbol{\psi}$. Sdružené apriorní rozdělení pro $(\boldsymbol{\psi}', \boldsymbol{\theta}', K)'$ je opět zadáno hierarchicky jako

$$f(\boldsymbol{\psi}, \boldsymbol{\theta}, K) = f(\boldsymbol{\psi} \mid \boldsymbol{\theta}, K) \times f(\boldsymbol{\theta} \mid K) \times f(K). \quad (9)$$

Metodou MCMC je zkonstruován Markovův řetězec $\{(\boldsymbol{\psi}^{(t)}, \boldsymbol{\theta}^{(t)}, K^{(t)}) : t = 1, \dots, T\}$, jehož limitním rozdělením je aposteriorní rozdělení $f(\boldsymbol{\psi}, \boldsymbol{\theta}, K \mid [\text{data}])$. Podřetězec $\{(\boldsymbol{\theta}^{(t)}, K^{(t)}) : t = 1, \dots, T\}$ má potom zřejmě $f(\boldsymbol{\theta}, K \mid [\text{data}])$ (aposteriorní rozdělení, jež nás primárně zajímá) za své limitní rozdělení. Charakteristiky aposteriorního rozdělení jsou následně odhadnuty pomocí výběrových charakteristik zkonstruovaného Markovova řetězce, podrobněji viz např. Robert a Casella [9].

V kontextu směsí a cenzorovaných dat je výhodné uvážit dva druhy skrytých parametrů: (i) příslušnosti do komponent označené dále jako $\mathbf{r} = (r_1, \dots, r_n)'$, $r_i \in \{1, \dots, K\}$, (ii) nepozorované hodnoty (posunutých a škálovaných) cenzorovaných dat, tj., hodnoty $\mathbf{y}_1^*, \dots, \mathbf{y}_n^*$ pro ta pozorování, jež jsou cenzorována. Pro jednoduchost značení se další výklad omezí na situaci, kdy jsou všechna pozorování cenzorovaná a vektor $\boldsymbol{\psi}$ je tedy roven $\boldsymbol{\psi} = (\mathbf{y}_1^{*'}, \dots, \mathbf{y}_n^{*'}, r_1, \dots, r_n)'$. První faktor součinu (9) dále faktorizujeme jako

$$f(\boldsymbol{\psi} \mid \boldsymbol{\theta}, K) = \prod_{i=1}^n f(\mathbf{y}_i^*, r_i \mid \boldsymbol{\theta}, K) = \prod_{i=1}^n \{f(\mathbf{y}_i^* \mid r_i, \boldsymbol{\theta}, K) \times f(r_i \mid \boldsymbol{\theta}, K)\},$$

kde

$$(10)$$

$$f(\mathbf{y}_i^* \mid r_i, \boldsymbol{\theta}, K) = f(\mathbf{y}_i^* \mid r_i, \boldsymbol{\theta}) = \varphi(\mathbf{y}_i^* \mid \boldsymbol{\mu}_{r_i}, \boldsymbol{\Sigma}_{r_i})$$

$$f(r_i \mid \boldsymbol{\theta}, K) \equiv P(r_i = k \mid \boldsymbol{\theta}, K) = w_k \quad (k = 1, \dots, K).$$

Snadno lze nahlédnout, že kombinace sdruženého apriorního rozdělení (9) a věrohodnosti

$$\begin{aligned} L_{\psi, \theta}(\boldsymbol{\psi}, \boldsymbol{\theta}, K) &= f([\text{data}] \mid \boldsymbol{\psi}, \boldsymbol{\theta}, K) \\ &= \prod_{i=1}^n f(\boldsymbol{l}_i^*, \boldsymbol{u}_i^* \mid \boldsymbol{y}_i^*) \propto \prod_{i=1}^n I(y_{i,1}^* \in [l_{i,1}^*, u_{i,1}^*], \dots, y_{i,p}^* \in [l_{i,p}^*, u_{i,p}^*]) \end{aligned}$$

získané při využití původních parametrů $\boldsymbol{\theta}$, K a skrytých parametrů $\boldsymbol{\psi}$ vede ke sdruženému aposteriornímu rozdělení

$$f(\boldsymbol{\psi}, \boldsymbol{\theta}, K \mid [\text{data}]) \propto L_{\psi, \theta}(\boldsymbol{\psi}, \boldsymbol{\theta}, K) \times f(\boldsymbol{\psi}, \boldsymbol{\theta}, K), \quad (11)$$

jehož marginální rozdělení $f(\boldsymbol{\theta}, K \mid [\text{data}])$ je rovno původnímu aposteriornímu rozdělení (8).

Konkrétní detaily konstrukce MCMC jsou uvedeny v práci Komárek [3]. Pro případ pevně daného K je v balíčku `mixAK` implementována MCMC simulace pro libovolnou dimenzi p odezvy \boldsymbol{Y} . V případě současného odhadu počet složek směsi K a zbylých parametrů směsi je MCMC simulace v současné době implementována pouze pro jednorozměrná data ($p = 1$).

3.2 Aposteriorní prediktivní hustota

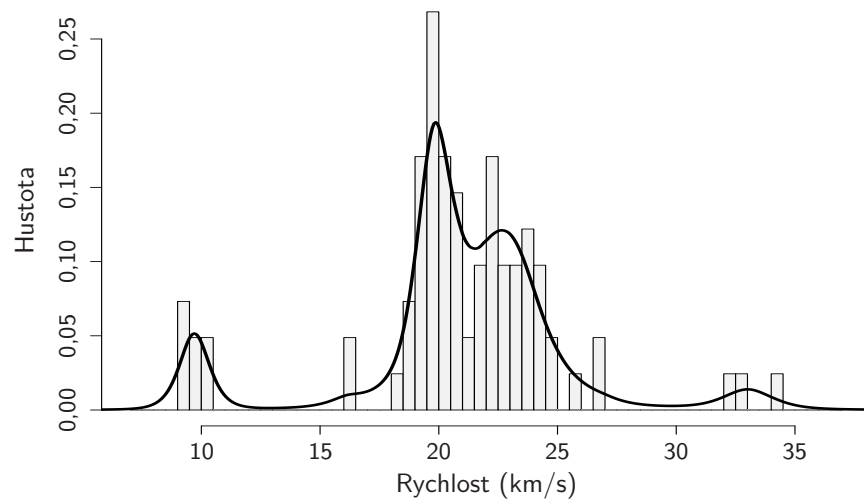
Jedním z primárních cílů popisovaných metod je odhad hustoty náhodného vektoru \boldsymbol{Y} , případně jeho posunutě a škálované varianty \boldsymbol{Y}^* . Vhodným odhadem je aposteriorní prediktivní hustota $E[g_y(\boldsymbol{y}) \mid [\text{data}]]$ případně $E[g_{y^*}(\boldsymbol{y}^*) \mid [\text{data}]]$. Pomocí vygenerovaného Markovova řetězce lze tyto aproximovat jako

$$\begin{aligned} \hat{g}_y(\boldsymbol{y}) &= \frac{1}{T} \sum_{t=1}^T \left\{ |\boldsymbol{S}|^{-1} \sum_{k=1}^{K^{(t)}} w_k^{(t)} \varphi(\boldsymbol{S}^{-1}(\boldsymbol{y} - \boldsymbol{m}) \mid \boldsymbol{\mu}_k^{(t)}, \boldsymbol{\Sigma}_k^{(t)}) \right\} \\ &\approx E[g_y(\boldsymbol{y}) \mid [\text{data}]]. \quad (12) \end{aligned}$$

Analogicky získáme výraz pro $\hat{g}_{y^*}(\boldsymbol{y}^*) \approx E[g_{y^*}(\boldsymbol{y}^*) \mid [\text{data}]]$.

3.3 Devianční informační kritérium

Spiegelhalter a kol. [11] navrhli obecný přístup k porovnávání modelů různé složitosti na základě náhodného výběru z aposteriorního rozdělení (respektive na základě Markovova řetězce s příslušným limitním rozdělením) založený na deviančním informačním kritériu (DIC). V diskuzi k článku Spiegelhaltera a kol. ukázala Richardsonová, jakým způsobem je možné pomocí aposteriorní prediktivní hustoty definovat DIC pro směšové modely a jakým způsobem je poté možno DIC využít k volbě optimálního počtu složek směsi. Problematika DIC v kontextu směsí byla dále rozpracována v práci Celeux a kol. [1]. V balíčku `mixAK` je implementován výpočet DIC dle Richardsonové (označený v práci Celeuxe a kol. jako DIC_3).



Obrázek 1: Data Galaxie: histogram a odhad aposteriorní prediktivní hustoty.

4 Příklad: Data Galaxie

Data Galaxie udávající rychlosti (v km/s) 82 vzdálených galaxií byla v kontextu modelování směsí poprvé použita v práci Roeder [10]. Richardsonová a Green [8] na těchto datech ilustrovali odhad hustoty pomocí transdimenzionálního MCMC. Jejich analýzu lze pomocí balíčku `mixAK` jednoduše zreprodukovat následujícím způsobem:

```
### Načtení dat
> data(Galaxy)

### Specifikace apriorního rozdělení a dalších parametrů MCMC
### odpovídající volbám Richardsonové a Greena
> GalaxyPrior <- list(priorK="uniform", Kmax=30, delta=1,
+   priormuQ="independentC", xi=21.73, D=630.5121,
+   zeta=2*2, g=0.2, h=0.016/2)
> parRJMCMC <- list(par.u1=c(2, 2), par.u2=c(2, 2), par.u3=c(1, 1))

### MCMC simulace
> GalaxyModel <- NMixMCMC(y0=Galaxy, prior=GalaxyPrior,
+   RJMCMC=parRJMCMC,
+   nMCMC=c(burn=100000, keep=100000, thin=1, info=10000),
+   scale=list(shift=0, scale=1))

### Základní charakteristiky aposteriorního rozdělení
> print(GalaxyModel)

Normal mixture with at most 30 components estimated using RJ-MCMC
=====
Posterior distribution of K:
```



```
-----
      1      2      3      4      5      6      7      8
0.00012 0.00452 0.03598 0.11724 0.22201 0.24206 0.18451 0.10610
      9     10     11     12     13     14     15
0.05116 0.02188 0.00849 0.00380 0.00149 0.00055 0.00009
```

Deviance information criteria:

```
-----
      DIC      pD      D.bar D.in.bar
DIC3 442.4403 22.00288 420.4375 398.4346
```

Posterior summary statistics for moments of mixture for original data:

Mean:

```
      Mean      Std.Dev.      Min.      2.5%      1st Qu.
20.8599025 0.7842305 1.3313071 19.6654139 20.4947485
      Median      3rd Qu.      97.5%      Max.
20.8639968 21.2289407 22.0401376 42.8685378
```

Standard deviation:

```
      Mean      Std.Dev.      Min.      2.5%      1st Qu.
5.0154316 1.0198730 0.1992252 3.7957217 4.4628958
      Median      3rd Qu.      97.5%      Max.
4.8440500 5.2943971 7.6354345 28.4306825
```

Z výstupu lze např. vyčíst, že aposteriorně jsou nejvíce zastoupeny směsi s $K=5$, 6, respektive 7 komponentami vykazující $P(K = k | [\text{data}])$ rovnou 0,22, 0,24, respektive 0,18 ($k = 5, 6, 7$).

```
### Aposteriorní prediktivní hustota a její graf
> GalaxyPDens <- NMixPredDensMarg(GalaxyModel)
> plot(GalaxyPDens)
```

Posledně provedený příkaz nakreslí základní obrázek odhadu aposteriorní prediktivní hustoty. S pomocí standardních funkcí programu R lze následně využít spočtený objekt `GalaxyPDens` k nakreslení obrázku v uživatelsky požadovaném formátu, jak je ukázáno na obr. 1.

5 Závěr

V tomto článku jsme stručně představili R balíček `mixAK` využitelný k bayesovskému odhadu neznámého rozdělení pomocí směsí. U jednorozměrného rozdělení lze počet složek směsi odhadnout současně se zbylými parametry pomocí transdimenzionálního MCMC. Pro vícerozměrná rozdělení lze volbu počtu složek směsi založit na deviančním informačním kritériu. Představený balíček dále umožňuje cenzorování dat zprava, zleva i cenzorované. Více detailů spolu s příklady analýzy vícerozměrných a cenzorovaných dat lze nalézt v práci Komárek [3]. Podrobný technický popis syntaxe pro analýzu představenou v tomto článku i práci [3] je k dispozici v rámci dokumentace balíčku `mixAK` dostupné po jeho nainstalování.

Literatura

- [1] Celeux G., Forbes F., Robert C.P., Titterington D.M. (2006). *Deviance information criteria for missing data models*. Bayesian Analysis, **1**, 651–674.
- [2] Diebolt J., Robert C.P. (1994). *Estimation of finite mixture distributions through Bayesian sampling*. Journal of the Royal Statistical Society, Series B, **56**, 363–375.
- [3] Komárek A. (2008). *A new R package for Bayesian estimation of multivariate normal mixtures allowing for selection of the number of components and interval-censored data*. V recenzním řízení, předběžná verze na CD ROBUST 2008.
- [4] Lunn D.J., Thomas A., Best N., Spiegelhalter D. (2000). *WinBUGS – A Bayesian modelling framework: Concepts, structure, and extensibility*. Statistics and Computing, **10**, 325–337.
- [5] Lunn D.J., Best N., Whittaker J. (2005). *Generic reversible jump MCMC using graphical models*. Tech. Rep. EPH-2005-01, Department of Epidemiology and Public Health, Imperial College, London.
- [6] McLachlan G.J., Basford K.E. (1988). *Mixture models: inference and applications to clustering*. New York: Marcel Dekker, Inc.
- [7] R Development Core Team (2008). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0. <http://www.R-project.org>.
- [8] Richardson S., Green P.J. (1997). *On Bayesian analysis of mixtures with unknown number of components (with Discussion)*. Journal of the Royal Statistical Society, Series B, **59**, 731–792.
- [9] Robert C.P., Casella G. (2004). *Monte Carlo statistical methods, second edition*. New York: Springer-Verlag.
- [10] Roeder K. (1990). *Density estimation with confidence sets exemplified by superclusters and voids in the galaxies*. Journal of the American Statistical Association, **85**, 617–624.
- [11] Spiegelhalter D.J., Best N.G., Carlin B.P., van der Linde A. (2002). *Bayesian measures of model complexity and fit (with Discussion)*. Journal of the Royal Statistical Society, Series B, **64**, 583–639.

Poděkování: Tato práce byla podporována grantem Ministerstva školství, mládeže a tělovýchovy České republiky MSM 0021620839.

Adresa: A. Komárek, MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8 – Karlín

E-mail: arnost.komarek@mff.cuni.cz

TEXT MINING OF ARTICLES IN LIFE SCIENCE

Milena Kovářová

Keywords: Text mining, article, life science.

Abstract: This article presents advantages that appear by the use of text mining methods in life science.

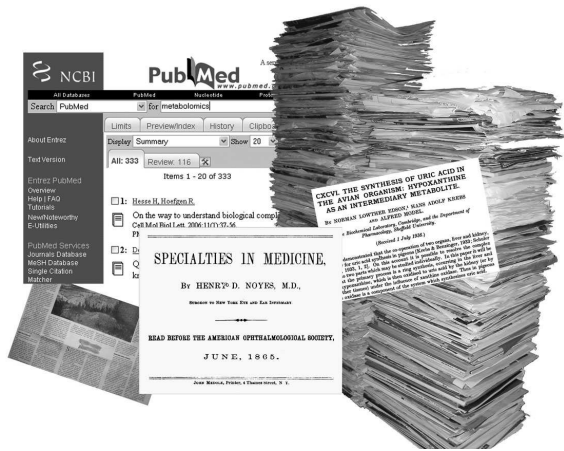
Abstrakt: Tento příspěvek popisuje výhody vznikající použitím text miningu v přírodních vědách.

1 Introduction

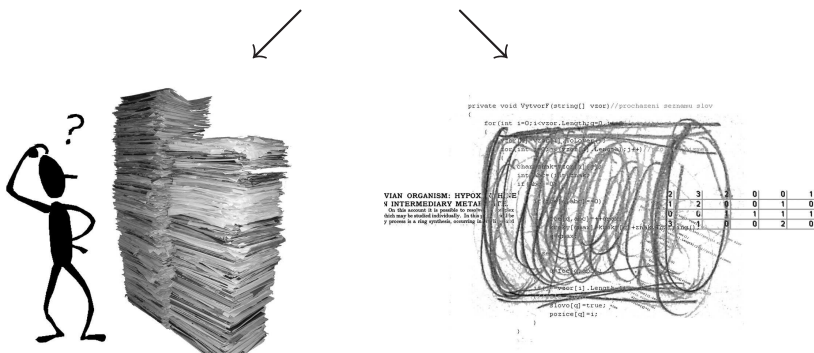
More than 80% of information is stored in textual form. A large number of papers appear every day. It is beyond power of anybody to go through even a part of these published works. A use of search machines yields excessive numbers of results and does not make a problem more simple. There are additional difficulties in biochemical research such as an abundance of terms with ambiguous nomenclature, acronyms and abbreviations. That is the reason for a boom of text mining methods in this area. Text mining methods transfer unstructured textual data into a structured data matrix, which can be evaluated with data mining methods.

2 Text mining

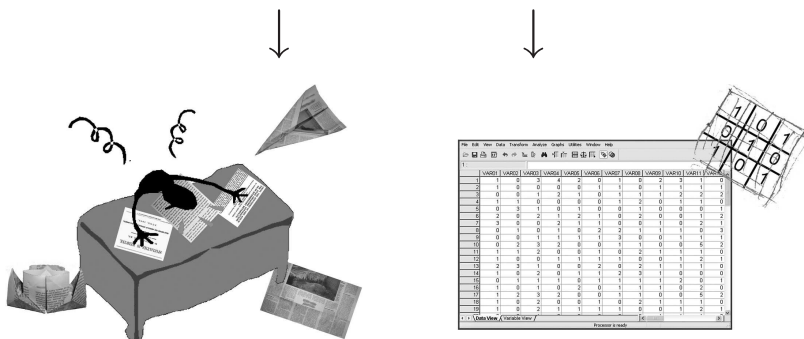
The discovery by computer of new, previously unknown information, by automatically extracting information from different written resources is text mining. A key element is the linking together of the extracted information to form new facts or new hypotheses to be explored further by more conventional means of experimentation. In text mining, the goal is to discover heretofore-unknown information, something that no one yet knows and so could not have yet written down. The most active and promising application area for text mining are biosciences. With biomedical literature increasing at a rate of several thousand papers per week, it is impossible to keep abreast of all developments. Therefore, automated means to manage the information overload are required. Text mining techniques, which involve the processes of information retrieval, information extraction and data mining, provide a way of solving this. By adding meaning to text, these techniques produce a more structured analysis of textual knowledge than simple word searches, and can provide powerful tools for the production and analysis of systems biology models.



Information Retrieval IR



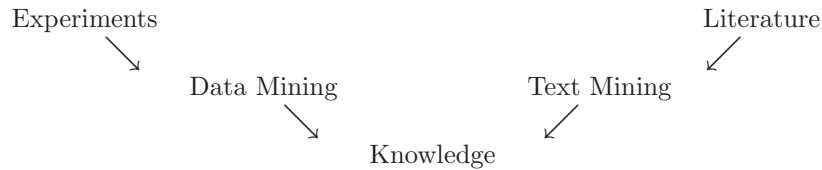
Information Extraction IE



Data Mining DM

3 The investigating of knowledge

To investigate all our knowledge experimentally is much more expensive and time consuming. The mining of literature can give us useful information and can lead to a better conclusion.



Both experiments and literature can produce data from which we can derive new knowledge.

4 Articles

The best known world document collection suitable for text mining is PubMed, the National Library of Medicine's online repository of citation-related information for biomedical research articles. PubMed has received significant attention from computer scientists interested in employing text mining techniques because this online service contains text-based document abstracts for more than 15 million articles on topics in the life sciences. The publication data of Pub Med's collected articles stretch from half of 19 century to the present. The collection is dynamic and grows. Every month is approximately 40,000 new biomedical abstracts added.

5 Example

An example, a potential relationship might be inferred between two proteins P1 and P2 by the pattern of (a) several articles mentioning the protein P1 in relation to the enzyme E1, (b) a few articles describing functional similarities between enzymes E1 and E2 without referring to any protein names, and (c) several articles linking enzyme E2 to protein P2. In this example, the information is not provided by any single document but rather from the totality of the collection. Text mining's methods of pattern analysis seek to discover co-occurrence relationships between concepts as reflected by the totality of the corpus at hand.

6 Conclusion

New knowledge can be investigated also in interdisciplinary research. A structure from theory of information can be recognized for example in theoretical biology, chemistry or life science generally. The text mining program can search for similarities without the dependence on a discipline in order to give us some suggestions.

References

- [1] Ananiadou S., Kell D.B., Tsujii J. (2006). *Trends in biotechnology*, **24**, 571–579.
- [2] Banville D.L. (2006). *Drug discov. today*, **11**, 35–42.
- [3] Feldman R., Sanger J. (2007). *The text mining handbook: advanced approaches in analyzing unstructured data*. Cambridge University Press.

Acknowledgement: The project was financed by project No. MSM6007665808 of the Ministry of Schools, Youth and Sport of the Czech Republic

Address: M. Kovářová, Institute of Physical Biology, University of South Bohemia and Institute of Systems Biology and Ecology ASCR, Academic and University Center, Zámek 136, 373 33 Nové Hradky, Czech Republic

E-mail: kovarova@greentech.cz

ZÁKLADY ŠTATISTIKY S FUZZY DÁTAMI

Pavol Kráľ, Marek Ďurica, Lenka Lašová

Kľúčové slová: Fuzzy množina, fuzzy náhodná premenná, vágne dáta.

Keywords: Fuzzy set, fuzzy random variable, vague data.

Abstrakt: Pod fuzzy dátami rozumieme náhodné vágne dáta vyjadrené prirodzeným jazykom. Cieľom príspevku je čiastočné zhrnutie súčasného stavu štatistiky s fuzzy dátami z pohľadu teórie fuzzy množín, pričom sa zameriavame len na fuzzy náhodnú premennú a popisnú štatistiku.

Abstract: Fuzzy data are random vague data expressed by a natural language. The main aim of our paper is a partial overview of the current state of statistics with fuzzy data from the fuzzy set theory point of view. We focus on the fuzzy random variable and descriptive statistics only.

1 Úvod

Za rok zrodu teórie fuzzy množín je považovaný rok 1965, keď Lotfi Zadeh publikoval svoj článok Fuzzy Sets, aj keď niektoré myšlienky tejto teórie je možné nájsť aj v starších prácach iných autorov. Zjednodušene povedané je základným cieľom tejto teórie najmä vytvorenie matematického aparátu pre popis neurčitosti sveta okolo nás. Táto neurčitosť je dôsledkom vágneho popisu sveta prostredníctvom prirodzeného jazyka. Hoci teória fuzzy množín popisuje kvalitatívne iný druh neurčitosti ako teória pravdepodobnosti, od jej vzniku je zaujímavou otázkou ich vzájomného vzťahu a možnosti využívať ich súčasne pri riešení praktických problémov. Je napríklad známe, že funkcia príslušnosti fuzzy množiny je interpretovateľná pravdepodobnostne, ale aj bez využitia teórie pravdepodobnosti. Vzhľadom na to, že moderná štatistika je úzko previazaná z teóriou pravdepodobnosti, úplne prirodzenou sa stáva otázka, aký je vzťah štatistiky a teórie fuzzy množín. Podľa nášho názoru tento vzťah má dve polohy, pričom motivácia obidvoch súvisí s praktickými aplikáciami teórie fuzzy množín v tzv. soft sciences. Prvou polohou je zovšeobecnenie známych štatistických metód s využitím teórie fuzzy množín. To znamená, že pracujeme s klasickými dátami, ale používaná štatistická metóda obsahuje zapracované fuzzy verzie vzdialeností, podobnosti a pod., často s „fuzzy“ interpretáciou výsledku. Príkladom je napríklad fuzzy zhlučková analýza, ktorá bola rozpracovaná už začiatkom sedemdesiatych rokov minulého storočia. V druhej polohe pracujeme s vágnymi dátami, ktoré sú priamo vyjadrené slovným popisom. Takéto dáta sa prirodzene vyskytujú v tzv. spoločenských vedách a získame ich najčastejšie z dotazníkových prieskumov. Keďže respondent miesto výberu konkrétnej odpovede uvádza svoje preferencie pre každú alternatívu, dochádza k menšej strate informácie a takýto dotazník by mal lepšie popisovať subjektívny názor respondenta.

Pre štatistické spracovanie týchto dát je potom nutné používať odlišné štatistické metódy. Je to jedna z oblastí, na ktoré sa upriamuje pozornosť výskumu v oblasti teórie fuzzy množín v posledných rokoch (pozri [3, 14]). Prezentovaný článok sa venuje práve druhému aspektu vzťahu teórie fuzzy množín a štatistiky. Vzhľadom na obmedzený rozsah sa obmedzíme len na veľmi stručné naznačenie základných koncepcií štatistiky s fuzzy dátami so zameraním na fuzzy náhodnú premennú a popisnú štatistiku.

2 Základné pojmy teórie fuzzy množín

Definícia 2.1. (Zadeh, 1965) *Nech X je základná množina. Fuzzy podmnožina A množiny X je jednoznačne určená svojou funkciou príslušnosti $\mu_A : A \rightarrow [0, 1]$.*

Množinu všetkých fuzzy podmnožín množiny X označujeme $\mathcal{F}(X)$. Miesto fuzzy podmnožina A základnej množiny X hovoríme pri pevne zvolenej základnej množine často zjednodušene o fuzzy množine A . Funkcia príslušnosti je zovšeobecnením pojmu charakteristická funkcia a nazývame ju vertikálnou reprezentáciou fuzzy množiny.

Ďalšími elementárnymi pojmami teórie fuzzy množín sú suport, jadro a α -rez fuzzy množiny.

Definícia 2.2. *Nech X je základná množina. Suportom fuzzy množiny A nazývame množinu $\text{Supp}(A)$, pre ktorú platí*

$$\text{Supp}_A = \{x \in X; \mu_A(x) > 0\}.$$

Definícia 2.3. *Nech X je základná množina. Jadrom fuzzy množiny A nazývame množinu $\text{Core}(A)$, pre ktorú platí*

$$\text{Core}(A) = \{x \in X; \mu_A(x) = 1\}.$$

Definícia 2.4. *Nech X je základná množina a nech $\alpha \in (0, 1)$. Alfa rezom fuzzy množiny A nazývame množinu A_α , pre ktorú platí*

$$A_\alpha = \{x \in X; \mu_A(x) \geq \alpha\}.$$

α -rezy predstavujú tzv. horizontálnu reprezentáciu fuzzy množín.

Na popis neurčitosti v reálnych dátach môžeme použiť špeciálny typ fuzzy množín, fuzzy intervaly. Je ich možné definovať nasledujúcim spôsobom:

Definícia 2.5. *Fuzzy interval A je fuzzy podmnožina množiny \mathbb{R} , ktorá spĺňa nasledujúce podmienky*

- je normálna, t. j. existuje $x \in \mathbb{R}$, pre ktoré $\mu_A(x) = 1$,
- suport je ohraničená množina,
- každý α -rez je uzavretý interval.

Fuzzy intervaly, ktorých množina $\{x \in \mathbb{R}; \text{pre ktoré } \mu_A(x) = 1\}$ má jediný prvok, nazývame fuzzy čísla. Fuzzy čísla sú zovšeobecnením reálnych čísel a môžeme pre ne definovať aritmetické operácie na základe Zadehovo princípu neurčitosti s využitím t -normy.

Definícia 2.6. (Schweizer, Sklar, 1958) Funkcia $T : [0, 1]^2 \rightarrow [0, 1]$, ktorá pre všetky $x, y, z \in [0, 1]$ spĺňa nasledujúce podmienky:

- $T(x, y) = T(y, x)$,
- $T(x, T(y, z)) = T(T(x, y), z)$,
- $y \leq z \Rightarrow T(x, y) \leq T(x, z)$,
- $T(x, 1) = x$,

sa nazýva t -norma.

Definícia 2.7. (Zadehov princíp rozšírenia) Nech A, B sú ľubovoľné fuzzy čísla. Aritmetickú operáciu $*_T$ na množine všetkých fuzzy čísel definujeme nasledujúcim spôsobom

$$\mu_{A*_TB}(z) = \sup\{T(\mu_A(x), \mu_B(y)); z = x * y, x \in \mathbb{R}, y \in \mathbb{R}\},$$

kde $*$ je aritmetická operácia definovaná na množine reálnych čísel a T je t -norma.

Predchádzajúce definície je možné jednoducho rozšíriť pre \mathbb{R}^n .

3 Fuzzy náhodná premenná

V tejto časti sa zameriame na pojem fuzzy náhodná premenná, ktorý je nevyhnutne potrebný pre budovanie príslušnej štatistickej teórie. Prístupy k fuzzy náhodnej premennej môžeme rozdeliť do dvoch skupín. Prvú skupinu predstavujú koncepcie, ktoré fuzzy náhodnú premennú viažu na podkladovú klasickú náhodnú premennú, t. j. hodnoty fuzzy náhodnej premennej predstavujú fuzzy popis reálnych hodnôt výsledkov experimentu. Tento prístup je prezentovaný v prácach [9, 12, 13]. V druhej skupine sú prístupy, ktoré každému experimentálnemu výsledku priradujú priamo fuzzy hodnotu. Tento prístup je prezentovaný v práci [16]. Vzťah medzi týmito koncepciami vyjadruje napríklad nasledujúca veta.

Veta 3.1. [6] Uvažujme pravdepodobnostný priestor (Ω, \mathcal{A}, P) a zobrazenie $\chi : \Omega \rightarrow \mathcal{F}_c(\mathbb{R})$, kde $\mathcal{F}_c(\mathbb{R})$ je množina všetkých konvexných podmnožín (\mathbb{R}) . Potom χ je fuzzy náhodná premenná v zmysle [9, 12, 13] práve vtedy, keď je fuzzy náhodnou premennou aj v zmysle [16].

V ďalšom texte sa zameriame na prístup v [16].

Definícia 3.1. Uvažujme pravdepodobnostný priestor (Ω, \mathcal{A}, P) . Zobrazenie $\chi : \Omega \rightarrow \mathcal{F}(\mathbb{R}^n)$ sa nazýva fuzzy náhodná premenná (niekedy aj náhodná fuzzy množina), ak pre každé $\alpha \in \langle 0, 1 \rangle$ zobrazenie $\chi_\alpha : \Omega \rightarrow \mathcal{K}(\mathbb{R}^n)$ je také, že pre každé $\omega \in \Omega$

$$\chi_\alpha(\omega) = (\chi(\omega))_\alpha$$

je náhodná množina.

Príklad 3.1. Uvažujme náhodný experiment, v ktorom náhodne vyberáme študenta zo skupiny študentov, ktorí práve absolvovali test z predmetu X . Uvažujme náhodnú premennú $\chi =$ klasifikácia náhodne vybraného študenta vzhľadom na jeho vnímanie náročnosti absolvovaného testu. V tomto príklade neexistuje prototyp jednotky náročnosti testu. Jednotlivé stupne vnímania náročnosti testu vieme porovnávať len vzájomne, t.j. neexistuje podkladová klasická náhodná premenná. Ak využijeme prístup z [16], môžeme túto situáciu modelovať nasledujúcim spôsobom. Pravdepodobnostný priestor je (Ω, \mathcal{A}, P) , kde $\Omega = \{\text{študenti účastníci sa testu}\}$, $\mathcal{A} = \mathcal{P}(\Omega)$. Študenta vyberáme náhodne, pričom uvažujeme rovnakú pravdepodobnosť výberu pre každého študenta, preto $P(\omega) = \frac{1}{\text{kard}(\Omega)}$ pre každé $\omega \in \Omega$. Klasifikáciu študentov vzhľadom na ich vnímanie obtiažnosti testu môžeme považovať za fuzzy náhodnú premennú χ asociovanú s (Ω, \mathcal{A}, P) , ktorá môže nadobúdať napríklad hodnoty: x_1 - veľmi veľká náročnosť, x_2 - veľká náročnosť, x_3 - stredná náročnosť, x_4 - nižšia náročnosť, x_5 - veľmi nízka náročnosť. Tieto hodnoty môžeme modelovať napríklad pomocou lineárnych fuzzy čísel (pozri [2]) so suportom $\langle 0\%, 100\% \rangle$.

Zaujímavou otázkou samozrejme je, akým spôsobom sú zavedené pojmy s pojmom náhodná premenná úzko súvisiace. My sa obmedzíme len na strednú hodnotu, rozptyl, kovarianciu a koreláciu. Najprv si pripomenieme definície Hausdorffovej metriky a Steinerovho bodu, ktoré budeme v ďalšom texte využívať. Hausdorffova metrika je základnou metrikou používanou na meranie vzdialeností dvoch fuzzy množín priestore. Steinerov bod slúži ako referenčný bod pre lokalizáciu priestorového útvaru (v našom prípade fuzzy množiny). Na rozdiel od ťažiska nedochádza k jeho zmene pri rovnomernom raste útvaru vo všetkých smeroch.

Hausdorffova metrika na priestore $\mathcal{K}(\mathbb{R}^P)$: pre každé $A, B \in \mathcal{K}(\mathbb{R}^P)$

$$d_H(A, B) = \max \left\{ \sup_{a \in A} \inf_{b \in B} |a - b|, \sup_{b \in B} \inf_{a \in A} |a - b| \right\}.$$

Pre Steinerovho bod fuzzy množiny A , budeme ho označovať σ_A , platí

$$\sigma_A = n \cdot \int_0^1 \int_{S_{n-1}} u \cdot s_A(u) \mu(du) d\alpha.$$

kde μ je normovaná Lebesguova miera na jednotkovom kruhu

$$S^{n-1}(\mu(S^{n-1})) = 1.$$

Definícia 3.2. [16] Máme daný pravdepodobnostný priestor (Ω, \mathcal{A}, P) . Nech $\chi : \Omega \rightarrow \mathcal{F}(\mathbb{R}^P)$ je fuzzy náhodná premenná, ktorá je integrovateľná v zmysle $d_H(\{0\}, \chi_0) \in L^1(\Omega, \mathcal{A}, P)$, potom fuzzy náhodná stredná hodnota je taká hodnota $E(\chi|P) \in F(\mathbb{R}^P)$, že pre všetky $\alpha \in [0, 1]$

$$(E(\chi|P))_\alpha = \{E(\chi|P) | X : \Omega \rightarrow \mathbb{R}^P, X \in (\Omega, \mathcal{A}, P), X \in \chi_\alpha\} \\ = \text{Aumannov integrál náhodnej množiny } \chi_\alpha.$$

Príklad 3.2. [6] Majme pravdepodobnostný priestor (Ω, \mathcal{A}, P) a fuzzy náhodnú premennú $\chi : \Omega \rightarrow \mathcal{F}(\mathbb{R}^P)$ ktorej hodnoty sú zrezané ihlany s podstavami ležiacimi v rovine tvorenej osami x a y a centrovanými v bode $(0, 0, 0)$. Súradnica z bodu, ktorým prechádza rezová rovina rovnobežná s podstavou ihlana je reálna náhodná premenná $\beta : \Omega \rightarrow \mathbb{R}$ majúca rovnomerné rozdelenie $\mathcal{U}_{[0,1]}$ a polomer je tiež reálna funkcia $\rho : \Omega \rightarrow \mathbb{R}$ majúca χ^2 rozdelenie s jedným stupňom voľnosti.

V tomto prípade, hodnota fuzzy náhodnej premennej v bode ω je daná

$$\chi(\omega)(x, y) = \begin{cases} 0, & \text{ak } x^2 + y^2 > [\rho(\omega)]^2; \\ \gamma(\omega), & \text{ak } x^2 + y^2 = [\gamma(\omega)]^2, \\ & \text{kde } \rho(\omega)(1 - \beta(\omega)) \leq \gamma(\omega) \leq \rho(\omega); \\ \beta(\omega), & \text{ak } 0 < x^2 + y^2 < [\rho(\omega)]^2 [1 - \beta(\omega)]^2; \\ 1, & \text{ak } 0 = x^2 + y^2. \end{cases}$$

a teda α -rezy χ_α sú dané

$$\chi_\alpha(\omega)(x, y) = \\ = \begin{cases} \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 \leq [\rho(\omega)]^2 (1 - \alpha)^2\} & \text{ak } 0 \leq \alpha \leq \beta(\omega), \\ \{(0, 0)\} & \text{ak } \alpha > \beta(\omega). \end{cases}$$

Ak teraz vypočítame Aumannovým integrálom strednú hodnotu α -rezu, dostaneme kruh so stredom $(0, 0)$ a polomerom $1 - \alpha$.

V literatúre sa často uvádzajú ďalšie dve definície strednej hodnoty

1. definícia odvodená pomocou Bochnerovej strednej hodnoty príslušnej suport funkcie fuzzy náhodnej premennej X (pozri [17]),
2. stredná hodnota fuzzy náhodnej premennej definovaná v metrickom priestore normálnych kompaktných konvexných fuzzy množín s metriku d_2 (definovaná pomocou L_2 metriky na priestore Lebesgueovskych merateľných funkcií) je uvedená v práci [11].

Všetky tri definície sú rovnocenné, čo je dokázané v práci [11].

Definícia 3.3. [11] Máme daný pravdepodobnostný priestor (Ω, \mathcal{A}, P) . Nech $\chi : \Omega \rightarrow \mathcal{F}(\mathbb{R}^P)$ je fuzzy náhodná premenná, ktorá je integrovateľná v zmysle $d_H(\{0\}, \chi_0) \in L^1(\Omega, \mathcal{A}, P)$ a označme μ jej strednú hodnotu. Potom fuzzy rozptyl náhodnej premennej χ definujeme

$$\text{var}(\chi|P) = E\left((\chi - \mu)^2 | P\right).$$

V nasledujúcej vete uvedieme niekoľko vlastností strednej hodnoty a rozptylu fuzzy náhodnej premennej.

Veta 3.2. [11] *Majme fuzzy náhodné premenné X, Y . Nech x je integrovateľná náhodná premenná, $A \in \mathcal{F}_C$ (množina všetkých normálnych kompaktných fuzzy podmnožín v \mathbb{R}^P) a λ, γ sú reálne čísla.*

1. $E(\lambda X + \gamma Y) = \lambda EX + \gamma EY$,
2. $\text{var}(X) = E\|X\|_2^2 - \|EX\|_2^2$,
3. $\text{var}(\lambda X) = \lambda^2 \text{var}(X)$,
4. $\text{var}(xA) = \|A\|_2^2 \cdot \text{var}(x)$,
5. $\text{var}(A + X) = \text{var}(X)$,
6. $\text{var}(xX) = E\|X\|_2^2 \cdot \text{var}(x) + Ex^2 \cdot \text{var}(X)$, ak x, X sú stochasticky nezávislé,
7. $\text{var}(X + Y) = \text{var}(X) + \text{var}(Y)$, ak X, Y sú nezávislé fuzzy náhodné premenné.

Kovarianciu medzi dvoma náhodnými premennými budeme definovať ako skalárny súčin ich suportov nasledujúcim spôsobom (Körner v [11]).

$$\text{cov}(X, Y) = E\langle s_X - s_{EX}, s_Y - s_{EY} \rangle,$$

kde skalárnym súčinom je

$$\langle s_A, s_B \rangle = \langle A, B \rangle = n \cdot \int_0^1 \int_{S^{n-1}} s_A(\alpha, u) \cdot s_B(\alpha, u) \mu(du) d\alpha, A, B \in \mathcal{F}_C.$$

Potom korelačný koeficient medzi dvoma náhodnými premennými definujeme

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X) \text{var}(Y)}}.$$

Pre kovarianciu a korelačný koeficient platia vlastnosti uvedené v nasledujúcej vete.

Veta 3.3. [11] *Nech X, Y je dvojica fuzzy náhodných premenných. Nech $\sigma_X(\sigma_Y)$ je Steinerov bod pre fuzzy náhodnú premennú X (Y). Nech $X_0 = X - \sigma_X$ a $Y_0 = Y - \sigma_Y$. Potom*

1. $\text{cov}(X, Y) = \text{cov}(X_0, Y_0) + \text{cov}(\sigma_X, \sigma_Y)$,
2. $\text{var}(X) = \text{cov}(X, X)$,

3. $\text{cov}(X, Y) = 0$, ak X, Y sú nezávislé,
4. $|\rho(X, Y)| \leq 1$,
5. $\text{cov}(X, Y) = E \langle X, Y \rangle - \langle EX, EY \rangle$.

Z hľadiska praktického použitia i z hľadiska teoretického sú dôležité analógie slabého a silného zákona veľkých čísel a centrálne limitné vety. Tieto boli skúmané pri použití rôznych metrík. Časť týchto výsledkov má síce hlavne teoretický charakter, ale na základe niektorých z nich boli vytvorené napríklad asymptotické procedúry na testovanie hypotéz o fuzzy strednej hodnote fuzzy náhodnej premennej (pozri napríklad [10]).

4 Popisná štatistika

V tejto časti sa zameriame na charakteristiky fuzzy dát, ktoré budú zovšeobecnením klasických pojmov aritmetický priemer, modus a medián. V prípade, že meraním získame priamo n fuzzy kvantít (fuzzy čísel) $f x_1, \dots, f x_n$, môžeme pre výpočet fuzzy aritmetického priemeru $f \bar{x}$ použiť nasledujúci analógiu klasického vzťahu, kde operácie sčítania a delenia sú definované na základe princípu rozšírenia. Výsledkom je opäť fuzzy kvantita. Pri praktickom použití sa ako vstup používajú najmä trojuholníkové fuzzy čísla a operácie sú založené na vhodnej t -norme, aby ostala zachovaná linearita fuzzy čísel (pozri [2]). Analogicky je možné definovať použitím vzťahov klasickej popisnej štatistiky modus a medián. V prípade kvalitatívnych dát, ktoré získavame napríklad pri spracovaní dotazníkov, je ale problematické získať priamo fuzzy kvantitu. V takom prípade zvyčajne získame len subjektívnu preferenciu k možným hodnotám fuzzy náhodnej premennej. Priemer a modus môžeme potom definovať napríklad nasledujúcim spôsobom (pozri [14]).

Definícia 4.1. *Nech X je základná množina, nech $L = \{L_1, L_2, \dots, L_k\}$ je množina lingvistických faktorov. Nech*

$$\left\{ F x_i = \left\{ \frac{\mu_{iL_1}}{L_1}, \frac{\mu_{iL_2}}{L_2}, \dots, \frac{\mu_{iL_n}}{L_k} \right\}, i = 1, 2, \dots, n \right\}$$

je realizácia náhodného fuzzy výberu, $\mu_{ij} (\sum_{j=1}^k \mu_{ij} = 1)$ je príslušnosť x_i vzhľadom na L_j . Fuzzy výberový priemer je definovaný ako

$$F \bar{x} = \left\{ \frac{\frac{1}{n} \sum_{i=1}^n \mu_{iL_1}}{L_1}, \frac{\frac{1}{n} \sum_{i=1}^n \mu_{iL_2}}{L_2}, \dots, \frac{\frac{1}{n} \sum_{i=1}^n \mu_{iL_n}}{L_n} \right\}.$$

Nech $S_j = \sum_{i=1}^n \mu_{ij}, j = 1, 2, \dots, k$. Fuzzy modus je definovaný nasledujúcim spôsobom

$$F \hat{x} = \left\{ L_j; S_j = \max_{1 \leq i \leq k} S_i \right\}.$$

Příklad 4.1. (Záujem o výrobok) L_1 : veľmi nízky = 1, L_2 : nízky = 2, L_3 : stredný = 3, L_4 : veľký = 4, L_5 : veľmi veľký = 5

	1	2	3	4	5
Fx_1	0,7	0,3	0	0	0
Fx_2	0	0,9	0,1	0	0
Fx_3	0	0	0,7	0,3	0
Fx_4	0	0	0	0,8	0,2
Fx_5	0	0	0	0,2	0,8

$$F\bar{x} = \left\{ \frac{0,14}{1}, \frac{0,24}{2}, \frac{0,16}{3}, \frac{0,26}{4}, \frac{0,2}{5} \right\}$$

$$S_1 = 0,7, S_2 = 1,2, S_3 = 0,8, S_4 = 1,3, S_5 = 1,$$

$$F\hat{x} = 4(\text{ veľký záujem })$$

5 Záver

V článku sme sa pokúsili veľmi stručne naznačiť základnú koncepciu fuzzy náhodnej premennej a popisnej štatistiky s fuzzy dátami. Zvolili sme ukážky práve z týchto dvoch oblastí, pretože koncepcia fuzzy náhodnej premennej predstavuje teoretický základ, bez ktorého by nebolo možné budovať ucelenú štatistickú teóriu pre fuzzy dáta a popisná štatistika reprezentuje základný krok pri snahe o praktickú aplikáciu tejto teórie. Hoci sme pre obmedzený priestor úplne opomenuli výberové skúmanie a regresnú analýzu, článok by mal umožniť čitateľovi získať pomerne jasnú elementárnu predstavu o možnostiach prezentovanej teórie. V prípade záujmu môže čitateľ nájsť úplnejší prehľad napríklad v prácach [6, 14]. Za najväčšiu slabinu prezentovanej teórie je možné označiť to, že hoci bol zavedený pojem normality pre fuzzy náhodnú premennú (pozri [15]), z praktického hľadiska nepredstavuje úplnú analógiu normálneho rozdelenia v klasickom prípade. Ďalšou nevýhodou je, že je možné pri definíciách využiť rôzne metriky, Zadehov princíp rozšírenia založený na rôznych t-normách a pod., čím vzniká množstvo teoretických výsledkov, ale výrazne to sťažuje reálnu aplikáciu teórie.

V súčasnosti sa pozornosť v tejto oblasti zameriava práve na konštrukciu testovacích techník pre testy (fuzzy) hypotéz s fuzzy dátami, regresnú analýzu, techniky bootstrapu, ale naďalej sa venuje pozornosť i teoretickým aspektom fuzzy náhodnej premennej a fuzzy náhodných množín. Vykonávajú sa aj simulačné štúdie s účelom porovnať jednotlivé techniky, empiricky overiť niektoré predpoklady a výsledky teórie fuzzy náhodnej premennej. Skúmajú sa aj simulačné techniky na simulovanie fuzzy náhodných premenných pre potreby modelovania reálnych situácií. V oblasti aplikácií je predmetom záujmu napríklad konštrukcia rôznych kontrolných systémov, systémov na podporu rozhodovania a oceňovacích modelov, v ktorých sa prirodzeným spôsobom

súčasne vyskytuje neurčitost' i náhodnosť. Podrobnosti môže nájsť čitateľ napríklad v prácach [1, 4, 5, 7, 8, 18].

Predpokladáme, že ďalší výskum bude pokračovať vo všetkých vyššie uvedených oblastiach so zameraním na viacrozmerne dáta. Predpokladáme aj intenzívny teoretický i aplikačný výskum možného rozšírenia uvedenej teórie na zovšeobecnenia fuzzy množín, najmä intervalovo ohodnotené fuzzy množiny (v tomto prípade stupeň príslušnosti nie je určený jedným reálnym číslom, ale intervalom) a intuicionistické fuzzy množiny v zmysle Atanassova (v tomto prípade okrem funkcie príslušnosti máme aj funkciu nepríslušnosti). Okrem toho sa v posledných rokoch venuje stále väčšia pozornosť aj skúmaniu tzv. neurčitých pravdepodobností.

Autori by sa v ďalších prácach chceli venovať z teoretického, ale najmä z praktického hľadiska štatistickému spracovaniu fuzzy dát ekonomického charakteru s dôrazom na komparatívnu analýzu predností a nedostatkov možných prístupov i možné použitie intervalovo ohodnotených fuzzy množín.

Literatúra

- [1] Bag S., Chakraborty D., Roy A. R. (2008). *A production inventory model with fuzzy random demand and with flexibility and reliability considerations*, Computers & Industrial Engineering, In Press.
- [2] Bojadziev G., Bojadziev M. (2007). *Fuzzy logic for business, finance, and management*. Advances in Fuzzy Systems: Applications and Theory **23**, World Scientific.
- [3] Buckley J.J. (2006). *Fuzzy probability and statistics*, Springer.
- [4] Colubi A. (2009). *Statistical inference about the means of fuzzy random variables: Applications to the analysis of fuzzy- and real-valued data*, Fuzzy Sets and Systems, **160**, 344–356.
- [5] Fu Ke-Ang, Zhang Li-Xin (2008). *Strong laws of large numbers for arrays of rowwise independent random compact sets and fuzzy random sets*, Fuzzy Sets and Systems, **159**, 3360–3368.
- [6] Gil M.Á., López-Díaz M., Ralescu D.A. (2006). *Overview on the development of fuzzy random variables*. Fuzzy Sets and Systems **157**, 2546–2557.
- [7] González-Rodríguez G., Blanco Á., Colubi A., Asunción Lubiano M. (2009). *Estimation of a simple linear regression model for fuzzy random variables*, Fuzzy Sets and Systems, **160**, 357–370.
- [8] González-Rodríguez G., Colubi A., Trutschnig W. (2008). *Simulation of fuzzy random variables*, Information Sciences, In Press.
- [9] Kruse R., Meyer K.D. (1987). *Statistics with vague data*, Reidel Publ. Co. Dordrecht.
- [10] Körner R. (2000). *An asymptotic α -test for the expectation of random fuzzy variables*, J. Statist. Plann. Inference **83**, 331–346.

- [11] Körner R. (1997). *On the variance of fuzzy random variables*, Fuzzy Sets and Systems **92**, 83–93.
- [12] Kwakernaak H.(1978). *Fuzzy random variables. Part I: definitions and theorems*. Inform. Sci. **15**, 1 – 29.
- [13] Kwakernaak H.(1979). *Fuzzy random variables. Part II: algorithms and examples for the discrete case*. Inform. Sci. **17**, 253 – 278.
- [14] Nguyen H.T., Wu B. (2006). *Fundamentals of statistics with fuzzy data*, Springer.
- [15] Puri M.L., Ralescu D.A. (1985). *Fuzzy random variables*. J. Math. Anal. Appl. **114**, 409–422.
- [16] Puri M.L., Ralescu, D.A. (1986). *The concept of normality for fuzzy random variables*. Ann. Probab. **11**, 1373–1379.
- [17] Vitale R.A.(1988). *An alternate formulation of mean value for random geometric figures*, J. Microscopy **151**, 197–204.
- [18] Xu Jiuping, Liu Yonggang (2008). *Multi-objective decision making model under fuzzy random environment and its application to inventory problems*, Information Sciences, **178**, 2899–2914.

Podakovanie: Táto práca bola podporená grantom KEGA 3/5214/07.

Adresa: P. Kráľ, EF UMB, Tajovského 10, 975 90 Banská Bystrica, Ústav vedy a výskumu UMB, Cesta na amfiteáter 1, 974 01 Banská Bystrica
M. Ďurica, FPEDAS ŽU, Univerzitná 1, 010 26 Žilina
M. Ďurica, L. Lašová FPV UMB, Tajovského 40, 974 01 Banská Bystrica

E-mail: pavol.kral@umb.sk, durica@fpv.umb.sk, lasova@fpv.umb.sk

TESTING HYPOTHESES ABOUT THE PARAMETERS OF THE SHIFTED AND RESCALED WIENER PROCESS

Andrea Kvitkovičová

Keywords: Wiener process, constant drift rate and instantaneous variance, uniformly most powerful tests, unbiased tests, Rényi divergences.

Abstract: The paper deals with testing hypotheses about the parameters a and b^2 of the random process $\{bw_t + at; t \geq 0\}$, where $\{w_t; t \geq 0\}$ is the Wiener process, $a \geq 0$ and $b^2 > 0$. Tests are based on the first time when the process reaches a pre-specified positive boundary. We focus on tests about the drift rate a and on testing both parameters simultaneously.

Abstrakt: V príspevku sa zaoberáme testovaním hypotéz o parametroch a a b^2 náhodného procesu $\{bw_t + at; t \geq 0\}$, kde $\{w_t; t \geq 0\}$ je Wienerov proces, $a \geq 0$ a $b^2 > 0$. Testy sú založené na čase, kedy proces prvýkrát dosiahne vopred zvolenú kladnú hranicu. Dôraz kladieme na testy hypotéz o smernici a a na testy hypotéz o oboch parametroch súčasne.

1 Introduction

Random processes can be used to describe development of various characteristics in many practical situations. Even when these characteristics do change continually over time, information, which we can use, is often limited. For example, the process can only be observed in a set of discrete time points. Or, we can only observe the time points when the process reaches a pre-specified set of values.

In many situations the second approach follows naturally from the type of the problem. As an example, let us consider a process, which is only observable with difficulties or not at all, but when it reaches a given value, we observe a certain event. In what follows we will focus on this setting.

We will assume that the underlying process is a Wiener process with a constant non-negative drift rate and instantaneous variance. The information available will be the first time when the process reaches a pre-specified positive boundary. This can be used as a starting model in various fields, such as medicine, biology, environmental sciences, social sciences and others (see [4] for examples).

Using only the first hitting time, we will test hypotheses about the drift rate and the instantaneous variance of the process. Working with limited information usually decreases quality of results. It is therefore important to ask whether the available information is sufficient to answer the questions of interest in a satisfactory way. We will study the power of the constructed tests to address this question.

We will start with a more detailed problem formulation in the following section. Later, in Section 3 we will test hypotheses about the drift rate, and in Section 4 we will study several ways of testing both parameters simultaneously. Finally, we will summarize the results in Section 5.

2 Problem formulation and first considerations

Let us have a random process $Y_t = bw_t + at$, where $\{w_t; t \geq 0\}$ is a Wiener process, $a \geq 0$ is a drift rate and $b^2 > 0$ is an instantaneous variance. Instead of the whole trajectory of the process, let us observe the first time when the process reaches a boundary point $A \geq 0$. Thus, instead of $\{Y_t; t \geq 0\}$ we observe $X = \inf\{t \geq 0; Y_t \geq A\}$.

The loss of information leads to considering a and b^2 parameters of a random variable, instead of a random process. It has been shown (see for example [1]) that the first hitting time X is a continuous random variable with the density $f_{A,a,b}(x)$ with respect to the Lebesgue measure, where

$$f_{A,a,b}(x) = \sqrt{\frac{A^2}{2\pi b^2}} \frac{1}{x^{\frac{3}{2}}} \exp\left\{-\frac{1}{2xb^2}(A-ax)^2\right\} \mathbb{I}\{x > 0\} \quad \text{for all } x \in \mathbb{R}.$$

We can notice that the distribution of X given by the density $f_{A,a,b}(x) = f_{\theta}(x)$ constitutes an exponential family with the natural parametrization $\theta = (\theta_1, \theta_2)^\top$ with $\theta_1 = \frac{1}{b^2}$ and $\theta_2 = \frac{a^2}{b^2}$. The natural parameter space is $\Theta = (0, \infty) \times [0, \infty)$.

For a positive drift rate we have an inverse Gaussian distribution, whereas in case of no drift we obtain an inverse gamma distribution. Let us realize that a non-negative drift rate guarantees that the first hitting time is a.s. finite. However, its expected value is infinite if $a = 0$. For a positive drift the expectation is finite decreasing function of the drift rate.

Considering the nature of experiment, we can expect that higher values of a with b^2 kept fixed lead to higher probability of reaching the boundary early. The same holds for higher values of b^2 with a kept fixed. Density $f_{A,a,b}(x)$ indeed has one peak, which is reached earlier for larger a or b^2 . However, when both parameters can be varied at the same time, their influence on the first hitting time is mixed. As a result, if both parameters are unknown, more observations will be needed to derive conclusions about a and/or b^2 .

For $n > 1$ independent observations of the first hitting time, the maximum likelihood estimators of the parameters a and b^2 are given by $\hat{a} = \frac{A}{\bar{x}}$ and $\hat{b}^2 = \frac{A^2}{n} \sum_{j=1}^n \left(\frac{1}{X_j} - \frac{1}{\bar{x}}\right)$, where $\bar{x} = \frac{1}{n} \sum_{j=1}^n X_j$. It can be shown (see for example [1]) that these are consistent estimators, provided that $a > 0$.

Independent observations X_1, \dots, X_n can be obtained by observing n independent realizations of the process Y_t and the corresponding first hitting times of A . Alternatively, we could observe the first times $Z_A, Z_{2A}, \dots, Z_{nA}$, when a single realization of the process reaches the boundaries $A, 2A, \dots, nA$, respectively. Let us consider random variables Z_1, \dots, Z_n such that $Z_1 = Z_A$

and $Z_i = Z_{iA} - Z_{(i-1)A}$ for $2 \leq i \leq n$. It follows from the strong Markov property of the Wiener process that $\mathcal{L}((X_1, \dots, X_n)^\top) = \mathcal{L}((Z_1, \dots, Z_n)^\top)$ (see [1] for details of the proof). This gives us an opportunity to conduct an experiment in two different ways and to use the observations from any of the settings in a unified testing approach discussed below.

3 Tests about the drift rate

Techniques for testing hypotheses about one parameter component in a distribution coming from an exponential family are well-developed. We have employed strategies discussed in [2] to construct uniformly most powerful or uniformly most powerful unbiased tests about a and about b^2 . Here we will shortly summarize the results for the drift rate. The derivation of results as well as tests about the instantaneous variance can be found in [1].

Before we start, let us realize a special importance of hypotheses assuming $a = 0$, since these lead to a Wiener process multiplied by a constant.

In what follows, we will assume that we have observations X_1, \dots, X_n obtained in one of the ways discussed at the end of the previous section.

3.1 Tests assuming a known variance

Far easiest case emerges when we can assume that the variance b^2 is known. Under this assumption we can construct a uniformly most powerful test of the hypothesis $H_1 : a \leq a_0$ against the alternative $K_1 : a > a_0$ for a given value $a_0 \geq 0$.

The testing function is given by $\psi(\mathbf{x}) = \mathbb{I} \left\{ \sum_{j=1}^n x_j \leq C \right\}$ with a constant C satisfying $\alpha = \exp \left\{ \frac{2nAa_0}{b^2} \right\} \Phi \left(\frac{-nA - a_0C}{b\sqrt{C}} \right) + \Phi \left(\frac{-nA + a_0C}{b\sqrt{C}} \right)$. The power at any $a > a_0$ is given by the power function $\beta_{A, a_0, b, n}(a) = \Phi \left(\frac{-nA + aC}{b\sqrt{C}} \right) + \exp \left\{ \frac{2nAa}{b^2} \right\} \Phi \left(\frac{-nA - aC}{b\sqrt{C}} \right)$.

Let us consider $X_i = Z_{iA} - Z_{(i-1)A}$ as described above. We can notice that to test the hypothesis it suffices to observe the hitting time Z_{nA} of the last point only. The hypothesis is rejected, if the last point is reached “too early”. Let us further realize that if the boundary is not reached until a certain time, we will not reject the hypothesis. Thus, to draw a conclusion we even do not have to wait until the boundary is reached. Based on similar considerations we can construct tests using more limited information on the number of boundaries reached until time T (see [1] for details).

Power $\beta_{A, a_0, b, n}(a)$ is shown as a function of a and b^2 in the left panel of Figure 1. It can be easily shown that it is an increasing function of the drift rate a . For the special case when $a_0 = 0$ we can even show that it is an increasing function of $\frac{nAa}{b^2}$. This gives an insight into the way how the power depends on each of the parameters a and b^2 , and A and n , when we keep the remaining parameters fixed. Further, it can be shown that for sufficiently large values of $\frac{nAa}{b^2}$ the power is arbitrarily close to one.

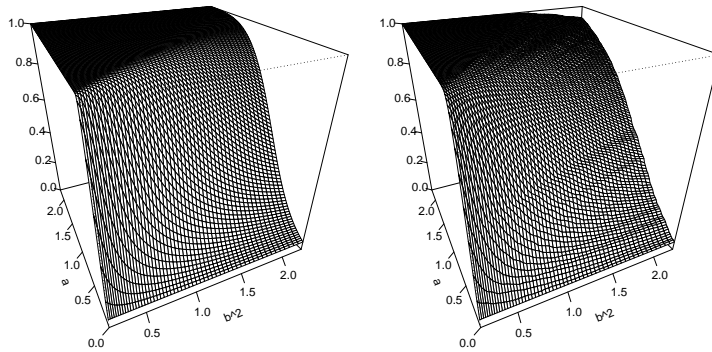


Figure 1: Left panel: power of a uniformly most powerful test of $H_1 : a = 0$ against $K_1 : a > 0$ assuming a known value of b^2 . Right panel: power of a uniformly most powerful unbiased test of $H_2 : a = 0$ against $K_2 : a > 0$ assuming an unknown value of b^2 . Both are shown as functions of a and b^2 with $A = 1$, $n = 10$ and $\alpha = 0.05$ fixed.

The simple setting of a known variance gives us opportunities to optimize the testing design with respect to time, costs and other aspects of the experiment (see [1] for more details).

3.2 Tests assuming an unknown variance

If we consider an unknown instantaneous variance b^2 , the situation becomes less insightful. With the variance being a nuisance parameter, we can only construct a uniformly most powerful unbiased test of the hypothesis $H_2 : a = 0$ against the alternative $K_2 : a > 0$. Developing a test of a more general hypothesis H_1 considered in the previous paragraph is not straightforward. This is due to the form of the natural parametrization, where a and b^2 are combined into one parameter component θ_2 .

The testing function of H_2 against K_2 is given by $\psi(\mathbf{x}) = \mathbb{I} \left\{ \left(\sum_{j=1}^n x_j \right) \times \left(\sum_{j=1}^n \frac{1}{x_j} \right) \leq n^2 \left(\frac{n-1}{F_{1, n-1, 1-\alpha}} + 1 \right) \right\}$. We can notice that it now takes into account both unknown drift rate via considering $\sum_{j=1}^n x_j$ and unknown variability via considering $\sum_{j=1}^n \frac{1}{x_j}$.

Let us realize that we cannot simplify the experiment to observing only until a pre-specified time point. This is rather inconvenient, since the expected experiment duration under the null hypothesis is infinite.

Further, it is not straightforward to study the power dependence on the parameters. For concrete values of A and n it is shown in the right panel of Figure 1 as a function of a and b^2 . In general, the power at any

$a > 0$ can be calculated as $\int_{F_{1, n-1, 1-\alpha}}^{\infty} \int_0^{\infty} \frac{y^{\frac{n}{2}-1} v^{-\frac{1}{2}}}{\Gamma(\frac{1}{2}) \Gamma(\frac{n-1}{2}) 2^{\frac{n}{2}} (n-1)^{\frac{1}{2}}} \times$
 $\times \exp\left\{-\frac{(b^2vy - n(n-1)aA)^2}{2(n-1)b^4vy} - \frac{y}{2}\right\} dy dv.$

4 Testing both parameters simultaneously

When testing hypotheses about a multidimensional parameter, we can choose from a variety of techniques. Perhaps one of the easiest approaches is to consider a concrete alternative of interest. We can then apply the classical Neyman-Pearson theory. The resulting most powerful test can be expected to reach high power against a set of simple alternatives in the direction of the original simple alternative. Even though this can be used in some situations, we will not study this approach here.

In what follows, we will focus on tests which can be used in an asymptotic way. We will briefly explore opportunities offered by a class of tests based on the Rényi divergences. These have already been applied in a slightly different setting in [3].

For two distributions coming from the same exponential family and differing only in the value of parameter θ , the Rényi divergence $D_r(f_{\mathbf{X},\theta^1}, f_{\mathbf{X},\theta^2})$ is given by

$$D_r(f_{\mathbf{X},\theta^1}, f_{\mathbf{X},\theta^2}) = \frac{1}{r(r-1)} \log \int_{\mathbf{S}} \left(\frac{f_{\mathbf{X},\theta^1}(\mathbf{x})}{f_{\mathbf{X},\theta^2}(\mathbf{x})} \right)^r f_{\mathbf{X},\theta^2}(\mathbf{x}) d\mu(\mathbf{x})$$

for $r \in \mathbb{R} \setminus \{0, 1\}$; $D_1(f_{\mathbf{X},\theta^1}, f_{\mathbf{X},\theta^2}) = \int_{\mathbf{S}} \log \frac{f_{\mathbf{X},\theta^1}(\mathbf{x})}{f_{\mathbf{X},\theta^2}(\mathbf{x})} f_{\mathbf{X},\theta^1}(\mathbf{x}) d\mu(\mathbf{x})$, and $D_0(f_{\mathbf{X},\theta^1}, f_{\mathbf{X},\theta^2}) = D_1(f_{\mathbf{X},\theta^2}, f_{\mathbf{X},\theta^1})$. These have been studied in much detail previously. One of their applications follows from the fact that under certain assumptions (see [1] for details), the asymptotic distribution of $2D_r(\widehat{\theta}_n, \theta)$ for large n is χ_k^2 , where $\widehat{\theta}_n$ is the maximum likelihood estimator of the true parameter value θ and k is the dimension of θ . We can notice that the generalized likelihood ratio test is obtained as a special case for $r = 1$.

In our situation, in terms of the maximum likelihood estimators \hat{a} and \hat{b}^2 , the test statistics are given by

$$\frac{r(r-1)}{n} D_{r,n}(\widehat{\theta}_n, \theta_0) = A(1-r) \frac{a_0}{\hat{b}_0^2} + \frac{r}{2} \log \frac{\hat{b}_0^2}{\hat{b}^2} - \frac{1}{2} \log \left(r \frac{\hat{b}_0^2}{\hat{b}^2} + (1-r) \right) +$$

$$+ Ar \frac{1}{\hat{b}_0^2} \frac{\hat{b}_0^2}{\hat{b}^2} \hat{a} - \frac{A}{\hat{b}_0^2} \sqrt{\left(r \frac{\hat{b}_0^2}{\hat{b}^2} + (1-r) \right) \left(r \frac{\hat{b}_0^2}{\hat{b}^2} \hat{a}^2 + (1-r)a_0^2 \right)}$$

for the values of the hypothesized θ_0 and $\widehat{\theta}_n$, for which the divergence is finite, and for $r \neq 0, 1$. The estimators \hat{a} and \hat{b}^2 depend on n but for simplicity we do not use the subscript here. Further,

$$\frac{1}{n} D_{1,n}(\widehat{\boldsymbol{\theta}}_n, \boldsymbol{\theta}^0) = -\frac{1}{2} - A \frac{a_0}{b_0^2} - \frac{1}{2} \log \frac{\widehat{b}^2}{b_0^2} + \frac{1}{2} \frac{\widehat{b}^2}{b_0^2} + \frac{A}{2} \frac{a_0^2}{b_0^2} \frac{1}{\widehat{a}} + \frac{A}{2} \frac{1}{b_0^2} \widehat{a}$$

for $\widehat{a} \neq 0$. Moreover, for $a_0 \neq 0$ we have

$$\frac{1}{n} D_{0,n}(\widehat{\boldsymbol{\theta}}_n, \boldsymbol{\theta}^0) = -\frac{1}{2} - \frac{1}{2} \log \frac{b_0^2}{\widehat{b}^2} + \frac{1}{2} \frac{b_0^2}{\widehat{b}^2} + \frac{A}{2} \frac{a_0}{b_0^2} \frac{b_0^2}{\widehat{b}^2} + \frac{A}{2} \frac{1}{b_0^2 a_0} \frac{b_0^2}{\widehat{b}^2} \widehat{a}^2 - A \frac{1}{b_0^2} \frac{b_0^2}{\widehat{b}^2} \widehat{a}.$$

If the number of observations is sufficient, these statistics can be used to test the hypothesis $H_3 : (a, b^2)^\top = (a_0, b_0^2)^\top$ against the alternative $K_3 : (a, b^2)^\top \neq (a_0, b_0^2)^\top$. However, perhaps the most interesting hypothesis of this type, the hypotheses that we observe the Wiener process, cannot be tested asymptotically in this way (see [1]).

The problem could be solved by using the exact distribution of the test statistic. For this reason and for several others not mentioned here, the exact distribution is an interesting issue. However, it is beyond the scope of the present paper to derive it.

Instead of that, let us now study a motivation to consider such a numerous family of tests. We can notice that for a given hypothesized value $(a_0, b_0^2)^\top$ tests based on the Rényi divergences differing in the value of r will have different critical regions. Some of the regions will be more sensitive to departures from a and some others will be more sensitive to departures from b^2 . Several examples of this phenomenon are shown in Figure 2.

However, a critical region more sensitive to departures from a than from b^2 does not guarantee that a test for a given fixed n will be more sensitive to departures from a than from b^2 . Thorough understanding of behaviour of the maximum likelihood estimators \widehat{a} and \widehat{b}^2 would be needed to study the correspondence between the shape of the critical region and the power of the test.

Further, we can notice that for given values of $r \notin [0, 1]$ and $(a_0, b_0^2)^\top$ there are regions where the divergence is infinite (white areas in Figure 2). A finite divergence value is reached when $\frac{r}{b^2} + \frac{(1-r)}{b_0^2} > 0$ and $\frac{r\widehat{a}^2}{b^2} + \frac{(1-r)a_0^2}{b_0^2} \geq 0$.

The finite-values areas are shown for several values of r in Figure 3. The shape of these areas can be derived analytically and it can be shown that for $|r| \rightarrow \infty$ the true parameter value “moves to the boundary” of the finite-divergence-values area. However, for each fixed n we can find an open ball containing the true parameter value, where the divergence is finite. Thus, the probability of obtaining an infinite divergence tends to zero for large n under the null hypothesis. However, infinite values are an issue in deriving the exact test statistic distribution and they are also of interest when choosing from the family of tests.

Let us remark that an infinite value cannot be reached for $r \in (0, 1)$ (see [1]) and it occurs with zero probability for $r = 0$ or $r = 1$. However, $r = 0$ cannot be used when we need to test for $a_0 = 0$.

5 Conclusion and discussion

We have explored quality of tests concerning parameters of a random process and constructed using a limited information. We have seen that even when we base our decision on an incomplete information, we are still able to construct a variety of tests with convenient properties. The choice of the test has to be carefully considered, depending on the set of alternatives of interest and on several other issues.

There still remain plenty of interesting issues not addressed by the present paper. Let us mention at least the infinite expected experiment time under some hypotheses, experiment design optimized with respect to other criteria than power, a rigorous study of powers of tests under the assumption of unknown variance, optimized choice of r for the tests based on the Rényi divergences, and many others. These are still to be explored.

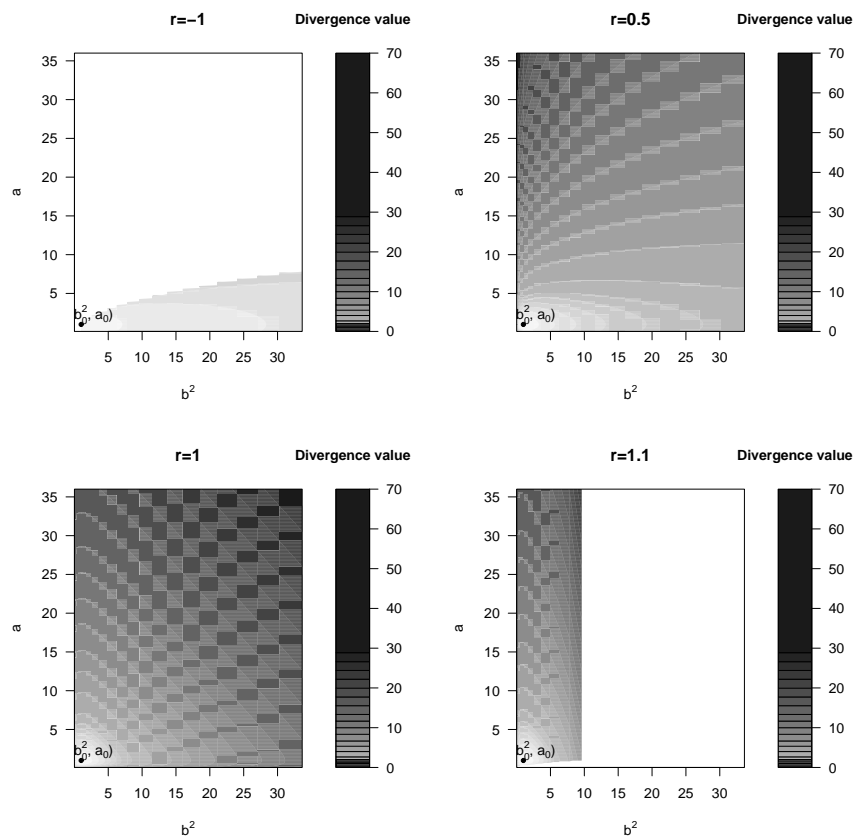


Figure 2: Values of the Rényi divergences for $(a_0, b_0^2)^\top = (1, 1)^\top$ and varying r .

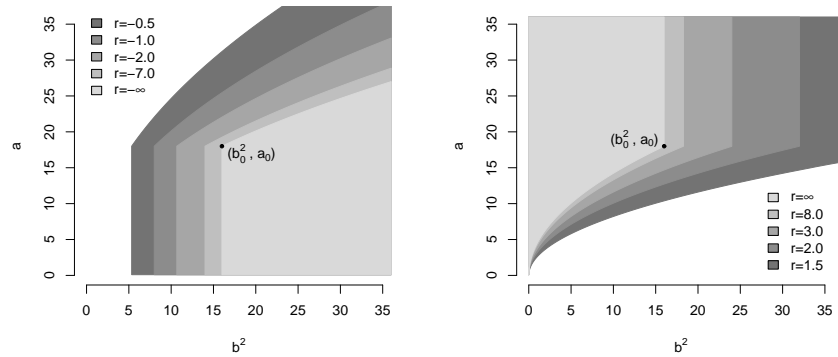


Figure 3: Areas with a finite Rényi divergence for $(a_0, b_0^2)^\top = (18, 16)^\top$ and for various values of r . The finite values areas are nested within each other; higher values of $|r|$ lead to smaller finite-values areas. Limiting areas are shown as well.

References

- [1] Kvitkovičová A. (2008). *Statistical inference for random processes*. Master Thesis, Charles University, Prague.
- [2] Lehmann E.L. (1986). *Testing statistical hypotheses*. Springer, New York.
- [3] Rivas M.J., Santos M.T., Morales D. (2005). *Rényi test statistics for partially observed diffusion processes*. Journal of Statistical Planning and Inference, **127**, 91 – 102.
- [4] Seshadri V. (1999). *The inverse Gaussian distribution: statistical theory and applications*. Springer, New York.

Acknowledgement: I would like to thank doc. RNDr. Daniel Hlubinka, Ph.D. for valuable suggestions and comments, which helped me during my work on the project. The project was supported by grant GAČR 201/08/0486.

Address: A. Kvitkovičová, MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8 – Karlín

E-mail: kvitko@karlin.mff.cuni.cz

MODEL PŘENOSU GENŮ V POPULACI JAKO ZÁKLAD PRO HOSPODÁRNÉ NAKLÁDÁNÍ S GENOVÝMI ZDROJI

Petr Lachout, Milan Lstibůrek

Klíčová slova: geneticky podmíněná proměnlivost, příbuznost, inbreeding.

Abstrakt: Účelem tohoto příspěvku je teoreticky popsat přenos genů v obecné populaci. Představuje se v něm model vhodný pro předávání genů z rodičů na potomky. Pro využití v aplikovaných oborech jsou uvedeny přesnější formule.

Abstract: The purpose of this paper is to layout a solid theoretical foundation to gene transfer in general population. It presents a model convenient for gene transfer from parents to offsprings. More accurate computations of introduced model are presented to account for realistic scenarios in applied fields.

1 Úvod

Jedním z často diskutovaných témat v aplikovaných oborech (zemědělství, lesnictví, ochrana přírody) je záchrana a reprodukce genových zdrojů. V praktickém hospodaření s genovými zdroji často dochází k hledání kompromisu mezi produkcí a ochranou. V ideálním případě je takový kompromis stanoven na základě exaktních vědeckých postupů. V dílčích aplikacích se však často setkáváme s přístupy, které jsou od tohoto ideálního stavu velmi vzdáleny (viz četné diskuze na téma genových zdrojů v národních parcích). Je pak možné spekulovat o hospodárnosti takových postupů.

Velmi často používaným pojmem je tzv. biodiverzita. Ve výše uvedených programech nás primárně zajímají kauzální složky proměnlivosti (polymorfismu), které mají genetický základ. V praxi často neumíme popsat přesný fyziologický mechanismus a nedokážeme jednoduše popsat vztah mezi polymorfismem na úrovni DNA a sledovanou proměnlivostí řady adaptačních znaků na úrovni fenotypu. Toto především platí u tzv. kvantitativních znaků, jejichž exprese bývá ovlivněna velkým počtem genů s relativně malým účinkem. Obecně můžeme tvrdit, že takové znaky v populacích převládají; viz [4]. Dále bylo empiricky prokázáno, že převažující složka genetického rozptylu je aditivní povahy, tj. má základ v přímých (aditivních) účincích jednotlivých genů ovlivňujících expresi sledovaných znaků; viz [3].

Ke kvantifikaci kauzálních složek rozptylu se vychází z kovariance mezi příbuznými jedinci. Konkrétní příbuzenský vztah implikuje vyšší pravděpodobnost sdílení alel identických svým původem (tzv. coancestry). Pokud známe příbuzenské vazby v populaci, můžeme jednoduše odhadnout kauzální složky genetického rozptylu; viz [1].

Cílem této studie je matematický popis přenosu genů z jedné generace na druhou. Tento model se dá využít např. při konstrukci modelů matematického programování, kde je cílem dosažení kompromisu mezi produkcí a ochranou, např. maximalizací selekčního kritéria s omezením ztráty relativní genové diversity apod.; viz [2]. Tyto modely pak mají přímé uplatnění v aplikovaných disciplínách.

Informace, z kterých v článku vycházíme jsou úplný rodokmen sledované populace a úplná znalost rodičovské generace.

2 Matematický model pro přenos genů

Pozorujeme populaci, kterou budeme označovat jako neprázdnou množinu \mathcal{K} . V této množině je určena podmnožina $\mathcal{K}_0 \subset \mathcal{K}$, která reprezentuje rodičovskou generaci.

Populace se vyvíjí tak, že dvojice jedinců (rodiče) dává vznik novému jedinci (potomek). Pro obecnější aplikaci připustíme situaci, kdy oba rodiče jsou jeden a ten samý jedinec. Mezi rodiči rozlišujeme a budeme je označovat Rodič-I a Rodič-II. Význam je otec a matka. Nový jedinec (potomek) získává geny pouze od svých rodičů. Mechanismus dědičnosti je obecně dobře popsán, např. v rámci středoškolské biologie. Množinu všech genů ve sledované populaci budeme označovat jako neprázdnou množinou \mathcal{G} .

Každý jedinec $a \in \mathcal{K}$ má na sledovaném lokusu dvojici alel (g_1, g_2) . Alelu g_1 získal od Rodiče-I a alelu g_2 od Rodiče-II. U jedinců v rodičovské generaci známe genotyp. U ostatních jedinců genotyp neznáme, ale na základě rodokmenu a znalosti mechanismu přenosu genů dokážeme určit jeho rozdělení pravděpodobnosti. Označme si genotyp jedince $a \in \mathcal{K}$ symbolem Q_a .

Pro jedince $a, b \in \mathcal{K}$ a alely $i, j \in \mathcal{G}$ označme

$$\tilde{W}(a; i) - \text{pravděpodobnost, že } a \text{ bude předávat alelu } i, \quad (1)$$

$$W(a, b; i, j) - \text{pravděpodobnost, že } a \text{ předá alelu } i \text{ a } b \text{ alelu } j. \quad (2)$$

Tyto pravděpodobnosti splňují evidentní rovnosti:

$$\sum_{i \in \mathcal{G}} \tilde{W}(a; i) = 1, \quad (3)$$

$$\sum_{j \in \mathcal{G}} W(a, b; i, j) = \tilde{W}(a; i), \quad (4)$$

$$\sum_{i \in \mathcal{G}} \sum_{j \in \mathcal{G}} W(a, b; i, j) = 1, \quad (5)$$

$$W(a, b; i, j) = W(b, a; j, i). \quad (6)$$

Tyto pravděpodobnosti umožňují studovat genetickou příbuznost jedinců.

Uvažujeme, že máme dáno ohodnocení $h : \mathcal{G}^2 \rightarrow \mathcal{R}$, které měří „kvalitu“ každé dvojice genů. „Genetickou kvalitu“ jedince $a \in \mathcal{K}$ měříme pomocí střední hodnoty

$$\tilde{H}(\mathbf{h}; a) = \sum_{i \in \mathcal{G}} h(i, i) \tilde{W}(a; i). \quad (7)$$

Dva jedince $a, b \in \mathcal{K}$ srovnáváme pomocí „koeficientu podobnosti“, což je střední hodnota

$$H(\mathbf{h}; a, b) = \sum_{i, j \in \mathcal{G}} h(i, j) W(a, b; i, j). \quad (8)$$

Evidentně platí

$$H(\mathbf{h}; b, a) = H(\bar{\mathbf{h}}; a, b), \quad (9)$$

kde $\bar{\mathbf{h}}$ značí transponovanou funkci, tj. $\bar{h}(j, i) = h(i, j)$ pro každé $i, j \in \mathcal{G}$.

Poznamenejme, že pro symetrickou funkci, tj. $\bar{\mathbf{h}} = \mathbf{h}$, je koeficient podobnosti symetrická funkce, tj. $H(\mathbf{h}; b, a) = H(\mathbf{h}; a, b)$ pro každé $a, b \in \mathcal{K}$.

3 Vzorečky pro přenos genů

Nyní upřesníme tvorbu pravděpodobností přenosu genů. Pro další text přidáme požadavky:

1. Pravděpodobnost předání alely rodičem svému potomkovi je pro každou alelu rodiče stejná.
2. Přiřazení alel potomka je nezávislé na přiřazení alel u ostatních jedinců.

Pro pravděpodobnosti přenosu genů platí následující vzorečky:

3.1 Obecné vztahy

Když $a \in \mathcal{K}$ a $i \in \mathcal{G}$, pak

$$\tilde{W}(a; i) = \frac{1}{2} \sum_{\chi \in \mathcal{G}} \text{Prob}(Q_a = (i, \chi)) + \frac{1}{2} \sum_{\varphi \in \mathcal{G}} \text{Prob}(Q_a = (\varphi, i)). \quad (10)$$

Když $a, b \in \mathcal{K}$ a $i, j \in \mathcal{G}$, pak

$$\begin{aligned} W(a, b; i, j) &= \frac{1}{4} \sum_{\chi \in \mathcal{G}} \sum_{\omega \in \mathcal{G}} \text{Prob}(Q_a = (i, \chi), Q_b = (j, \omega)) \\ &+ \frac{1}{4} \sum_{\varphi \in \mathcal{G}} \sum_{\omega \in \mathcal{G}} \text{Prob}(Q_a = (\varphi, i), Q_b = (j, \omega)) \\ &+ \frac{1}{4} \sum_{\chi \in \mathcal{G}} \sum_{\psi \in \mathcal{G}} \text{Prob}(Q_a = (i, \chi), Q_b = (\psi, j)) \\ &+ \frac{1}{4} \sum_{\varphi \in \mathcal{G}} \sum_{\psi \in \mathcal{G}} \text{Prob}(Q_a = (\varphi, i), Q_b = (\psi, j)). \end{aligned} \quad (11)$$

Pokud $a = b$ nebo a je jedním z rodičů b nebo b je jedním z rodičů a , pak některé ze sčítanců v (11) budou nulové.

1. Pro $a = b$ a $i = j$ získáme

$$W(a, a; i, i) = \frac{1}{2} \text{Prob}(Q_a = (i, i)) + \frac{1}{2} \tilde{W}(a; i). \quad (12)$$

2. Pro $a = b$ a $i \neq j$ získáme

$$W(a, a; i, j) = \frac{1}{4} \text{Prob}(Q_a = (i, j)) + \frac{1}{4} \text{Prob}(Q_a = (j, i)). \quad (13)$$

3. Když a je Rodič-I pro b and $i = j$, pak zjistíme

$$\begin{aligned} W(a, b; i, i) &= \frac{1}{4} \text{Prob}(Q_a = (i, i)) \\ &+ \frac{1}{8} \sum_{\chi \in \mathcal{G}} \text{Prob}(Q_a = (i, \chi)) \\ &+ \frac{1}{8} \sum_{\varphi \in \mathcal{G}} \text{Prob}(Q_a = (\varphi, i)) \\ &+ \frac{1}{8} \sum_{\chi \in \mathcal{G}} \text{Prob}(Q_a = (i, \chi), Q_b = (i, i)) \\ &+ \frac{1}{8} \sum_{\varphi \in \mathcal{G}} \text{Prob}(Q_a = (\varphi, i), Q_b = (i, i)) \\ &+ \frac{1}{8} \sum_{\chi \in \mathcal{G}} \text{Prob}(Q_a = (i, \chi), Q_b = (\chi, i)) \\ &+ \frac{1}{8} \sum_{\varphi \in \mathcal{G}} \text{Prob}(Q_a = (\varphi, i), Q_b = (\varphi, i)). \end{aligned} \quad (14)$$

4. když a je Rodič-I b a $i \neq j$, pak zjistíme

$$\begin{aligned} W(a, b; i, j) &= \frac{1}{8} \text{Prob}(Q_a = (i, j)) \\ &+ \frac{1}{8} \text{Prob}(Q_a = (j, i)) \\ &+ \frac{1}{8} \sum_{\chi \in \mathcal{G}} \text{Prob}(Q_a = (i, \chi), Q_b = (i, j)) \\ &+ \frac{1}{8} \sum_{\varphi \in \mathcal{G}} \text{Prob}(Q_a = (\varphi, i), Q_b = (i, j)) \\ &+ \frac{1}{8} \sum_{\chi \in \mathcal{G}} \text{Prob}(Q_a = (i, \chi), Q_b = (\chi, j)) \\ &+ \frac{1}{8} \sum_{\varphi \in \mathcal{G}} \text{Prob}(Q_a = (\varphi, i), Q_b = (\varphi, j)). \end{aligned} \quad (15)$$

5. Analogické výsledky lze odvodit i v ostatních případech, tj. „ a je Rodič-II b “, „ b je Rodič-I a “ a „ b je Rodič-II a “.

3.2 Rekurzivní vztahy

Pro pravděpodobnosti přenosu genů platí následující rekurzivní vzorečky:

1. Když $a, b \in \mathcal{K}$ a $i, j \in \mathcal{G}$, pak

$$\sum_{i \in \mathcal{G}} \tilde{W}(a; i) = 1, \quad (16)$$

$$\sum_{j \in \mathcal{G}} W(a, b; i, j) = \tilde{W}(a; i), \quad (17)$$

$$\sum_{i \in \mathcal{G}} \sum_{j \in \mathcal{G}} W(a, b; i, j) = 1, \quad (18)$$

$$W(a, b; i, j) = W(b, a; j, i). \quad (19)$$

2. Když $P \in \mathcal{K}$ je potomek rodičů A, B , $\gamma \in \mathcal{K}$ a $i, j \in \mathcal{G}$, pak

$$\tilde{W}(P; i) = \frac{1}{2} \tilde{W}(A; i) + \frac{1}{2} \tilde{W}(B; i), \quad (20)$$

$$W(P, \gamma; i, j) = \frac{1}{2} W(A, \gamma; i, j) + \frac{1}{2} W(B, \gamma; i, j). \quad (21)$$

3. Když $P \in \mathcal{K}$ je potomek rodičů A, B a $i \in \mathcal{G}$, pak

$$W(P, P; i, i) = \frac{1}{2} W(A, B; i, i) + \frac{1}{4} \tilde{W}(A; i) + \frac{1}{4} \tilde{W}(B; i). \quad (22)$$

4. Když $P \in \mathcal{K}$ je potomek rodičů A, B a $i, j \in \mathcal{G}$, $i \neq j$, pak

$$W(P, P; i, j) = \frac{1}{4} W(A, B; i, j) + \frac{1}{4} W(A, B; j, i). \quad (23)$$

4 Rekurzivní vzorečky pro koeficient podobnosti

Pro genetickou kvalitu a koeficient podobnosti platí následující vztahy:

1. Platí

$$H(h; b, a) = H(\bar{h}; a, b), \quad (24)$$

kde \bar{h} je transponovaná funkce definovaná předpisem

$$\bar{h}(i, j) = h(j, i) \quad \text{pro všechny } i, j \in \mathcal{G}. \quad (25)$$

2. Koeficient podobnosti je symetrická funkce, tj.

$$H(h; b, a) = H(h; a, b), \quad (26)$$

kdykoli h je symetrická funkce, tj. $h(j, i) = h(i, j)$ pro všechny $i, j \in \mathcal{G}$.

3. Když $P \in \mathcal{K}$ je potomek rodičů A, B a $\gamma \in \mathcal{K}$, $P \neq \gamma$, pak

$$H(h; P, \gamma) = \frac{1}{2}H(h; A, \gamma) + \frac{1}{2}H(h; B, \gamma). \quad (27)$$

4. Když $P \in \mathcal{K}$ je potomek rodičů A, B , pak

$$\tilde{H}(h; P) = \frac{1}{2}\tilde{H}(h; A) + \frac{1}{2}\tilde{H}(h; B), \quad (28)$$

$$H(h; P, P) = \frac{1}{4}H(h; A, B) + \frac{1}{4}H(h; B, A) \quad (29)$$

$$+ \frac{1}{4}\tilde{H}(h; A) + \frac{1}{4}\tilde{H}(h; B). \quad (30)$$

5 Míra příbuznosti

Speciální volba funkce h se používá jako „míra příbuznosti“ dvou jedinců. Tato volba je

$$\begin{aligned} h(i, j) &= 1 \quad \text{if } i = j, \\ &= 0 \quad \text{if } i \neq j. \end{aligned} \quad (31)$$

Potom je míra příbuznosti dána vzorcem

$$C_{a,b} = H(h; a, b) = \sum_{i \in \mathcal{G}} W(a, b; i, i). \quad (32)$$

Genetická kvalita všech jedinců je pro tuto volbu stejná. Pro každé $a \in \mathcal{K}$ platí

$$\tilde{H}(h; a) = \sum_{i \in \mathcal{G}} \tilde{W}(a; i) = 1. \quad (33)$$

Vlastnosti míry příbuznosti jsou zvláštním případem vlastností z kapitoly 4.

1. Míra příbuznost je symetrická, tj.

$$C_{b,a} = C_{a,b} \quad \text{pro každé } a, b \in \mathcal{K}, \quad (34)$$

neboť h je symetrická funkce.

2. Když $P \in \mathcal{K}$ je potomek rodičů A, B a $\gamma \in \mathcal{K}$, $P \neq \gamma$, pak

$$C_{P,\gamma} = \frac{1}{2}C_{A,\gamma} + \frac{1}{2}C_{B,\gamma}. \quad (35)$$

3. Když $P \in \mathcal{K}$ je potomek rodičů A, B , pak

$$C_{P,P} = \frac{1}{2}C_{A,B} + \frac{1}{2}. \quad (36)$$

5.1 Pozitivní definitnost

V této kapitole ukážeme, že za obecných předpokladů je míra příbuznosti pozitivně definitní.

Věta 1. Předpokládejme, že každý jedinec rodičovské generace má dva různé geny a každý z genů se vyskytuje v základní generaci pouze jednou. Pak je míra příbuznosti C pozitivně definitní funkce a $\frac{1}{2} \leq C_{a,a} < 1$ pro každé $a \in \mathcal{K}$.

Důkaz: Bez újmy na obecnosti budeme předpokládat, že jedinci sledované populace jsou uspořádáni tak, že jedinci z rodičovské generace mají nejnižší indexy a ostatní jedinci jsou očíslováni chronologicky, podle času jejich zrodu.

Z vlastnosti (36) okamžitě plyne, že $\frac{1}{2} \leq C_{a,a}$ pro každé $a \in \mathcal{K}$. Je tedy třeba ukázat pouze pozitivní definitnost a že $C_{a,a} < 1$ pro každé $a \in \mathcal{K}$.

Pozitivní definitnost funkce znamená, že matice $(C_{a,b}, a \in \mathcal{K}, b \in \mathcal{K})$ je pozitivně definitní. Vlastnost ukážeme indukcí přes rozměr matice. Budeme postupně přidávat potomky.

1. Levá horní submatice typu $K_0 \times K_0$ matice C má na diagonále $\frac{1}{2}$ a mimo ni 0. Je tedy pozitivně definitní a na diagonále má čísla menší nežli '1'.
2. Předpokládejme, že již víme, že levá horní submatice typu 'dxd' matice C je pozitivně definitní a na diagonále má čísla menší nežli '1'. Označme si tuto submatici symbolem D a submatici, která má o řádek a sloupec více, tedy je typu '(d+1)x(d+1)', symbolem Q . Přibývá nám jedinec 'i', jehož rodiči jsou 'g' a 'h'.

(a) Nový diagonální prvek je

$$Q_{d+1,d+1} = \frac{1}{2} + \frac{1}{2}D_{g,h} < \frac{1}{2} + \frac{1}{2} = 1.$$

(b) Matice Q je z matice D jednoznačně odvozena pomocí pravidel 2) a 3). Výpočet můžeme maticově zapsat následovně

$$Q = M^T D M + \Delta,$$

kde

$$M = \left(\text{II} \mid \begin{array}{l} \frac{1}{2} \delta_g + \frac{1}{2} \delta_h \end{array} \right),$$

II označuje jednotkovou matici, tj. na diagonále '1' a mimo ni '0', δ_j je sloupcový vektor, který má na j -tém místě '1' a jinak '0', Δ je matice, která je celá nulová až na poslední diagonální prvek, který je roven

$$\begin{aligned} \Delta_{d+1,d+1} &= \frac{1}{2} + \frac{1}{2}D_{g,h} - \frac{1}{4}(D_{g,g} + 2D_{g,h} + D_{h,h}) \\ &= \frac{1}{2} - \frac{1}{4}(D_{g,g} + D_{h,h}). \end{aligned}$$

Diagonální prvky matice D jsou menší než '1', proto $\Delta_{d+1,d+1} > 0$. Nyní ověříme pozitivní definitnost matice Q . Vezměme pro ten účel sloupcový vektor $\eta = \begin{pmatrix} \xi \\ c \end{pmatrix} \neq 0$, kde ξ je vektor o d -složkách a c je číslo.

Nyní jsou dvě možnosti.

- Když $c = 0$, pak $\xi \neq 0$ a

$$\eta^\top Q \eta = \xi^\top D \xi > 0.$$

- Když $c \neq 0$, pak

$$\begin{aligned} \eta^\top Q \eta &= \eta^\top (M^\top D M + \Delta) \eta \\ &= \eta^\top M^\top D M \eta + c^2 \Delta_{d+1,d+1} \\ &\geq c^2 \Delta_{d+1,d+1} > 0. \end{aligned}$$

Tím je ověřeno, že matice Q je také pozitivně definitní.

Využitím matematické indukce zjišťujeme, že matice C je pozitivně definitní a na diagonále má čísla menší nežli '1'.

□

Literatura

- [1] Falconer D.S., Mackay T.F.C. (1996). *Introduction to quantitative genetics*. Pearson Education Ltd.
- [2] Funda T., Lstibůrek, M., Lachout, P., Klápště, J., El-Kassaby Y. (2008). *Deployment strategies in advanced-generations seed orchards*. (v tisku).
- [3] Hill W.G., Goddard M.E., Visscher P.M. (2008). *Data and theory point to mainly additive genetic variance for complex traits*. PLoS Genet. **4**, 1–10.
- [4] Lynch M., Walsh B. (1998). *Genetics and analysis of quantitative traits*. Sinauer Assoc.

Poděkování: Tato práce byla podporována výzkumným záměrem MSM 0021620839, grantem GA ČR 201/08/0486 a grantem GA ČR 521/07/P337.

Adresa: P. Lachout, MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8 – Karlín
M. Lstibůrek, Česká zemědělská univerzita v Praze, Fakulta lesnická a dřevařská, Kamýčká 1176, 165 21 Praha 6 - Suchbátka

E-mail: lstiburek@gmail.com, lachout@karlin.mff.cuni.cz

ANALÝZA ROZMÍSTĚNÍ ELEKTRO- NICKÝCH PODATELEN OBCÍ V ČESKÉ REPUBLICĚ

Radka Lechnerová, Tomáš Lechner

Klíčová slova: E-government, elektronické podatelny, bodové procesy, sumární statistiky.

Abstrakt: E-Government jako jeden z prostředků transformačního procesu veřejné správy vytváří nové elektronické veřejné služby. Rozsah poskytování těchto služeb však dosud není na očekávané úrovni a neodpovídá plně platné legislativě, což vyvolává celou řadu otázek, proč tomu tak je. V článku zkoumáme aktuální funkční elektronické podatelny úřadů územních samosprávných celků, a to zejména s ohledem na jejich prostorové rozmístění v rámci České republiky, které analyzujeme prostřednictvím sumárních statistik bodových procesů. Tímto způsobem hledáme odpověď na otázku, zda mezi obcemi, které provozují elektronickou podatelnu, existují nějaké interakce vyplývající z jejich vzájemných poloh nebo zda jsou elektronické podatelny rozmístěny v rovnoměrně náhodně vybraných obcích. Získané výsledky jsou podstatné pro vytvoření celkového obrazu průběhu procesu E-Governmentu v České republice.

Abstract: E-Government as one of the instruments of transformation process of Public Administration creates new electronic public services. The scope of providing these services is not however at the expected level yet and it does not correspond to the legislative at all. This fact brings a lot of questions asking why. In the present contribution, we investigate current workable electronic registries of municipalities, especially with respect to their spatial distribution in the area of the Czech Republic. It is analysed by means of summary statistics of the point processes. In this way we are trying to find the answer to the question whether there is an interaction between municipalities providing an electronic registry and resulting from their respective positions or if electronic registries are provided by uniformly at random distributed municipalities of the Czech Republic. Obtained results are important for the creation of a complex view of the progress of E-Government process in the Czech Republic.

1 Úvod

Transformační proces veřejné správy probíhá v České republice již od počátku 90. let minulého století. V posledních letech jsou významnou měrou využívány informační a komunikační technologie. Jejich aplikace ve veřejné správě, jejímž cílem by mělo být celkové zefektivnění výkonu veřejné moci, se obecně

nazývá E-Government [3]. Na základě platných legislativních předpisů lze v tomto procesu identifikovat několik základních nástrojů [2], z nichž velmi významné postavení zastává elektronická podatelna jako prostředník poskytování elektronických služeb veřejné správy.

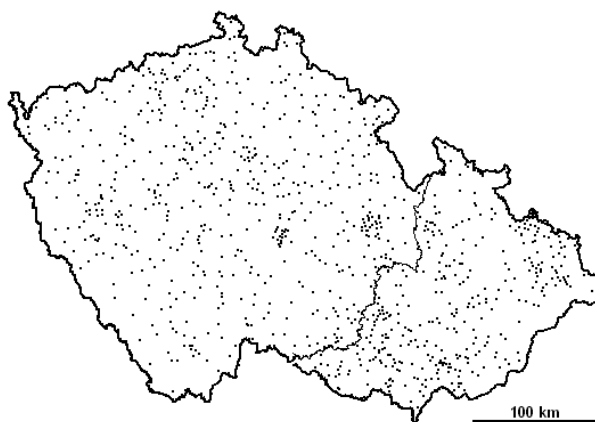
Elektronické podatelny v podobě nástroje zaručené elektronické komunikace jsou po právní stránce založeny na zákonu o elektronickém podpisu, který byl v ČR schválen v roce 2000. Následný vývoj nebyl rozhodně přímočarý, nicméně od 1. ledna 2005 platí vyhláška Ministerstva informatiky ČR o elektronických podatelnách a nařízení vlády č. 495/2004 Sb., které přikazuje všem orgánům veřejné moci provozovat elektronickou podatelnu [2]. Součástí tohoto nařízení je též povinnost sdělit adresu elektronické podatelny Ministerstvu informatiky ČR (v současnosti tuto úlohu převzalo Ministerstvo vnitra ČR), které tyto adresy neprodleně zveřejňuje na Portálu veřejné správy (PVS) [7].

Při našem zkoumání jsme tedy vycházeli ze seznamu adres elektronických podatelen publikovaných na PVS. Lze předpokládat, že existují orgány veřejné moci, které provozují elektronickou podatelnu, a při tom nedodržely tu část nařízení, která dává povinnost předat informaci o adrese Ministerstvu informatiky ČR, resp. Ministerstvu vnitra ČR. Na druhou stranu díky značně komplikované dostupnosti těchto adres, které nejsou v centrálním seznamu, nelze tyto podatelny považovat za plně funkční, a proto se jimi dále nezabýváme. Výběr seznamu adres dále zužujeme na oblast úřadů územně samosprávných celků (ÚSC), která se vyznačuje stejnorodostí pravidel a předpisů, jimiž se tyto orgány musí řídit, a je tedy vhodná k celkovému statistickému zpracování.

Podle platné legislativy by všechny vybrané úřady měly mít realizovanou elektronickou podatelnu, nicméně z aktuálně existujících 6249 obcí ji v únoru roku 2008 mělo funkční, dle adresáře zveřejněného na PVS, pouze 952 úřadů [7]. Poloha těchto obcí je znázorněna na obrázku 1. Z tohoto tristního stavu, kdy po více než 7 letech od nabytí účinnosti nařízení vlády dávající povinnost zřídit elektronické podatelny (původní nařízení vlády č. 304/2001 Sb., kterým se prováděl zákon č. 227/2000 Sb., o elektronickém podpisu) má tento nástroj realizováno pouze přibližně 15 % úřadů územních samosprávných celků, plynou otázky různého typu. My se nyní pokusíme nalézt odpovědi na to, jaké prostorové vztahy jsou mezi úřady, které elektronické podatelny realizovaly, a jak jsou funkční elektronické podatelny rozmístěny v rámci ČR.

2 Data

Jak již bylo zmíněno v úvodu, vycházíme z adresáře elektronických podatelen zveřejněného na Portálu veřejné správy [7]. Adresář byl stažen z příslušné webové stránky v únoru 2007 a v únoru 2008. Protože ve zmíněném období došlo pouze k zanedbatelné změně v řádu jednotek promile, rozhodli jsme se zkoumat pouze aktuální stav roku 2008.

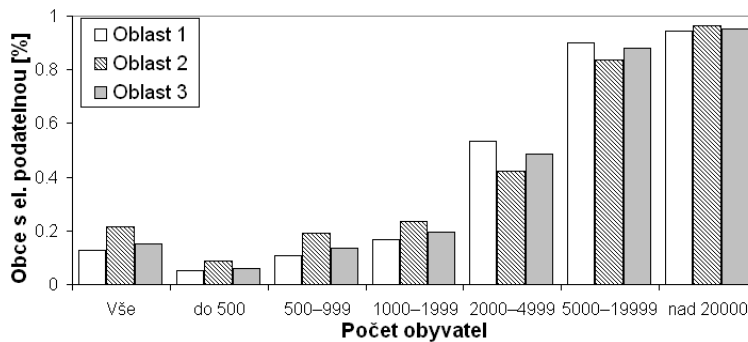


Obrázek 1: Na obrázku jsou vyznačeny polohy obcí České republiky, které mají elektronickou podatelnu, stav v únoru 2008. Slabá čára rozděluje Českou republiku na Oblast 1 (vlevo) a Oblast 2 (vpravo).

Pro zkoumání prostorového rozložení jsme získané adresářové údaje spárovali s daty prezentovanými Českým statistickým úřadem v rámci Územně identifikačního registru ÚIR-ZSJ [6]. Geografická data jsou zde uváděna v metrech v souřadnicovém systému Jednotné trigonometrické sítě katastrální (S-JTSK), který je definován v nařízení vlády č. 430/2006 Sb., o stanovení geodetických referenčních systémů a státních mapových děl závazných na území státu a zásadách jejich používání. Dle tohoto nařízení je systém S-JTSK určen Besselovým elipsoidem s parametry $a = 6377397,15508$ m, $b = 6356078,96290$ m, kde a je délka hlavní poloosy, b je délka vedlejší poloosy, Křovákovým konformním kuželovým zobrazením v obecné poloze a souborem souřadnic bodů z vyrovnání trigonometrických sítí. Výhodou tohoto souřadnicového systému je, že v něm můžeme provádět standardní geodetická měření. Vzdálenosti bodů lze získat s chybou, která může být maximálně v jednotkách metrů. V případě, kdy budeme aproximovat obce a města ČR bodem v místě polohy úřadu příslušného ÚSC, je tato chyba naprosto zanedbatelná.

3 Metody

Nechť $\mathbf{X} = \{X_1, \dots, X_k\}$ jsou body, které reprezentují polohy úřadů ÚSC, jimiž aproximujeme obce v ČR. Obce s elektronickou podatelnu nechť jsou reprezentovány body $\mathbf{X}' = \{X'_1, \dots, X'_n\}$, přičemž platí, že $\mathbf{X}' \subseteq \mathbf{X}$. Pozorovacím oknem nechť je území České republiky nebo její část. Naše analýza se zabývá otázkou, zda mezi body \mathbf{X}' jsou nebo nejsou nějaké interakce (přitažlivé či odpudivé). Budeme tedy testovat hypotézu, že body \mathbf{X}' jsou rovnoměrně náhodně rozdělené na množině bodů \mathbf{X} , proti alternativě, že mezi body \mathbf{X}' existují přitažlivé či odpudivé interakce za podmínky, že $\mathbf{X}' \subseteq \mathbf{X}$.



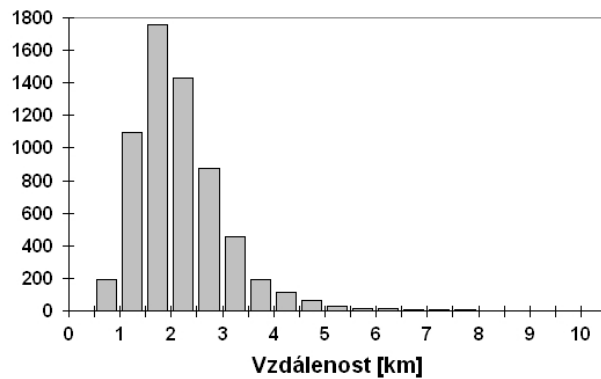
Obrázek 2: Procento obcí, které mají elektronickou podatelnu, v závislosti na počtu jejich obyvatel.

K naší analýze bodových procesů použijeme sumární statistiky, kterými lze dobře charakterizovat prostorové rozmístění bodů. Konkrétně jsme použili distribuční funkci nejbližších sousedů (G), sférickou kontaktní distribuční funkci (F) a párovou korelační funkci (g) (viz např. [4, 5]). Z jejich definic vyplývá, že distribuční funkce $G(r)$ resp. $F(r)$ určují pravděpodobnost, že se do vzdálenosti r od bodu procesu resp. libovolného bodu v prostoru vyskytuje (v případě funkce G jiný) bod procesu. Poznamenejme, že jsme použili Kaplan-Meierovy odhady pro funkce G a F [1] a Ripleyho odhad pro funkci g [5], které jsou implementovány v programu R (balíček SPATSTAT). Všechny zmíněné odhady zahrnují korekci na okrajové efekty.

Pro ověření výše uvedené hypotézy provedeme následující test. Pro rovnoměrně náhodný výběr N bodů z množiny bodů \mathbf{X} spočítáme výše uvedené statistiky. Tento postup opakujeme 199krát a poté ze všech simulací vykreslíme jako obálky vždy pátou největší a pátou nejmenší hodnotu v každém bodě grafu. Pokud by graf funkce spočtený pro \mathbf{X}' ležel mimo oblast vymezenou obálkami, znamenalo by to, že hypotézu o rovnoměrně náhodném rozmístění elektronických podatel v obcích ČR můžeme zamítnout na hladině testu 0,05.

4 Numerické výsledky

Pro zkoumaná data jsme nejprve spočetli základní statistické parametry. Z celkového počtu 6249 obcí mělo v únoru roku 2008 realizováno funkční elektronickou podatelnu přibližně 15,2 % obcí (viz obrázek 1). Pro podrobnější zkoumání jsme rozdělili obce podle velikostních kategorií (dle standardu používaného Českým statistickým úřadem). Podle výsledků prezentovaných v [2] jsou mezi jednotlivými kraji ČR nezanedbatelné rozdíly v procentech obcí, které povinnost provozovat elektronickou podatelnu plní. Z tohoto důvodu jsme si ČR, pro potřeby studia metodami popsány v předchozí kapi-



Obrázek 3: Histogram vzdáleností mezi nejbližšími obcemi v ČR.

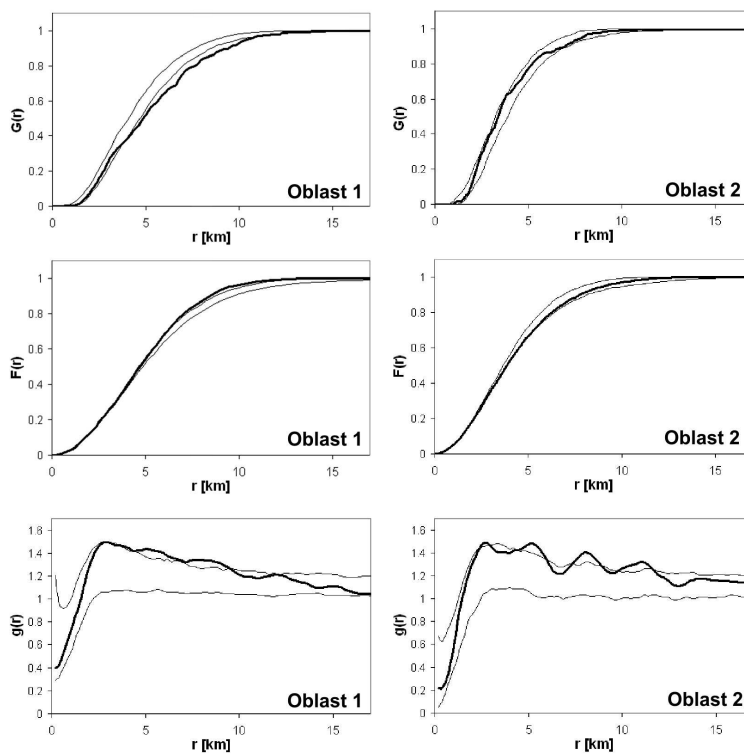
tole, rozdělili na dvě disjunktní souvislé oblasti a to tak, aby v jedné z nich byly zastoupeny kraje s lepším stavem plnění zmíněné povinnosti. Spolu s podmínkou souvislosti pak dostáváme následující oblasti:

- Oblast 1 je tvořena hl. m. Prahou a dále Středočeským, Jihočeským, Plzeňským, Karlovarským, Ústeckým, Libereckým, Královéhradeckým a Pardubickým krajem a krajem Vysočina.
- Oblast 2 je tvořena Jihomoravským, Olomouckým, Zlínským a Moravskoslezským krajem.
- Oblast 3 je vymezena celým územím ČR, tj. je sjednocením Oblasti 1 a Oblasti 2.

Obrázek 2 potvrzuje celkově lepší situaci v Oblasti 2, což také platí pro malé obce až do velikosti 1999 obyvatel. Mezi středními obcemi je stav lepší v Oblasti 1. Pro velké obce je situace v obou oblastech téměř shodná. V každém případě platí, že čím vyšší je velikostní kategorie, tím větší procento obcí má elektronickou podatelnu realizovanou. Zatímco z obcí do 500 obyvatel nabízí cestu zaručené elektronické komunikace pouze přibližně 6 % obcí, v rozmezí 2000 až 4999 obyvatel je to téměř polovina z nich a u velkých obcí je to více než 95 %.

Na obrázku 3 je znázorněn histogram vzdáleností mezi nejbližšími obcemi ČR. Z něj také vyplývá, že v rozmezí od 0,5 km do 10,4 km se od libovolné obce nachází další obec, přičemž převládají menší vzdálenosti. Podíl obcí, které jsou od sebe vzdáleny více než 3,5 km, nepřevyšuje 4 %.

Než začneme rozebírat výsledky plynoucí z grafů sumárních statistik, připomeňme, že se pohybujeme na množině obcí ČR. Při zkoumání prostorových závislostí se nejprve zaměříme na Oblast 1. Z grafů na obrázku 4 v levém sloupci vyplývá, že hypotézu o rovnoměrně náhodném rozmístění



Obrázek 4: Grafy distribuční funkce nejbližších sousedů (první řádek), sférické kontaktní distribuční funkce (druhý řádek) a párové korelační funkce (třetí řádek) spočtené pro data z Oblastí 1 a 2 (tučné křivky). Dále jsou zde zobrazeny obálky (tenké křivky) vymezující 95% interval spolehlivosti pro testování hypotézy o rovnoměrně náhodném rozmístění obcí s elektronickou podatelnou v obcích ČR.

elektronických podatelen v obcích na 5% hladině testu zamítáme. Nejprve se podíváme na grafy distribučních funkcí vzdáleností G a F (viz první dva obrázky vlevo). Z obou grafů vyplývá, že pro větší vzdálenosti jsou zde známky „odpudivých interakcí“ mezi obcemi s elektronickou podatelnou oproti případu, kdybychom vzali rovnoměrně náhodný výběr obcí ČR o stejném rozsahu. V případě funkce G (horní graf) od 3,8 km dále a v případě funkce F (prostřední graf) od 6 km dále. Pro menší vzdálenosti (do 3,8 km), tedy v případě vzdáleností typických pro většinu obcí, jsou obce s elektronickou podatelnou rozmístěny rovnoměrně náhodně na množině obcí ČR. Z grafu g vyplývá, že vzdálenosti od 3,6 km do 9,4 km mezi jakýmikoliv elektronickými podatelny se vyskytují častěji, než by odpovídalo rovnoměrně náhodnému výběru obcí, nicméně tento rozdíl není příliš výrazný.

Nyní se zaměříme na situaci v Oblasti 2, kterou si popíšeme na základě grafů na obrázku 4 v pravém sloupci. Pokud bychom brali v úvahu pouze výsledky plynoucí z grafů distribučních funkcí vzdáleností (horní a prostřední graf vpravo), testovanou hypotézu o rovnoměrně náhodném rozmístění obcí s elektronickou podatelnou v rámci obcí ČR bychom nemohli zamítnout, protože sledovaná křivka leží uvnitř oblasti ohraničené obálkami. Odlišný výsledek dává párová korelační funkce, kde se vyskytují 4 vrcholy, které se dostaly nad oblast vymezenou obálkami. Vzhledem k tomuto výsledku testovanou hypotézu opět zamítáme. Nejvyšších hodnot dosahuje funkce v bodech 2,8 km a 5,2 km, protože se jedná o poměrně malé vzdálenosti, můžeme zde předpokládat jistou spolupráci mezi blízkými obcemi.

Oblast 3 je sjednocením Oblastí 1 a 2, přičemž poměr rozlohy Oblasti 1 ku Oblasti 2 je přibližně 2,6. Proto není překvapivé, že výsledky pro Oblast 3 jsou obdobné jako v případě Oblasti 1, a proto zde grafy sledovaných funkcí neuvádíme, pouze popíšeme výsledky. Z distribuční funkce nejbližších sousedů vyplývají odpudivé interakce, od 4,5 km dále, mezi obcemi s elektronickou podatelnou oproti případu, kdybychom vzali rovnoměrně náhodný výběr obcí ČR o stejném rozsahu. Do 4,5 km jsou sledované obce rozmístěny rovnoměrně náhodně. Z grafu párové korelační funkce vyplývá, že podatelny se ve vzdálenostech mezi 2 km a 12,6 km vyskytují častěji a tento rozdíl oproti rovnoměrně náhodnému rozmístění v obcích ČR je větší než v případě Oblasti 1. Nejvyšších hodnot dosahuje funkce v intervalu mezi 2,8 km a 5 km, což odpovídá výsledkům pro Oblast 2.

5 Shrnutí

Po více než 7 letech od začátku platnosti prvního nařízení vlády ukládajícím všem orgánům veřejné moci povinnost zřídit elektronickou podatelnu má tento nástroj zaručené elektronické komunikace v ČR realizováno pouze 15 % úřadů územních samosprávných celků. Z detailního rozboru vyplývá, že procento obcí, které mají elektronickou podatelnu, roste s počtem obyvatel a to velmi významně.

Dále jsme testovali hypotézu, že obce s elektronickou podatelnou jsou rovnoměrně náhodně rozdělené na množině všech obcí ČR. Vlastní test jsme provedli jak pro celou ČR, tak pro dvě dílčí disjunktní oblasti. Pro všechny zkoumané Oblasti (1, 2, 3) hypotézu na 5% hladině testu zamítáme. Z detailního rozboru výsledků sumárních statistik plyne, že v Oblasti 1 se prosazují slabé odpudivé síly na vzdálenostech nad 3,8 km, avšak pro menší vzdálenosti obcí se výsledek blíží rovnoměrně náhodnému rozmístění. V Oblasti 2 jsou patrné nevýrazné shluky, které mohou ukazovat na spolupráci blízkých obcí. Ve všech zkoumaných Oblastech (1, 2, 3) jsou zjištěné odchylky od testované hypotézy nepříliš výrazné.

Tento výsledek neodpovídá záměru zákona č. 500/2004 Sb., správní řád, který zavedl možnost prostřednictvím veřejnoprávní smlouvy zajistit provozování elektronických podatelů pro malé obce obcemi s rozšířenou pů-

sobností. Pokud by se tento postup ve velké míře realizoval, museli bychom v rozmístění obcí s možností příjmu elektronických podání pozorovat výrazné shluky, což se neukázalo. Zdá se dokonce, že v oblasti zajištění provozu elektronických podatelen úřady ÚSC výrazně nespolupracují, neboť na malých vzdálenostech jsou obce s elektronickou podatelnou rovnoměrně náhodně rozmístěné. Tyto výsledky přispívají k vytvoření celkového obrazu stavu rozvoje E-Governmentu v ČR a ukazují, že je třeba hledat i jiné než legislativní stimuly pro jeho praktický rozvoj.

Literatura

- [1] Baddeley A.J., Gill R.D. (1997) *Kaplan-Meier estimators of interpoint distance distributions for spatial point processes*. *Annals of Statistics* **25**, 263–292.
- [2] Lechner T. (2007) *Základní nástroje E-Governmentu v ČR*. Sborník konference Veřejná ekonomika a správa 2007: Inovace ve veřejném sektoru jako faktor rozvoje územních samosprávných celků, Vaňková I. (Eds.). VŠB – Technická univerzita Ostrava, 216–223.
- [3] Mates P., Smejkal V. (2006) *E-Government v českém právu*. Linde Praha, Praha.
- [4] Möller J., Waagepetersen R.P. (2003) *Statistical inference and simulation for spatial point processes*. Chapman & Hall/CRC, Boca Raton.
- [5] Stoyan D., Stoyan H. (1994) *Fractals, random shapes and point fields: methods of geometrical statistics*. John Wiley and Sons, Chichester.
- [6] Český statistický úřad – územně identifikační registr:
http://www.czso.cz/csu/rso.nsf/i/prohlizec_uir_zsj
- [7] Portál veřejné správy (PVS):
http://portal.gov.cz/wps/portal/_s.155/696?kam=epodatelny

Poděkování: Tato práce byla podporována grantem GAAV IAA 101120604. Autoři děkují prof. Viktoru Benešovi a RNDr. Ivanu Saxlovi za cenné připomínky.

Adresa: R. Lechnerová, SVSEŠ, s.r.o., Lindnerova 575/1, 180 00 Praha 8-Libeň

T. Lechner, VŠE v Praze, Národohospodářská fakulta, katedra práva, nám. W. Churchilla 4, 130 67 Praha 3

E-mail: radka.lech@seznam.cz, lechner@triada.cz

SPLINE MODELS WITH CHANGE-POINTS

Matúš Maciak

Keywords: Local polynomial modelling, spline regression, B-splines, P-splines, change-point, jump-detection, semiparametric regression.

Abstract: In this work we will focus on a problem of estimating an unknown regression function based on spline approach, while taking into account also some sudden jumps - change-points. We will discuss some basic methods and their properties for a simple spline approach and afterwards we will extend them and we will implement an algorithm which would allow for change-points occurrences at some certain, pre-defined points - so called knots.

As the knots are given in advance, there is a need to come up with a strategy how to define knot points so that they can reflect real change-points occurrences. We will briefly point at two different strategies which are plausible here and we will introduce a reasonable combination of both of them. The proposed methods are applied to B19 parvovirus data in Belgium population and we have also used a simulation study for additional conclusions.

Abstrakt: V tejto práci sa zameriame na odhad neznámej regresnej funkcie pomocou splinov, pričom budeme uvažovať možnosť výskytu náhlych zmien - change-pointov. Spomenieme základné metódy, ich vlastnosti a následne navrhujeme algoritmus, ktorý implementuje skoky do neznámej regresnej funkcie v predom definovaných bodoch - uzloch.

Keďže sú uzly definované predom, je potrebné navrhnúť stratégiu, ktorá v určitom zmysle optimálne lokalizuje tieto uzly a predstavíme našu metódu, ktorá efektívne kombinuje dve bežne používané stratégie. Výsledný algoritmus aplikujeme na reálne data v belgickej populácii a pre porovnanie a datočné závery použijeme simulačnú štúdiu.

1 Introduction

Let us consider a simple random sample $\{(X_i, Y_i); i = 1, \dots, N \in \mathbb{N}\}$ given from a 2-dimensional space \mathbb{R}^2 where the distribution of each couple (X_i, Y_i) is described by a joint density function $f(x, y)$. We are interested in a dependence structure of the variable Y given the value of the regressor variable X which can be generally written as

$$E[Y|X = x] = m(x),$$

for some unknown function $m(\cdot)$. In case of spline modelling approach we will take a semiparametric point of view where we assume some parametric background for the regression function $m(\cdot)$ but we will pose no restriction on its shape (except some degree of smoothness).

Let us further consider a set of well-defined spline basis functions $\{\psi_j(x); j = 1, \dots, K\}$ of the order $n \in \mathbb{N}$, where each $\psi_j(x)$ is defined over the range of the variable X . In such a case the unknown regression function $m(\cdot)$ is estimated via OLS method as a linear combination given by

$$\widehat{m}(x) = \sum_{j=1}^K \widehat{\theta}_j \psi_j(x), \quad (1)$$

where the set of parameters $\widehat{\Theta} = (\widehat{\theta}_1, \dots, \widehat{\theta}_K)^\top$, also called the basis coefficients are defined by the minimization problem as

$$\widehat{\Theta} = \underset{\Theta \in \mathbb{R}^K}{\text{Argmin}} \left[\sum_{i=1}^N \left(Y_i - \sum_{j=1}^K \theta_j \psi_j(X_i) \right)^2 + \lambda \cdot F \left(\sum_{j=1}^K \theta_j \psi_j(X_i) \right) \right], \quad (2)$$

where $\lambda \in [0, \infty)$ is a smoothing parameter and $F(\cdot)$ is some penalty term which controls the amount of smoothness in the final fit. There are many different settings which are required for spline regression approach (like basis selection, knot points selection, amount of smoothness, etc.) but it is not the main purpose of this work to discuss it here. From now on we will consider a penalty term based on a sum of squared basis coefficients¹, truncated-power spline basis², and a GCV selection approach for parameter $\lambda \in [0, \infty]$.

2 Splines with change-points

Splines are originally defined as piecewise polynomial functions (which brings forward a nice interpretation option in a local sense) which are fully continuous up to the specific order $n \in \mathbb{N}$. However, sometimes it could be convenient to assume some discontinuities (finite number of jumps) while keeping the same property of having a nice interpretation for the final model. We will modify a traditional spline approach in order to accommodate for this new property. We will assume the same regression function as in (1) however, this time we will allow for discontinuity points which is formally written as

$$m(x) = m_0(x) + \sum_{t=1}^T \alpha_t \mathbb{I}_{\{x > x_t\}},$$

for $m_0(\cdot)$ to be $n - 1$ times continuously differentiable and some sequence $x_1 < \dots < x_T$ and non-zero values $\alpha_t \in \mathbb{R}$ for $t = 1, \dots, T$. One can even

¹One can consider many other penalty terms based on basis coefficients or those based on some degree of derivative. However, we have chosen a sum of squared coefficients as it nicely corresponds with the square loss function we have used for fitting of the model.

²The other reasonable choice is to use B-spline basis which is computationally more stable however, one can even show an equivalence between those two types in some sense.

consider a more structured model where we allow for change-points also in all order derivatives of the regression function $m(\cdot)$ given by

$$m^{(p)}(x) = m_0^{(p)}(x) + \sum_{t=1}^{T_p} \alpha_{pt} \mathbb{I}_{\{x > x_{pt}\}}, \quad \text{for some } p \in \{0, 1, \dots, n-1\}.$$

A reasonable way to implement this model into the original spline framework is to relief a continuity condition posed at knot points locations and to assume polynomial pieces which do not have to joint together any more with respect to any level of continuity. This would simply mean that the whole set of change-points $\{\{x_{p1}, \dots, x_{pT_p}\}; p = 0, \dots, n-1\}$ has to be a subset of the mesh of knot points and the set of basis functions is extended by adding some additional functions which can generate all these jumps - change-points.

Let us simplify a little bit and consider a case where every single knot point stands also for a change-point in the regression function itself and also in all its derivatives. Then the whole problem reduces to a simple basis extension, where we add additional functions to the original basis such that

$$\begin{aligned} \{\psi_j(x)\} &= \{\psi_j(x); j = 1, \dots, K = (n+1)(k+1)\} \\ &= \{\psi_{pj}(x); p = 0, 1, \dots, n, j = 0, 1, \dots, k\} \\ &= \left\{ \begin{array}{ccccccc} 1 & x & x^2 & \dots & x^n \\ \mathbb{I}_{\{x > \xi_1\}} & (x - \xi_1)\mathbb{I}_{\{x > \xi_1\}} & (x - \xi_1)^2\mathbb{I}_{\{x > \xi_1\}} & \dots & (x - \xi_1)^n\mathbb{I}_{\{x > \xi_1\}} \\ \dots & \dots & \dots & \dots & \dots \\ \mathbb{I}_{\{x > \xi_k\}} & (x - \xi_k)\mathbb{I}_{\{x > \xi_k\}} & (x - \xi_k)^2\mathbb{I}_{\{x > \xi_k\}} & \dots & (x - \xi_k)^n\mathbb{I}_{\{x > \xi_k\}} \end{array} \right\}, \end{aligned} \quad (3)$$

where $\xi_i, i = 1, \dots, k$ stand for unique inner knot points. With a matrix representation for functions $\{\psi_j(X_i)\}_{j=1}^{(n+1)(k+1)}$ evaluated at each data point $X_i, i = 1, \dots, N$, vector of coefficients $\Theta = (\theta_1, \theta_2, \dots, \theta_{(n+1)(k+1)})^\top$ and the Lagrange multipliers method we can write the solution as

$$\hat{Y} = \mathbb{X}_T (\mathbb{X}_T^\top \mathbb{X}_T + \lambda \mathbf{B})^{-1} \mathbb{X}_T^\top Y, \quad (4)$$

where \mathbb{X}_T is a matrix of basis functions evaluated for a single data point in each row, the matrix \mathbf{B} is a well-known penalty matrix and $\lambda \in [0, \infty)$ is an appropriate smoothing parameter which optimize bias-variance trade-off in some sense. One can easily see that basis functions $\{\psi_{p0}(x), \psi_{nj}(x); p = 0, \dots, n, j = 1, \dots, k\}$ just generate a simple spline regression fit and the other basis functions $\{\psi_{pj}(x); j = 1, \dots, k\}$ allow for change-points occurrences at the exact knot points locations in the p^{th} order derivative of the regression function $m(\cdot)$, for $p = 0, \dots, n-1$. However, this is mostly not the case when we deal with a real problem. On the other hand it easily follows from the basis definition (3) how one can manage change-points occurrences in the regression function $m(\cdot)$ and its derivatives: for example, if we expect some change-points in the function itself only all basis function $\psi_{pj}(x)$ for $j = 1, \dots, k$ and $p = 1, \dots, n-1$ have to be replaced by zero. If we also allow for a change-point in the first and the second derivative then $\psi_{pj}(x) \equiv 0$ for

$j = 1, \dots, k$ and $p = 3, \dots, n - 1$. Of course, we do not want to overparametrize our model and we do not want to include unrealistic change-point into the final fit. The optimal way how to make a selection and to identify basis function which we need to fit the final model will be discussed in Section 4.

2.1 Double penalization method

Before we actually discuss other aspects of spline fitting approach let us give a short comment about the way of penalizing the amount of smoothness.

The most common choice for the matrix \mathbf{B} in the equation (4) is a diagonal matrix $\mathbf{B} = \text{diag}\{0, 0, \mathbf{1}_{(n-1)}^\top, \mathbf{b}_{(1)}^\top, \dots, \mathbf{b}_{(k)}^\top\}$, where $\mathbf{b}_{(j)} = (0_{(1)}, \dots, 0_{(n)}, 1)^\top$. The first two zeros stand for a constant and linear trend - this means that for $\lambda \rightarrow \infty$ all penalized coefficients but these two will be shrunk to zero, resulting in a linear regression line. Of course, if one wants to keep a quadratic trend to be the one with the maximum amount of smoothness, we have to change the corresponding value from one to zero too. All other coefficients which brings additional flexibility into the model yield to penalization. However, as we already have an information about the knot-points locations we do not want to penalize coefficients which stand for change-points neither in order to keep all jumps in the final curve and not to smooth them.

On the other hand, sometimes it might be useful to have a tool to smooth even these jumps and to get a continuous straight line for maximum smoothness effect. However, if we just replace corresponding zeros by ones in the matrix \mathbf{B} all coefficients will be penalized in the same way and the sizes of jumps will shrink to zero quite fast which is not the case we would like.

In order to improve the performance of the fitting algorithm with respect to some model selection criteria (we have used GCV and BIC) we have proposed an extended way of penalizing the basis coefficients, so called *double penalization method*. The idea behind is to have a separate tools to control the amount of smoothness (variability) of the final estimate and the “smoothness” of jumps. In the original setting we have just defined the matrix \mathbf{B} to be a diagonal matrix with zeros and ones on its diagonal (zero for the constant and linear trend and for all change-points coefficients and ones otherwise). As we do not penalize for coefficients responsible for change-points the effect of $\lambda \rightarrow \infty$ will be a set of straight discontinuous lines separated by change-points in the zero and the first derivative. Therefore, we have imposed two independent smoothing parameters λ_1 and λ_2 : the first smoothing parameter will take the same role as in the previous case and the second parameter will penalize basis coefficients responsible for the sizes of jumps. The minimization problem (4) can be rewritten in the following form

$$\hat{\mathbf{Y}} = \mathbb{X}_T (\mathbb{X}_T^\top \mathbb{X}_T + \lambda_1 \mathbf{B}_1 + \lambda_2 \mathbf{B}_2)^{-1} \mathbb{X}_T^\top \mathbf{Y}, \quad (5)$$

where both matrices \mathbf{B}_1 and \mathbf{B}_2 are diagonal matrices with ones for the corresponding coefficients which have to be penalized and zeros otherwise. The

final models provide much better in this setting in case of simulation study and real data problem too (see also Section 5). One can easily think of an obvious extension and to try to implement a multiple penalization where each level of change-points coefficients will be penalized separately. However, the additional time expense we have to accept is much higher in this case and it is not in a balance with a gain we can get any more.

3 Extension to a logistic regression

In order to implement spline approach with change-points also in case of 0-1 discrete type of response we have considered a different approach. The problem behind is that the binary variable which is of interest does not follow normal distribution any more but it comes from a more general family of exponential distributions given by

$$f(y_i) = \exp \left\{ \frac{y_i \zeta_i + b(\zeta_i)}{a(\phi)} + c(y_i, \phi) \right\}, \quad \text{for } i = 1, \dots, N,$$

where ζ_i is a canonical parameter we want to model and ϕ is a dispersion parameter. Functions $a(\cdot)$, $b(\cdot)$ and $c(\cdot)$ are some common real functions where we assume at least an existence of the second derivative of $b(\cdot)$ which is bounded away from zero. The final fit in this case is given in sense of maximizing the likelihood function which is however, not given in explicit way any more. Therefore, we have proposed an iterative algorithm based on Newton-Raphson procedure which can implement penalty term even with respect to double penalization approach³.

Let Θ^* be a vector of coefficients (in the same sense of interpretation as in the previous case) given by the previous step of this iterative procedure. Then the updated value Θ^{new} is given by

$$\Theta^{\text{new}} = \Theta^* - (\mathbf{H}(\Theta^*))^{-1} \cdot \mathbf{S}(\Theta^*), \tag{6}$$

where $\mathbf{S}(\Theta^*)$ is the score matrix and $\mathbf{H}(\Theta^*)$ is the Hessian matrix, both defined by the following expressions

$$\left. \begin{aligned} \mathbf{S}(\Theta^*) &= \mathbb{X}_T^\top \left(\mathbf{Y} - \frac{e^\zeta}{1+e^\zeta} \right) = \mathbb{X}_T^\top \left(\mathbf{Y} - \frac{\exp\{\mathbb{X}_T^\top \Theta^*\}}{1+\exp\{\mathbb{X}_T^\top \Theta^*\}} \right) = 0, \\ \mathbf{H}(\Theta^*) &= (-1) \cdot (\mathbb{X}_T^\top \mathbf{W} \mathbb{X}_T + \lambda \mathbf{B}), \\ \mathbf{H}(\Theta^*) &= (-1) \cdot (\mathbb{X}_T^\top \mathbf{W} \mathbb{X}_T + \lambda_1 \mathbf{B}_1 + \lambda_2 \mathbf{B}_2) \quad \text{respectively.} \end{aligned} \right\} \tag{7}$$

where \mathbb{X}_T is a matrix of basis functions evaluated at each X_i , for $i = 1, \dots, N$ and $\mathbf{W} = \text{diag}(b''(\zeta_1), \dots, \zeta_N)$. The iterative procedure (6) is repeated until convergence, defined as $|\Theta^{\text{new}} - \Theta^*| \leq \epsilon$ for some some small $\epsilon > 0$, where

$$|\Theta^{\text{new}} - \Theta^*| \leq \epsilon \iff \forall_{j=1, \dots, (k+1) \times (n+1)} |\theta_j^{\text{new}} - \theta_j^*| \leq \epsilon \quad \text{and} \quad \exists_j |\theta_j^{\text{new}} - \theta_j^*| < \epsilon.$$

³We have gently modified a penalized quaslikelihood algorithm used for fitting of mixed models, where the variance-covariance matrix of random effects is replaced by a penalization matrix \mathbf{B} .

4 Jump detection and knots positioning

We have already mentioned that the most important task within the spline modelling framework not only with change-points is a positioning of knot points. This time we are not even so much concerned about regular knot points positions (there were many strategies proposed, like equidistant knots, data driven knots based on quantile positions or fully automatic positioning based on minimizing of squared residuals between two neighbouring knots) but mostly we have to propose good locations for all possible jumps. There are actually two different approaches we can consider for splines with change-points: we can either use a full set of basis function defined by (3) for a dense mesh of knots and by implementing an L_1 penalty term for basis coefficients we can let the algorithm to reveal change-points by itself (so called lasso selection approach proposed by Tibshirani [4]) or we can detect possible change-points locations with some advance algorithms and to manually define the set of basis functions we need to fit the model. We have analyzed both of these approaches however, in later text we will just point out the method related to jump detection algorithm⁴.

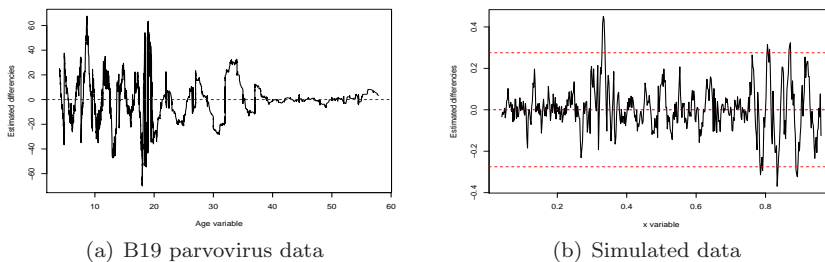


Figure 1: The performance of the jump detection algorithm for the regression function itself for two different settings: real data problem with the large variability and no significant indication of any possible jumps; On the other hand in case of simulated data with a smaller variability we even got the 95 % confidence boundary lines for a change-point indicator.

We will describe a method which brings together Stone’s optimization step with Qiu and Yandell’s localization procedure. The idea behind is to apply firstly the jump detection algorithm and to reveal some small regions inside the span of variable X where we can expect change-points occurrences. With a quite analogous way we can apply the same algorithm in order to find small areas for possible jumps in all order derivatives (see Figure 1).

⁴The most effective method would probably be the one proposed by Stone [3] where all knot points positions are treated as additional model parameters and they also yield to a minimization process. However, such a procedure would be too much time consuming therefore, we have proposed a modification together with a jump detection algorithm proposed by Qiu and Yandell [2].

Once we identify small areas for possible change-points we will optimize their locations via minimizing the GCV criterion by shifting the appropriate knot point in a small grid of points within this area. This means that only knot points related to jumps will yield to minimization, while all other knot points which are not related to any jump will be kept being fixed. This approach is a reasonable compromise between the optimizing of all knot positions proposed by Stone [3] and a manual selection which can provide to poorly.

One can see in Figure 1 that unlike the simulated data, in case of real data sample we did not get any significant indications for change-points regions (compare Figure 1(a) with much larger 95 % confidence bounds (not even in the Figure) and 1(b)). The reason can be seen nicely from Figure 2: the data variability in real data is too much high when comparing with a possible size of a jump. The jump detection algorithm does not have enough power as in case of simulated data where some evidence for discontinuity can be even visible with an eye. Therefore, in case of real data we have just considered an area with the highest peak and strongest oscillation to be an area for possible change-point occurrence.

5 Simulations and real data

Finally, we can take a look how well did new algorithms perform in case of real data (B19 parvovirus antibody level in a blood stream in a sample of Belgium population with a size of 3000 patients) and simulated data (given as a mixture of three normal distributions with a size of 500 points).

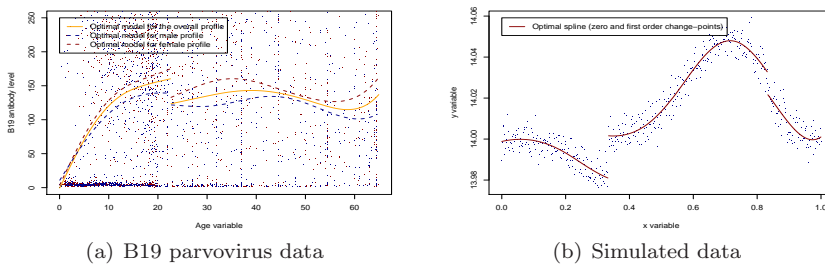


Figure 2: Final spline models for B19 parvovirus data in Belgium and simulated data.

By using the jump detection algorithm we have indicated small interval(s) inside a span of variable X and within these intervals we have found an optimal position for each knot point (by shifting it in a small grid of points) with respect to minimum GCV value. Different models were compared with BIC criterion rather than GCV as GCV criterion quite often tends to overfitting with large sample sizes. In all cases we have implemented a double penalization method with two independent smoothing parameters selected via GCV criterion. The final models can be seen in Figure 2.

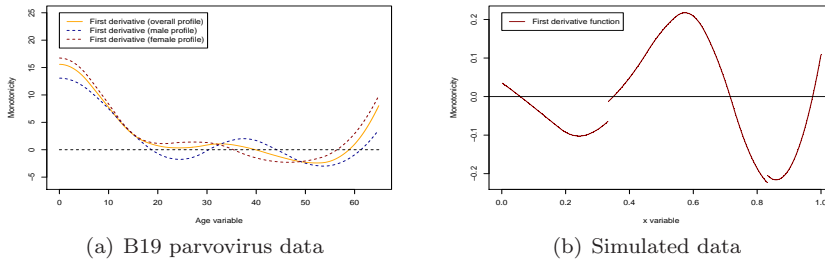


Figure 3: The first order derivative plot for B19 parvovirus and simulated data.

6 Conclusion

We have proposed and applied new methods based on spline approach while allowing for some discontinuities. However, a large variability together with relatively small jump sizes in case of B19 parvovirus data did not show the full potential of the algorithm and the models with no change-points were slightly preferred according to BIC criterion (but not GCV criterion). On the other hand, in case of simulated data where the variability was in a rate with the relative size of jumps the whole methods performed much better and the final model taken by BIC selection criterion was actually the same as the original model where data were simulated from.

References

- [1] Maciak M. (2008). *Spline models with change-points, applied to seroprevalence data*. Master Thesis, Universiteit Hasselt, Diepenbeek, Belgium.
- [2] Qui P. and Yandell B. (1998). *A local polynomial jump detection algorithm in nonparametric regression*. *Technometrics* **40**, 141–152.
- [3] Stone Ch. (1977). *Consistent nonparametric regression*. *The Annals of Statistics* **5**, 595–620.
- [4] Tibshirani R. (1996). *Regression shrinkage and selection via the lasso*. *Journal of Royal Statistical Society* **58**, No. 3, 267–288.

Acknowledgement: The work was supported by grant GAČR 201/09/0775.

Address: M. Maciak, MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8 – Karlín

E-mail: maciak@karlin.mff.cuni.cz

RATIO TYPE STATISTICS FOR DETECTION OF CHANGES IN LINEAR REGRESSION MODELS

Barbora Madurkayová

Keywords: Linear regression, ratio type statistics, change point detection.

Abstract: Linear regression models are relatively frequently used models in statistical analysis, with possible applications e.g. in econometrics, biology or climatology. Here we study procedures for detection of changes in parameters of linear regression models, particularly test procedures based on ratio type statistics.

To the most remarkable advantages of ratio type test statistics belongs the fact, that when computing the test statistics, it is not necessary to estimate variance of the underlying model. This property makes the ratio type statistics a suitable alternative of classical (non-ratio) statistic – most of all in situations, when it is difficult to find a suitable variance estimate.

In this work, we focus on the properties of ratio type test statistics for detection of changes in linear regression models and demonstrate these properties both on simulated and real data.

Abstrakt: Modely lineárnej regresie sú pomerne často používané modely v štatistickej analýze, s možnými aplikáciami napríklad v ekonometrii, biológii či klimatológii. Budeme študovať postupy detekcie zmien v parametroch modelov lineárnej regresie, konkrétne pôjde o postupy založené na štatistických podielového typu.

Medzi najvýraznejšie prednosti štatistík podielového typu patrí skutočnosť, že pri výpočte testovej štatistiky nie je potrebné odhadovať rozptyl skúmaného modelu. Vďaka tejto vlastnosti sú štatistiky podielového typu vhodnou alternatívou klasických (nepodielových) štatistík – predovšetkým v takých situáciách, v ktorých je obtiažne nájsť vhodný odhad rozptylu.

V tejto práci sa zameriame na vlastnosti štatistík podielového typu pre detekciu zmien v modeloch lineárnej regresie a budeme demonštrovať tieto vlastnosti na simulovaných a reálnych dátach.

1 Introduction

We assume to have a set of observations Y_1, \dots, Y_n obtained at time ordered points and that these data follow a linear regression model. Particularly, we are interested in studying a situation, where a change in regression parameters may occur at some time point k^* . We may formally describe such situation as following:

$$Y_k = \mathbf{h}^T(k/n)\boldsymbol{\beta} + \mathbf{h}^T(k/n)\boldsymbol{\delta}I\{k > k^*\} + e_k, \quad k = 1, \dots, n, \quad (1)$$

where $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$, $\boldsymbol{\delta} = \boldsymbol{\delta}_n = (\delta_1, \dots, \delta_p)^T$ and $k^* = k_n^*$ are unknown parameters. $\mathbf{h}^T(t) = (h_1(t), \dots, h_p(t))^T$ is such that $h_1(t) = 1$ for $t \in [0, 1]$ and $h_j(t)$, $j = 2, \dots, p$ are continuously differentiable functions on $[0, 1]$. The random error terms e_1, \dots, e_n are generally assumed to satisfy the so called functional central limit theorem. However, here we are going to focus on a simpler situation, where error terms e_1, \dots, e_n are independent and identically distributed (i.i.d.) random variables, satisfying $E e_k = 0$ and $\text{Var } e_k = \sigma^2 > 0$ for $k = 1, \dots, n$.

The basic question, we are trying to answer, is whether a change in regression parameters occurred at some unknown time point k^* , or not. Using the above introduced notation, the null hypothesis of no change can be expressed as

$$H_0 : k^* = n. \quad (2)$$

We are going to test this null hypothesis against the alternative of change occurring at k^* :

$$H_1 : k^* < n, \boldsymbol{\delta} \neq \mathbf{0}. \quad (3)$$

2 General form of a ratio type test statistic

One possible form of a ratio test statistic that may be used in change-point analysis is the following:

$$\max_{n\gamma \leq k \leq n-n\gamma} \frac{\max_{1 \leq j \leq k} V_{j,k}}{\max_{k \leq j \leq n} \tilde{V}_{j,k}}, \quad (4)$$

where $V_{j,k}$ for $j = 1, \dots, k$ denotes a statistic based on observations Y_1, \dots, Y_k and $\tilde{V}_{j,k}$ for $j = k+1, \dots, n$ is a similar statistic based on observations Y_{k+1}, \dots, Y_n .

This type of statistic was studied for example in [3] or [2], where a ratio statistic of two CUSUM statistic was used for detection of abrupt changes in location. A report about a ratio statistic that may be used for detection of gradual changes in location may be found in [5].

The basic motivation for studying ratio type test statistics lies in the fact that when computing such test statistic, it is not necessary to estimate variance of the underlying model. This property makes the ratio type statistics a suitable alternative of classical (non-ratio) statistic – most of all in situations, when it is difficult to find a suitable variance estimate.

3 Ratio type test statistic for detection of change in regression parameters

For situation described above, test statistics based on weighted partial sums of residuals are often used, i.e. statistics of the form

$$\mathbf{S}_k = \sum_{i=1}^k \mathbf{h}(i/n) (Y_i - \mathbf{h}^T(i/n)\mathbf{b}_n), \quad k = 1, \dots, n, \quad (5)$$

where \mathbf{b}_n is an L_2 -estimate of parameter of the regression parameter based on observations Y_1, \dots, Y_n .

Let us denote

$$\mathbf{S}_{j,k} = \sum_{i=1}^j \mathbf{h}(i/n) (Y_i - \mathbf{h}^T(i/n)\mathbf{b}_k), \quad j, k = 1, \dots, n, \quad j \leq k. \quad (6)$$

Here, \mathbf{b}_k denotes an L_2 -estimate of of the regression parameter based on observations Y_1, \dots, Y_k .

And, similarly

$$\tilde{\mathbf{S}}_{j,k} = \sum_{i=k+1}^j \mathbf{h}(i/n) (Y_i - \mathbf{h}^T(i/n)\tilde{\mathbf{b}}_k), \quad j, k = 1, \dots, n, \quad j > k \quad (7)$$

where $\tilde{\mathbf{b}}_k$ is an L_2 -estimate of the regression parameter based on observations Y_{k+1}, \dots, Y_n .

Further we denote

$$\mathbf{C}_{j,k} = \sum_{i=j+1}^k \mathbf{h}(i/n)\mathbf{h}^T(i/n), \quad j, k = 1, \dots, n. \quad (8)$$

Using this notation, we may now define the ratio type test statistic

$$U_n = \max_{n\gamma \leq k \leq n-n\gamma} \frac{\max_{1 \leq j \leq k} \mathbf{S}_{j,k}^T \mathbf{C}_{1,k}^{-1} \mathbf{S}_{j,k}}{\max_{k \leq j \leq n} \tilde{\mathbf{S}}_{j,k}^T \mathbf{C}_{k+1,n}^{-1} \tilde{\mathbf{S}}_{j,k}}. \quad (9)$$

4 Asymptotic properties

Theorem 1 (Null hypothesis) Let us assume that Y_1, \dots, Y_n follow model (1) and that the null hypothesis (2) described above is true. Let us further assume that e_1, \dots, e_n are i.i.d. random variables, satisfying $E e_k = 0$ and $\text{Var } e_k = \sigma^2 > 0$ for $k = 1, \dots, n$ and that $\mathbf{h}^T(t) = (h_1(t), \dots, h_p(t))^T$ is such that $h_1(t) = 1$ for $t \in [0, 1]$ and $h_j(t)$, $j = 2, \dots, p$ are continuously differentiable functions on $[0, 1]$ with $\int_0^1 h_j(t) dt = 0$, $j = 2, \dots, p$.

Then, as $n \rightarrow \infty$, U_n converges in distribution to

$$U = \sup_{\gamma \leq t \leq 1-\gamma} \frac{\sup_{0 \leq s \leq t} \mathbf{S}^T(s, t) \mathbf{C}(0, t)^{-1} \mathbf{S}(s, t)}{\sup_{t \leq s \leq 1} \tilde{\mathbf{S}}^T(s, t) \mathbf{C}(t, 1)^{-1} \tilde{\mathbf{S}}(s, t)}. \tag{10}$$

where

$$\mathbf{C}(s, t) = \lim_{n \rightarrow \infty} \frac{1}{n} \mathbf{C}_{[ns], [nt]}, \tag{11}$$

$$\mathbf{S}(s, t) = \int_0^s \mathbf{h}(x) dB_1(x) - \mathbf{C}(0, s) \mathbf{C}(0, t)^{-1} \int_0^t \mathbf{h}(x) dB_1(x), \tag{12}$$

for $s, t \in [0, 1]$ such that $s \leq t$,

$$\tilde{\mathbf{S}}(s, t) = \int_t^s \mathbf{h}(x) dB_2(x) - \mathbf{C}(t, s) \mathbf{C}(t, 1)^{-1} \int_t^1 \mathbf{h}(x) dB_2(x). \tag{13}$$

for $s, t \in [0, 1]$ such that $t \leq s$ and where $\{B_1(u), 0 \leq u \leq 1\}$ and $\{B_2(u), 0 \leq u \leq 1\}$ are independent Brownian bridges.

Proof. Using the arguments of the proof of Theorem 2.1 in [4], for $k = [nt]$ we get that

$$\frac{1}{\sigma^2} \max_{1 \leq j \leq k} \left\{ \mathbf{S}_{j,k}^T \mathbf{C}_{1,k}^{-1} \mathbf{S}_{j,k} \right\} \xrightarrow{d} \sup_{0 \leq s \leq t} \left\{ \mathbf{S}^T(s, t) \mathbf{C}(0, t)^{-1} \mathbf{S}(s, t) \right\}, \text{ as } n \rightarrow \infty \tag{14}$$

and similarly, for $k = [nt]$

$$\frac{1}{\sigma^2} \max_{k \leq j \leq n} \left\{ \tilde{\mathbf{S}}_{j,k}^T \mathbf{C}_{k+1,n}^{-1} \tilde{\mathbf{S}}_{j,k} \right\} \xrightarrow{d} \sup_{t \leq s \leq 1} \left\{ \tilde{\mathbf{S}}^T(s, t) \mathbf{C}(t, 1)^{-1} \tilde{\mathbf{S}}(s, t) \right\}, \text{ as } n \rightarrow \infty, \tag{15}$$

where by the symbol \xrightarrow{d} we denote convergence in distribution. Since both statistics are independent for any relevant choice of k , it directly follows that

$$U_n \xrightarrow{d} U, \text{ as } n \rightarrow \infty. \tag{16}$$

□

The next theorem describes a situation under the local alternative.

Theorem 2 (Local alternative) Let us suppose that Y_1, \dots, Y_n follow model (1) and the alternative (3) and that the same assumptions about random error terms and $\mathbf{h}(t)$ as in Theorem 1 hold. Let us further assume that $k^* = [nt]$ for some $0 < t < 1$ and that $\boldsymbol{\delta}_n$ satisfies the conditions

$$\|\boldsymbol{\delta}_n\| \rightarrow 0 \text{ and } n^{1/2} \|\boldsymbol{\delta}_n\| \rightarrow \infty. \tag{17}$$

Then, for $\gamma < t < 1 - \gamma$:

$$U_n \xrightarrow{P} \infty, \text{ as } n \rightarrow \infty. \tag{18}$$

Proof. Let us take $k = \lfloor nr \rfloor$ for some $t < r < 1 - \gamma$. Then by Theorem 2.2 in [4] it follows that

$$\max_{1 \leq j \leq k} \left\{ \mathbf{S}_{j,k}^T \mathbf{C}_{1,k}^{-1} \mathbf{S}_{j,k} \right\} \xrightarrow{P} \infty.$$

For $k = \lfloor nr \rfloor$, the denominator in (9) has the same distribution as under the null hypothesis, and is therefore bounded in probability. It follows that the maximum of the ratio also has to go to infinity, as $n \rightarrow \infty$.

This proves the consistency of the studied test statistic under the given assumptions. \square

5 Applications

According to Theorem 1, U_n converges to a limit distribution under the null hypothesis. The explicit form of the limit distribution is not known. The critical values may be determined by simulation of the limit distribution or by bootstrap method. Since by Theorem 2 U_n converges in probability to infinity, we are going to reject the null hypothesis for large values of the test statistic.

5.1 Simulated data

In this section we present some applications of the proposed ratio statistic to simulated data from normal and Laplace distributions. In simulation, we took $p = 3$, $h_1(x) = 1$, $h_2(x) = x - 1/2$ and $h_3(x) = 4x^2 - 4x + 2/3$, $\gamma = 0.1$ and several choices of $\boldsymbol{\delta}$.

On Figure 1 and 2, we can see the values of ratio

$$Q_k = \frac{\max_{1 \leq j \leq k} \mathbf{S}_{j,k}^T \mathbf{C}_{1,k}^{-1} \mathbf{S}_{j,k}}{\max_{k \leq j \leq n} \tilde{\mathbf{S}}_{j,k}^T \mathbf{C}_{k+1,n}^{-1} \tilde{\mathbf{S}}_{j,k}}, \quad (19)$$

computed for $k : n\gamma \leq k \leq n - n\gamma$ with $\gamma = 0.1$. Simulated 95% critical values for each of the two distributions are depicted by the horizontal dotted line.

If we compare several choices of $\boldsymbol{\delta}$ with the same norm, we may see, that the method is more sensitive to changes in parameters corresponding to the constant and linear element of $\mathbf{h}(x)$, rather than the quadratic one. We may also see, that under the alternative, the values of Q_k tend to increase with k and the maximum value is obtained for k close to n .

5.2 Surface temperature data

The analyzed data come from a large dataset based on long term surface temperature measurements at several meteorological stations around the world (for more details see [6], dataset HadCRUT3). On Figure 4, we may see the

δ	N(0,1)	Lap(0,1)	N(0,1)	Lap(0,1)
	90%	90%	95%	95%
(0,1,0)	0.130	0.086	0.075	0.039
(0,0,1)	0.117	0.089	0.069	0.032
(0,1,1)	0.171	0.098	0.105	0.045
(1,1,1)	0.568	0.485	0.450	0.335
(0,3,0)	0.462	0.356	0.335	0.206
(0,0,3)	0.163	0.081	0.101	0.034
(0,3,3)	0.713	0.619	0.576	0.420
(3,3,3)	1.000	1.000	0.999	0.996
(0,5,0)	0.858	0.750	0.757	0.588
(0,0,5)	0.259	0.173	0.154	0.088
(0,5,5)	0.986	0.960	0.959	0.890
(5,5,5)	1.000	1.000	1.000	1.000

Table 1: Simulated 90% and 95% rejection rates of U_{100} ; $\gamma = 0.1$, $k^* = n/2 = 50$.

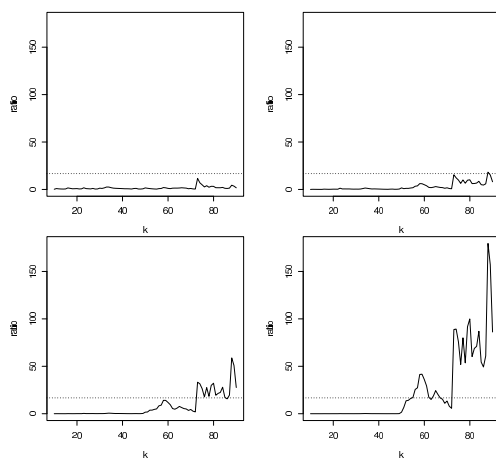


Figure 1: The values of Q_k for simulated normal distribution samples with parameters $\mu = 0$, $\sigma = 1$, $n = 100$, $\gamma = 0.1$. The upper left figure refers to the null hypothesis. The other figures refer to alternatives with $k^* = n/2 = 50$ and $\delta = (0, 0, \sqrt{27})$ (upper right), $\delta = (0, \sqrt{27}, 0)$ (lower left), $\delta = (3, 3, 3)$ (lower right).

data together with estimated regression curves. The data represent temperature anomalies, i.e. differences from what is expected to be measured in some particular area at some particular time of the year. Each observation corresponds to monthly measurements at the chosen area located in the South Pacific Ocean, close to New Zealand (the center of the 5×5 degree area is located at 177.5W and 32.5S). The data covers the period of years 1947-1987.

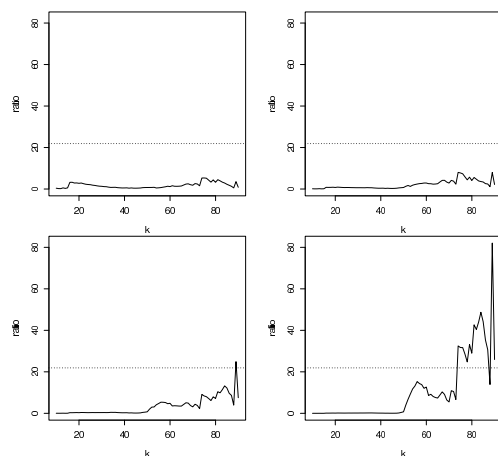


Figure 2: The values of Q_k for simulated Laplace distribution samples with parameters $\mu = 0$, $b = 1$, $n = 100$, $\gamma = 0.1$. The upper left figure refers to the null hypothesis. The other figures refer to alternatives with $k^* = n/2 = 50$ and $\delta = (0, 0, \sqrt{27})$ (upper right), $\delta = (0, \sqrt{27}, 0)$ (lower left), $\delta = (3, 3, 3)$ (lower right).

Again, we took $p = 3$, $h_1(x) = 1$, $h_2(x) = x - 1/2$ and $h_3(x) = 4x^2 - 4x + 2/3$, $\gamma = 0.1$. After applying the method described above we reject null hypothesis of no change. The values of the ratio Q_k defined by (19) are to see on Figure 3. We estimate the time of change k^* by maximizing the numerator in (9):

$$\hat{k}^* = \arg \max_k S_{j,k}^T C_{1,k}^{-1} S_{j,k}. \tag{20}$$

The estimates of the regression parameters can be then obtained as

$$\mathbf{b}_{\hat{k}^*} = C_{1,\hat{k}^*}^{-1} \sum_{i=1}^{\hat{k}^*} \mathbf{h}(i/n) Y_i, \quad \tilde{\mathbf{b}}_{\hat{k}^*} = C_{\hat{k}^*+1,n}^{-1} \sum_{i=\hat{k}^*+1}^n \mathbf{h}(i/n) Y_i \tag{21}$$

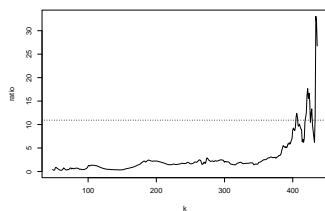


Figure 3: The values of Q_k for surface temperature data. The simulated critical value is depicted by the dotted line.

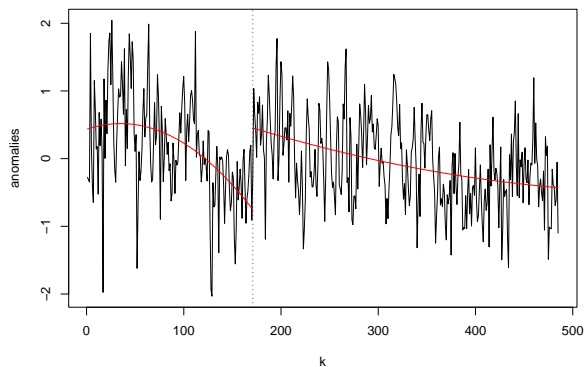


Figure 4: The surface temperature data. Estimated regression curves are depicted by the red lines.

References

- [1] Aue, A., Horváth, Z., Hušková, M., Kokoszka, P. (2008). Testing for changes in polynomial regression. *Bernoulli* 14:637–660.
- [2] Horváth, L., Horváth, Z., and Hušková, M. (2008). Ratio tests for change point detection. *Beyond Parametrics in Interdisciplinary Research: Festschrift in Honor of Professor P. K. Sen*, 1:293–304.
- [3] Hušková, M. (2007). Ratio type test statistics for detection of changes in time series. *Bulletin of the International Statistical Institute, Proceedings of the 56th Session, Lisboa 2007*, (976).
- [4] Hušková, M. and Picek, J. (2005). Bootstrap in detection of changes in linear regression. *The Indian Journal of Statistics Special Issue on Quantile Regression and Related Methods*, 67:200–226.
- [5] Madurkayová, B. (2007). Ratio tests for gradual changes. *Proceedings of WDS 2007, Prague*.
- [6] Met Office Hadley Centre observations datasets. HadCRUT3. Available from: <http://www.metoffice.gov.uk/research/hadleycentre/obsdata/>.

Acknowledgement: The work was supported by grant GAČR 201/05/H007 and GAUK B-MAT 116209.

Address: B. Madurkayová, MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8 – Karlín

E-mail: madurka@karlin.mff.cuni.cz

TAJNÝ OSTROV TEMPLÁŘŮ

(Příběh z cyklu „Statistický pohled na měření v české a světové literatuře“)

Jaroslav Marek

Klíčová slova: Dvouetapový regresní model, podmínka typu II, BLUE.

Abstrakt: Článek je věnován regresnímu modelu s podmínkou typu II. Cílem je ukázat aplikační možnosti teorie odhadu. Ústřední téma knihy Tajný ostrov templářů, rytíři, kněz a poklad autorů Erlinga Haagensena a Henry Lincolna je poloha kostelů na ostrově Bornholm, které mají být postaveny v posvátném útvaru a mají vytvářet pěticípou hvězdu, jejíž střed tvoří kostel Osterlands. Cílem článku je najít odhady souřadnic všech kostelů za předpokladu, že byly postaveny v posvátné konfiguraci.

Abstract: The article deals with regression model with constraints of type II. The aim is to show the application possibilities of a particular theory of parameter estimation. The central theme in the book The Templars' Secret Island, the Knights, the Priest and the Treasure by Erling Haagenzen and Henry Lincoln, seems to be the construction of many Bornholm's churches according to sacred geometry, most notably the five pointed star centered on the Østerlars church. The aim of article is make the estimator of the coordinates of all churches under the constraints that all these points are located on a places according to the sacred geometry.

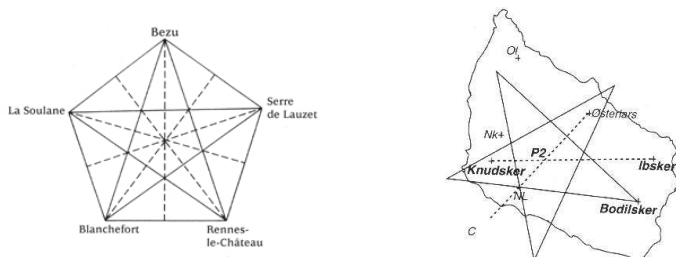
1 Mystické příběhy

Čím dál častěji se můžeme setkat s různými úžasnými objevy a teoriemi. Kniha [1] je typickým příkladem. Autoři Erling Haagenzen a Henry Lincoln v ní seznámí čtenáře s mystickou teorií, při jejíž konstrukci provedou čtenáře Egyptem, Palestinou, Francií a konečně po ostrově Bornholm. Čtenář je seznámen s tajemstvím pyramid, ukrývající číslo π a délku poledníku. Dále je seznámen s údajně hlavním úkolem templářů, kterým má být uchovat a ukrýt to, co našli rytíři Bernarda z Clairvauz v r. 1127 v Jeruzalémě. To kvůli čemu neváhají členové tajné organizace Opus Dei v knize D. Browna Šifra Mistra Leonarda vraždit — Svatý grál.

Autoři Haagenzen a Lincoln čtenáře zavedou do Jeruzaléma dobytého v r. 586 králem Nabukadnazarem a v r. 1099 křižáckou výpravou zorganizovanou na popud papeže Urbana II. Dále sdělují čtenáři, že templáři prokazatelně našli Svatý grál v Jeruzalému v r. 1127, kdy podali Sv. Bernardovi zprávu „poslání splněno“. Nový úkryt, pro vše co našli, měl být vybudován ve Francii ve středu pentagramu (viz obrázek 1 vlevo). Ovšem v r. 1307 král

Filip Sličný nechává templáře uvěznit a zavraždit. V r. 1314 je upálen velmistr Jacques de Molay. Ale v té době je to, co našli, již ukryto na jednom malém dánském ostrově Bornholm — na místě, které lze vypočítat z polohy kostelů (viz obrázek 1 vpravo).

Autoři knihy zjišťují, že stavitelé bornholmských kostelů si přáli sestavit kruh o obvodu 56 anglických mil. To se jim podařilo — podle autorů knihy — s přesností 0,4 mm. Úhly v pentagramu templáři vytyčili s přesností 0,00000015 stupně. Čtenář je nakonec seznámen s tzv. geometrickým klíčem. Ten nám umožní sestavit podmínky a regresní model, když úloha povede na dvouetapový regresní model s podmínkou typu II.



Obrázek 1: Mystické rozmístění templářských kostelů v jižní Francii a na ostrově Bornholm.

Při studiu polohy bornholmských kostelů autoři zjišťují překvapivé skutečnosti o templáři a o jejich vědomostech, ke kterým by se měli vyslovit vědci různých vědních disciplín, když už zástupci církve mohou jejich tvrzení jen těžko vyvracet.

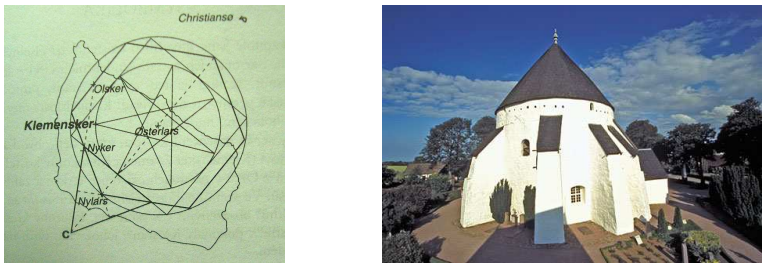
Matematici by mohli zpochybnit fakta o schopnostech templářských rytířů provést trisekci úhlu, přesně zkonstruovat odmocniny z čísla 2, 3, 5 a 7. Architekti by mohli posoudit, zda je pravda, že kostel Nylars je postaven jako přesná kopie Božího hrobu v Jeruzalémě. Historici by se mohli věnovat tvrzením o působení templářů na Bornholmu. Jazykovědci by mohli posoudit tvrzení, že slovo Bornholm je odvozeno od Burgundska.

K dovednostem templářů provést trisekci úhlu nebo vytyčit vzdálenost danou jistým číslem násobeným podílem odmocniny ze sedmi a ze tří, a to ještě po moři, se nebudeme vyslovovat. Hlavním cílem tohoto článku bude pouze posoudit tvrzení o konfiguraci kostelů na ostrově Bornholm.

Z jednoho z mnoha schémat v [1] nakonec vyplyne, že kostely nejsou situovány ve vrcholech pěticípé hvězdy — ale v bodech na obr. 2. Čtenář může být zklamán — pokud si náčrtku povšimne.

2 Vytvoření regresního modelu

Autoři uvádí (viz schéma na obr. 2 vlevo) následující skutečnosti (S představuje střed mystické konfigurace — kostel Østerlars):



Obrázek 2: Posvátná konfigurace templářských kostelů na ostrově Bornholm a kostel Østerlars.

- 1) kostel Nylars se nachází na kružnici k se středem $S = [\gamma_1, \gamma_2]$ a s poloměrem γ_3 ,
- 2) kostel Olsker se nachází na kružnici se středem $S = [\gamma_1, \gamma_2]$ a s poloměrem $\Theta_1^2 \times \gamma_3 = \sqrt{\left(\frac{3}{2} - \sqrt{2}\right)^2 + \left(\frac{\sqrt{2}}{2}\right)^2} \times \gamma_3$,
- 3) kostel Nyker se nachází na kružnici se středem S a s poloměrem $\Theta_2^2 \times \gamma_3 = \sqrt{\left(-\frac{\sqrt{7}}{4} + \frac{3}{2} - \sqrt{2}\right)^2 + \left(\frac{\sqrt{3}}{2} - \frac{4\sqrt{7}}{3}\right)^2} \times \gamma_3$,
- 4) kostel Klemensker se nachází na kružnici se středem $S = [\gamma_1, \gamma_2]$ a s poloměrem $\Theta_3^2 \times \gamma_3 = \frac{1}{\sqrt{2}} \times \gamma_3$,
- 5) kostel Rutsker se nachází na kružnici k .

Autoři knihy uvádí, že mají k dispozici souřadnice kostelů změřené dánským kartografickým úřadem Kort & Matrikelstyrelsen. Údaje v tabulce jsou uvedeny v metrech.

bod	Nylars	Olsker	Nyker	Klemensker	Rutsker	Østerlars
X	54 748.37	55 653.27	57 665.44	55 525.51	58 866.14	45 385.33
Y	45 803.24	63 844.61	53 086.37	57 063.72	61 535.90	56 658.79

Pro výpočet použijeme dvouetapový regresní model (viz str. 258), ten nám umožňuje najít odhady parametrů β a γ splňující podmínky. Dále tento model neprovede korekci vektorového parametru Θ , nalezené odhady ale budou respektovat nejistoty v určení toho parametru určené maticí $\Sigma_{1,1} = \text{Diag}(\omega_1^2, \omega_1^2, \omega_1^2)$. Ve druhé etapě budeme uvažovat kovarianční matici $\Sigma_{2,2} = \text{Diag}(\omega_2^2, \dots, \omega_2^2)$ o rozměru 10×10 .

Uvažujme, že hodnoty Θ_1, Θ_2 a Θ_3 byly templáři určeny a změřeny (v 1. etapě) jako $(\hat{\Theta}_1, \hat{\Theta}_2, \hat{\Theta}_3) = (0,870263933, 0,890359258, 0,707106781)$, což jsou hodnoty použité autory při výpočtu. Vzhledem k tvrzení autorů budeme uvažovat přesnost určení těchto hodnot $\omega_1^2 = 0.0001^2$.

V druhé etapě měření budeme uvažovat disperzi $\omega_2^2 = 25^2 \text{ m}^2$ pro měřené souřadnice $(y_1, y_2), \dots, (y_9, y_{10}) = (54748.37, 45803.24, 55653.27, 63844.61, 57665.44, 53086.37, 55525.51, 57063.27, 58866.14, 61535.30)$ kostelů Nylars,

Olsker, Nyker, Klemensker a Rutsker s rovinnými souřadnicemi $(\beta_{2i-1}, \beta_{2i})$ pro $i = 1, 2, 3, 4, 5$.

Podmínky lze zapsat takto

$$\begin{aligned} g_1(\Theta, \beta, \gamma) &= (\beta_1 - \gamma_1)^2 + (\beta_2 - \gamma_2)^2 - \gamma_3^2 &= 0, \\ g_2(\Theta, \beta, \gamma) &= (\beta_3 - \gamma_1)^2 + (\beta_4 - \gamma_2)^2 - \Theta_1^2 \times \gamma_3^2 &= 0, \\ g_3(\Theta, \beta, \gamma) &= (\beta_5 - \gamma_1)^2 + (\beta_6 - \gamma_2)^2 - \Theta_2^2 \times \gamma_3^2 &= 0, \\ g_4(\Theta, \beta, \gamma) &= (\beta_7 - \gamma_1)^2 + (\beta_8 - \gamma_2)^2 - \Theta_3^2 \times \gamma_3^2 &= 0, \\ g_5(\Theta, \beta, \gamma) &= (\beta_9 - \gamma_1)^2 + (\beta_{10} - \gamma_2)^2 - \gamma_3^2 &= 0. \end{aligned}$$

Získané podmínky $\mathbf{g}(\Theta, \beta, \gamma)$ budeme linearizovat. Užijeme Taylorův rozvoj — lineární verze podmínky $\mathbf{g}(\Theta, \beta, \gamma) = 0$ je $\mathbf{B}^* \delta\beta + \mathbf{C}^* \delta\Theta + \mathbf{G}\gamma + \mathbf{a}^* = 0$, kde matice $\mathbf{B}^* = \frac{\partial \mathbf{g}(\Theta^0, \beta^0, \gamma^0)}{\partial \beta'}$, $\mathbf{C}^* = \frac{\partial \mathbf{g}(\Theta^0, \beta^0, \gamma^0)}{\partial \Theta'}$, $\mathbf{G} = \frac{\partial \mathbf{g}(\Theta^0, \beta^0, \gamma^0)}{\partial \gamma'}$, $\mathbf{a}^* = \mathbf{g}(\Theta^0, \beta^0, \gamma^0)$ v přibližném bodě $(\Theta^0, \beta^0, \gamma^0)$.

V našem případě např.

$$\mathbf{G}^0 = \begin{pmatrix} 2(\gamma_1^0 - \beta_1^0), & 2(\gamma_2^0 - \beta_2^0), & -2\gamma_3^0 \\ 2(\gamma_1^0 - \beta_3^0), & 2(\gamma_2^0 - \beta_4^0), & -2\gamma_3^0(\Theta_1^0)^2 \\ 2(\gamma_1^0 - \beta_5^0), & 2(\gamma_2^0 - \beta_6^0), & -2\gamma_3^0(\Theta_2^0)^2 \\ 2(\gamma_1^0 - \beta_7^0), & 2(\gamma_2^0 - \beta_8^0), & -2\gamma_3^0(\Theta_3^0)^2 \\ 2(\gamma_1^0 - \beta_9^0), & 2(\gamma_2^0 - \beta_{10}^0), & -2\gamma_3^0 \end{pmatrix}.$$

Matice \mathbf{C}^* a \mathbf{B}^* nebudeme podrobně uvádět.

Dvouetapový model připojovacího měření je

$$\begin{pmatrix} \hat{\Theta} \\ \mathbf{Y} - \mathbf{D}\hat{\Theta} \end{pmatrix} \sim_n \left(\begin{pmatrix} \Theta \\ \mathbf{X}\beta \end{pmatrix}, \begin{pmatrix} \Sigma_{1,1}, & -\Sigma_{1,1}\mathbf{D}' \\ -\mathbf{D}\Sigma_{1,1}, & \Sigma_{2,2} + \mathbf{D}\Sigma_{1,1}\mathbf{D}' \end{pmatrix} \right).$$

Parametrickým prostorem dvouetapového modelu s podmínkou typu II je

$$\underline{\Theta} = \{(\Theta', \beta') : \mathbf{B}^* \beta + \mathbf{C}^* \Theta + \mathbf{G}\gamma + \mathbf{a}^* = \mathbf{0}\},$$

kde \mathbf{B}^* , \mathbf{C}^* , \mathbf{G} jsou dané matice $q \times k_2$, $q \times k_1$, $q \times k_3$ a kde \mathbf{a}^* je daný vektor s délkou q , tak, že $\mathcal{M}(\mathbf{C}^*) \subset \mathcal{M}(\mathbf{B}^*)$, $\mathbf{a}^* \in \mathcal{M}(\mathbf{C}^*)$ a $r(\mathbf{B}^*) = q < k_2$.

Z první etapy měření máme k dispozici nevychýlený odhad $\hat{\Theta}$.

Cílem je určit odhad parametru β na základě náhodného vektoru $\mathbf{Y} - \mathbf{D}\hat{\Theta}$, kde \mathbf{Y} je observační vektor z druhé etapy měření na základě odhadu $\hat{\Theta}$.

BLUE parametru $(\beta', \gamma')'$ v našem modelu je (viz [3])

$$\begin{aligned}
\hat{\beta} &= \left(\mathbf{I} - (\mathbf{X}'\Sigma_{\hat{\theta}}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' \left\{ [\mathbf{B}^*(\mathbf{X}'\Sigma_{\hat{\theta}}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}']^{-1} - \right. \right. \\
&\quad \left. \left. - [\mathbf{B}^*(\mathbf{X}'\Sigma_{\hat{\theta}}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}']^{-1} \mathbf{G} \left\{ \mathbf{G}' [\mathbf{B}^*(\mathbf{X}'\Sigma_{\hat{\theta}}^{-1}\mathbf{X})^{-1} \times \right. \right. \right. \\
&\quad \left. \left. \left. \times (\mathbf{B}^*)' + \mathbf{G}\mathbf{G}' \right\}^{-1} \mathbf{G} \right\}^{-1} \mathbf{G}' [\mathbf{B}^*(\mathbf{X}'\Sigma_{\hat{\theta}}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}']^{-1} \right\} \mathbf{B}^* \right) \times \\
&\quad \times (\mathbf{X}'\Sigma_{\hat{\theta}}^{-1}\mathbf{X})^{-1} \mathbf{X}'\Sigma_{\hat{\theta}}^{-1}(\mathbf{Y} - \mathbf{D}\Theta) - \\
&\quad - (\mathbf{X}'\Sigma_{\hat{\theta}}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' \left\{ [\mathbf{B}^*(\mathbf{X}'\Sigma_{\hat{\theta}}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}']^{-1} - \right. \\
&\quad \left. - [\mathbf{B}^*(\mathbf{X}'\Sigma_{\hat{\theta}}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}']^{-1} \mathbf{G} \left\{ \mathbf{G}' [\mathbf{B}^*(\mathbf{X}'\Sigma_{\hat{\theta}}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' + \right. \right. \\
&\quad \left. \left. \left. + \mathbf{G}\mathbf{G}' \right\}^{-1} \mathbf{G} \right\}^{-1} \mathbf{G}' [\mathbf{B}^*(\mathbf{X}'\Sigma_{\hat{\theta}}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}']^{-1} \right\} (\mathbf{a}^* + \mathbf{C}^*\Theta), \\
\hat{\gamma} &= - \left\{ \mathbf{G}' [\mathbf{B}^*(\mathbf{X}'\Sigma_{\hat{\theta}}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}']^{-1} \mathbf{G} \right\}^{-1} \mathbf{G}' \times \\
&\quad \times [\mathbf{B}^*(\mathbf{X}'\Sigma_{\hat{\theta}}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}']^{-1} [\mathbf{B}^*(\mathbf{X}'\Sigma_{\hat{\theta}}^{-1}\mathbf{X})^{-1} \mathbf{X}'\Sigma_{\hat{\theta}}^{-1} \mathbf{Y} + \\
&\quad \quad \quad + \mathbf{a}^* + \mathbf{C}^*\Theta].
\end{aligned}$$

Príslušné kovarianční matice jsou

$$\begin{aligned}
\text{Var}(\hat{\beta}) &= (\mathbf{X}'\Sigma_{\hat{\theta}}^{-1}\mathbf{X})^{-1} - (\mathbf{X}'\Sigma_{\hat{\theta}}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' [\mathbf{B}^*(\mathbf{X}'\Sigma_{\hat{\theta}}^{-1}\mathbf{X})^{-1} \times \\
&\quad \times (\mathbf{B}^*)' + \mathbf{G}\mathbf{G}']^{-1} \mathbf{B}^*(\mathbf{X}'\Sigma_{\hat{\theta}}^{-1}\mathbf{X})^{-1} + (\mathbf{X}'\Sigma_{\hat{\theta}}^{-1}\mathbf{X})^{-1} \times \\
&\quad \times (\mathbf{B}^*)' [\mathbf{B}^*(\mathbf{X}'\Sigma_{\hat{\theta}}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}']^{-1} \times \\
&\quad \times \mathbf{G} \left\{ \mathbf{G}' [\mathbf{B}^*(\mathbf{X}'\Sigma_{\hat{\theta}}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}']^{-1} \mathbf{G} \right\}^{-1} \mathbf{G}' \times \\
&\quad \times [\mathbf{B}^*(\mathbf{X}'\Sigma_{\hat{\theta}}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}']^{-1} \mathbf{B}^*(\mathbf{X}'\Sigma_{\hat{\theta}}^{-1}\mathbf{X})^{-1}, \\
\text{cov}(\hat{\beta}, \hat{\gamma}) &= -(\mathbf{X}'\Sigma_{\hat{\theta}}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' [\mathbf{B}^*(\mathbf{X}'\Sigma_{\hat{\theta}}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}']^{-1} \times \\
&\quad \times \mathbf{G} \left\{ \mathbf{G}' [\mathbf{B}^*(\mathbf{X}'\Sigma_{\hat{\theta}}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}']^{-1} \mathbf{G} \right\}^{-1}, \\
\text{Var}(\hat{\gamma}) &= \left\{ \mathbf{G}' [\mathbf{B}^*(\mathbf{X}'\Sigma_{\hat{\theta}}^{-1}\mathbf{X})^{-1}(\mathbf{B}^*)' + \mathbf{G}\mathbf{G}']^{-1} \mathbf{G} \right\}^{-1} - \mathbf{I}.
\end{aligned}$$

Nyní mimo jiné najdeme odhad souřadnic kostela Østerlars (viz fotografie na obr. 2 vpravo) a porovnáme jej se souřadnicemi z tabulky.

Numerickým výpočtem získáváme odhad

$$\hat{\beta} = (54748.43, 45803.18, 55633.42, 63830.72, 57652.33, 53090.18, 55530.98, 57063.94, 58893.57, 61545.22)' \text{ a } \hat{\gamma} = (45389.41, 56684.33, 14352.36)'.$$

Rovinné souřadnice kostela Østerlars byly změřeny hodnotami $(45385.33, 56658.79)'$, souřadnice určené z našeho regresního modelu jsou $(45389.41, 56684.33)'$. Diference je velmi malá.

3 Závěr

Autor tohoto článku se domníval, že s použitím regresních modelů prokáže nesmyslnost schématu. To se nepodařilo.

Ale při přípravě obrázků bylo zjištěno, že přímka spojující Nylars a Østerlars má probíhat nedalekým ostrovem Christiansø (viz schéma na obr. 2 vlevo). Avšak ze satelitního snímku na obr. 3 vyplývá, že tomu tak rozhodně není — kostely jsou označeny křížkem, Christiansø je nad přímkou v pravém horním rohu obrázku. Autoři si tedy na schématu Christiansø mírně posunuli. Přitom jsou v knize uvedeny i souřadnice bodu na Christiansø, kterým měla přímka procházet. Souřadnice bodu vyhovují jejich výpočtům a zkonstruovaným podmínkám. Je nepochybné, že autoři manipulovali s měřenými údaji, pokud si je rovnou všechny nevymysleli.



Obrázek 3: Ortofotomapa ostrova Bornholm s vyznačením polohy kostelů Nylars a Østerlars.

Možnost najít polohu kostelů na satelitních snímcích umožní čtenářům tohoto článku i jeho autorovi v budoucnosti provést výpočet z ověřených dat.

Literatura

- [1] Haagenen, E., Lincoln, H. (2003) *Tajný ostrov templářů, rytíři, kněz a poklad*. BB art, Praha.
- [2] Kubáček L., Kubáčková, L. (2000) *Statistics and Metrology (in Czech)*. Publishing House of Palacký University.
- [3] Kubáček, L., Marek, J. (2005) *Partial optimum estimator in two stage regression model with constraints and a problem of equivalence*. Math. Slovaca **55**, 477–494.

Poděkování: Článek byl podpořen výzkumným záměrem MSM 6198959214.

Adresa: J. Marek, PŘF UP, KMAaAM, Tomkova 40, 779 00 Olomouc

E-mail: marek@inf.upol.cz

ON THE ESTIMATION OF MUTUAL INFORMATION

Tomáš Marek, Petr Tichavský

Keywords: Mutual information, redundancy, adaptive histogram.

Abstract: The mutual information is useful measure of a random vector component dependence. It is important in many technical applications. The estimation methods are often based on the well known relation between the mutual information and the appropriate entropies. In 1999 Darbellay and Vajda [3] proposed a direct estimation methods. In this paper we compare some available estimation methods using different 2-D random distributions.

Abstrakt: Vzájemná informace je hojně užívanou mírou vzájemné závislosti jednotlivých složek vícerozměrných náhodných vektorů. Časté uplatnění nachází především v inženýrských aplikacích. Metody odhadu vzájemné informace většinou vychází ze známého vztahu mezi vzájemnou informací a entropiemi příslušných rozdělání, ale vzájemnou informaci je možné odhadnout také přímo. Na příkladech různých typů dvojrozměrných rozdělání srovnáme některé dostupné metody odhadu vzájemné informace.

1 Introduction

The mutual information is a very important tool in many engineering applications. Assuming 2-D random vector $(X, Y)^T$ with the joint density function $f_{X,Y}$ and marginal density functions f_X, f_Y , the mutual information $I(X, Y)$ is given as

$$I(X, Y) = \int_{\mathbf{R}^2} f_{X,Y}(x, y) \log \frac{f_{X,Y}(x, y)}{f_X(x)f_Y(y)} dx dy.$$

Traditional estimation methods are based on the well known relation between the mutual information and the appropriate entropies

$$\begin{aligned} I(X, Y) &= H(X) + H(Y) - H(X, Y), \\ &= H(X) - H(X|Y), \\ &= H(Y) - H(Y|X). \end{aligned}$$

In this paper we compare the available algorithms for direct computation of the mutual information with the estimate based on the maximum likelihood introduced by Miller [5]. In addition to the Miller-Madow's method we employ three methods described in the next section. All these methods of the direct computation of the mutual information are based on the 2-D histogram. The computer simulations bellow show the advantage of algorithms based on an adaptive histogram.

2 Methods and algorithms

2.1 Adaptive histogram methods

The adaptive histogram methods introduced by Darbellay [2] reach good efficiency. In this paper we use the algorithm published by Darbellay and Vajda [3]. The histogram generating process is based on the partitioning of the observation space into a finite number of nonoverlapping rectangular cells C_k , $1 \leq k \leq m$. The cells are generated by the recursive process described below.

Algorithm Mutin A

- (i) The initial (the largest) cell is the smallest rectangular cell containing all data pairs $(X, Y)^T$.
- (ii) Any cell containing less than two observations (data pairs) will not be partitioned.
- (iii) Every cell containing at least two observations is tentatively partitioned by dividing each one of its edges into two equiprobable halves. It means four new cells are tentatively generated instead given 'mother' cell.
- (iv) Assume that the partitioned 'mother' cell contains $N \leq 2$ observations. The new generated cells contain N_1, N_2, N_3 and N_4 observations. The partitioning is accepted if

$$T = \frac{4}{N} \sum_{i=1}^4 \left(N_i - \frac{N}{4} \right)^2 > \chi_3^2(0.95) = 7.81, \quad (1)$$

where T is the goodness of fit statistic T intuitively testing the local independence of marginals at this 'mother' cell. If $T < 7.81$ then the tentative partition is refused and the 'mother' cell is admitted to the final computation.

- (v) After stopping all (local) partitioning processes we denote generated cells C_k , $1 \leq k \leq m$ and the corresponding numbers of observations N_k . The estimate of the unknown mutual information $I(X, Y)$ we define as

$$\hat{I}_N(X, Y) = \sum_{k=1}^m \frac{N_k}{N} \log \frac{N_k/N}{(N_{x,k}/N)(N_{y,k}/N)}, \quad (2)$$

where $N_{x,k}$ is the number of observations that have the same x coordinate as observations in the cell C_k (analogically the $N_{y,k}$).

The computer simulation studies, e.g. Franěk [4], show that true expectation of the goodness of fit variable T defined in (1) is not constant but $ET \leq 3$. Keeping notation of algorithm *Mutin A*, it approximately holds

$$ET \approx 3 - \frac{N_{x,k}}{N_k} - \frac{N_{y,k}}{N_k}.$$

It follows that we should adjust the critical value of the goodness-of-fit test via (1). This adjustment decreases absolute value of the negative part of estimation bias, because the hypothesis of local independence is rejected in more cases and the partitioning process generates more cells with some positive contribution to the mutual information estimate (2). On the other hand the positive part of estimation bias slightly increases because of increasing the first type error probability in the test (1) which is not 0.05, but is lower in the algorithm *Mutin A*. The algorithm with adjusted critical values of the goodness-of-fit test will be called *Mutin B*.

2.2 Fixed histogram method

First we consider estimates using histograms based on a fixed partition of the observation space. There are two basic approaches to the partition making - the equidistant and the equiquantile partitioning of marginals, both followed by the 2-D product partition construction. The number of cells choice is the common problem of these methods. Generally, it is known that the optimal number of cells depends not only on the observation number but also on its 2-D distribution. This dependence is much stronger for estimators using equidistant cells.

In this paper, we denote *DirectD* $K \times K$ an algorithm based on the equidistantly generated partition containing K^2 rectangular cells. The cells are the Cartesian product of K equidistant intervals between maximal and minimal values on both marginals. The *DirectQ* $K \times K$ will denote an algorithm based on the partition also containing K^2 rectangular cells, but the cells are the Cartesian products of K intervals between maximal and minimal values on both marginals, which were chosen to contain approximately N/K observations. The estimates are calculated in the same way as in the case of the adaptive histogram methods, i.e. using the equation (2), where $m = K^2$.

2.3 Entropy method

The estimation of entropy is a well developed problem. There are many available methods in the literature. Consider independent identically distributed random vectors $\mathbf{X}_1, \dots, \mathbf{X}_N \in \mathbf{R}^p$ and a partition B_1, \dots, B_m of the space \mathbf{R}^p . Let us define a natural estimation of the entropy of \mathbf{X} 's distribution as

$$\hat{H}_N(\mathbf{p}_N) = - \sum_{i=1}^m p_{N,i} \log p_{N,i}, \quad (3)$$

where $\mathbf{p}_N = (p_{N,i})_{i=1}^m$ are relative frequencies of the sets B_i . This estimate is also called maximum likelihood, plug-in (see Antos and Kontoyiannis [1]) or naive (see Strong [7]). The asymptotical properties of $\hat{H}_N(\mathbf{p}_N)$ are summarized e.g. in Paninski [6]. Let the partition B_1, \dots, B_m , be fixed, $H(p)$ be the entropy of discrete distribution $\mathbf{p} = (p_i)_{i=1}^m$ such that for any $1 \leq i \leq m$ holds

$p_{N,i} \rightarrow p_i$ if $N \rightarrow \infty$, and $p_i > 0$. The known results about the asymptotic bias and variance are

$$\mathbb{E} \left(\widehat{H}_N(\mathbf{p}_N) - H(p) \right) = -\frac{m-1}{2N} + O(N^{-1}), \quad (4)$$

$$\text{Var} \widehat{H}_N(p_N) \leq \frac{(\log N)^2}{N}. \quad (5)$$

For the complete proof of (5) see Antos and Kontoyiannis [1]. The equation (4) is proved also by Miller [5].

At first we fix equidistant partition of marginals and the corresponding product 2-D partition of the observation space. Using the maximum likelihood entropy estimation with the Miller-Madow's bias correction

$$\widehat{H}_N(\mathbf{p}_N) = -\sum_{i=1}^m p_{N,i} \log p_{N,i} + \frac{\hat{m}-1}{2N}, \quad (6)$$

where \hat{m} is number of nonempty cells in used partition, we calculate the entropy estimates $\widehat{H}_N(X)$, $\widehat{H}_N(Y)$ and $\widehat{H}_N(X, Y)$. The *Miller's* estimate of the mutual information can be defined as $\widehat{H}_N(X) + \widehat{H}_N(Y) - \widehat{H}_N(X, Y)$. The equidistant (*DirectD* 5 × 5) construction of partition is used for entropy estimation in the next section.

3 Simulation results

The computer simulations show the estimation results for 2-D data with various true mutual information using methods described in the previous section. The two types of data were generated. At first, the observations were linear dependent data pairs $(X, Y)^T$ with

$$Y = bX + Z,$$

where $X \sim U(0, 1)$, b is linear dependence parameter and Z is random noise variable independent on X . The subplots in the right column of the figure 1 show comparable values of the standard deviations of all used methods without any strong dependence on the noise variable distribution. The left column subplots compare the mean values of the estimators. In the case of the Gaussian noise (a) the adaptive histogram methods are negligibly biased, the method *Mutin B* gives slightly higher values. The other methods have an observable negative bias. The dependence on the noise distribution is apparent in comparison of the Gaussian case (a), the uniform case (b) and the Cauchy case (c). The methods *Mutin A* and *Mutin B* have an observable negative bias in the uniform case, but a moderate positive bias in the Cauchy case. Both the other methods are much more biased. The Miller's estimator is inappropriate in the case of the Cauchy noise because of the partition problems during estimation of the needed entropies.

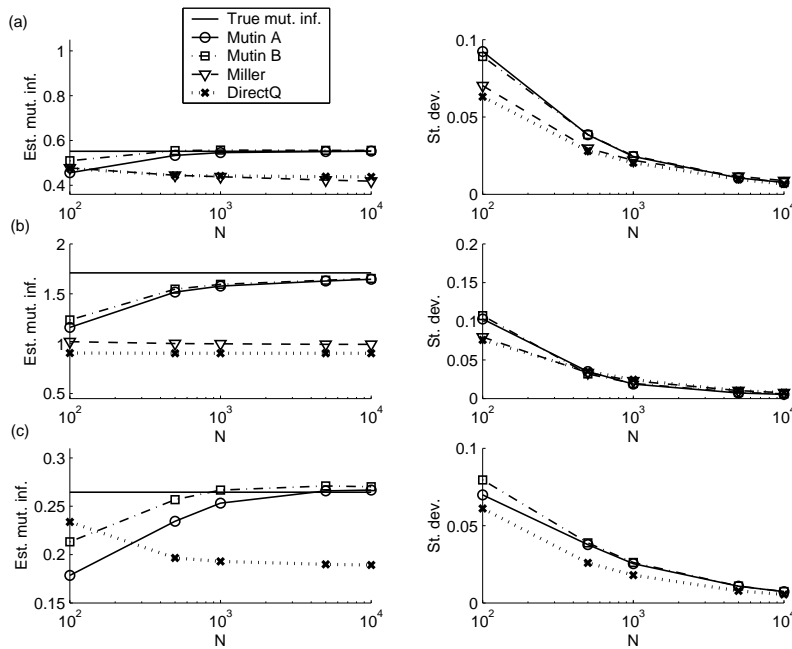


Figure 1: Results of 1000 simulated experiments with N random vectors $(X, Y)^T$ with uniformly distributed variable $X \sim U(0, 1)$ and $Y = 5X + Z$, where Z is random variable independent to X .

Figure (a): $Z \sim N(0, 1)$, $I(X, Y) = 0.5518$;

Figure (b): $Z \sim U(0, 1)$, $I(X, Y) = 1.7094$;

Figure (c): $Z \sim C(0, 1)$, $I(X, Y) = 0.2645$. The Miller's estimator is not displayed because of its extreme bias in the Cauchy distribution case.

The figure 2 shows results obtained in the case of data that have zero linear correlation. The data are uniformly distributed on the annulus with the center in the origin and the various width. The observed standard deviations of all used methods are similar as in the previous case. The methods *Mutin A* and *Mutin B* have a lower bias than both other ones again, but it is seen that absolute bias decreases if the true mutual information increases. It means that the ratio $MSE/I(X, Y)$ of both *Mutin* estimates is notably lower in the case of higher true mutual information.

4 Conclusions

The computer simulation shows that the estimation methods based on the adaptive partition of observation space are more efficient than conservative methods based on a fixed partition. Especially the bias part of the MSE is

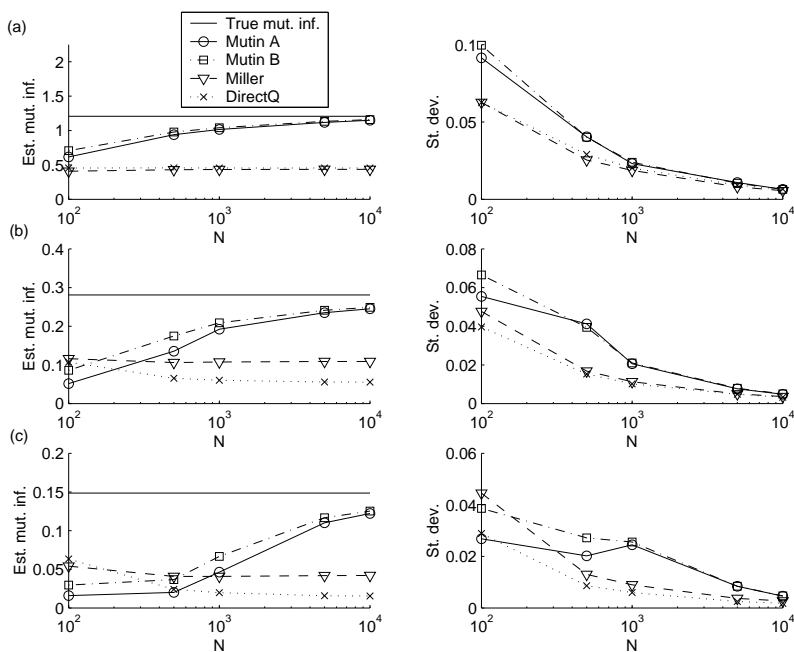


Figure 2: Results of 1000 simulated experiments with N random vectors $(X, Y)^T$ which are uniformly distributed on the annulus with center in the origin. The inner radius of this annulus is r_L and the outer radius is r_U .

Figure (a): $r_U = 1.1$ and $r_L = 0.9$, $I(X, Y) = 1.2081$;

Figure (b): $r_U = 1.5$ and $r_L = 0.5$, $I(X, Y) = 0.2809$;

Figure (c): $r_U = 1.9$ and $r_L = 0.1$, $I(X, Y) = 0.1485$.

seriously reduced. This reduction is stronger if the true mutual information is high and the methods work reasonably well also for data with heavy tailed distributions. The adaptive partitioning also removes the problem with the choice of the fixed partition. It is well known that the number of histogram cells has strong influence on the bias and the standard deviation of corresponding entropy estimators. Generally, the bias decreases and the standard deviation increases with increasing number of cells, see e.g. Paninski [6]. There are many methods optimizing the cells number with regard to the MSE in the literature. A different way to decrease MSE is to employ more accurate methods of density estimation as kernel estimates. We did not deal with this topic for a lack of space. Finally, it is seen that the methods employing the adaptive partitioning of the observation space are very user friendly regarding the implementation and the computation time.

References

- [1] Antos A., Kontoyiannis I. (2001). *Convergence properties of functional estimates for discrete distributions*. Random Structures and Algorithms, **19**, 163–193.
- [2] Darbellay G.A. (1999). *An estimator for the mutual information based on the criterion for independence*. Journal of the Computational Statistics and Data Analysis, **32**, 1–17.
- [3] Darbellay G.A., Vajda I. (1999). *Estimation of the information by an adaptive partitioning of the observation space*. IEEE Transactions on Information Theory, **45**, 1315–1321.
- [4] Franěk J. (2002). *A non parametric estimator of mutual information, redundancy and entropy for continuous random vectors (theory, implementation, applications)*. Ph.D. Thesis, Faculty of Nuclear Science and Physical Engineering Czech Technical University, Prague.
- [5] Miller G. (1955). *Note on the bias of information estimates*. In H. Quastler (Ed.), Information Theory in Psychology II-B, Glencoe, IL: Free Press, 95–100.
- [6] Paninski L. (2003). *Estimation of entropy and mutual information*. Neural Computation, **15**, 1191–1253.
- [7] Strong S., Koberle R., de Ruyter van Steveninck R., Bialek W. (1998). *Entropy and information in neural spike trains*. Physical Review Letters, **80**, 197–202.

Acknowledgement: Tato práce byla podporována projektem 1M0572 MŠMT ČR.

Address: T. Marek, P. Tichavský, ÚTIA, AV ČR, Pod Vodárenskou věží 4, 18208 Praha 8

E-mail: marek@utia.cas.cz

MULTIPLE CHANGES IN COEFFICIENTS OF AUTOREGRESSIVE MODELS

Miriam Marušiaková

Keywords: multiple change points, autoregressive model, F type test, bootstrap.

Abstract: We consider autoregressive models with multiple changes in coefficients occurring at unknown times. We deal with an F type test for detection of changes. We propose application of residual bootstrap to approximate critical values of the test.

Abstrakt: Uvažujeme autoregresné modely s mnohonásobnými zmenami v koeficientoch v neznámych časoch. Zaoberáme sa testom pre detekciu zmien založenom na F štatistikách. Navrhujeme použitie metódy bootstrap pre aproximáciu kritických hodnôt testu.

1 Introduction

We consider an autoregressive model of order p with possible multiple changes in coefficients. The model can be written as a linear regression model

$$y_i = \mathbf{x}'_i \boldsymbol{\beta}_j + e_i \quad i = t_{j-1} + 1, \dots, t_j, \quad j = 1, \dots, m + 1, \quad (1)$$

where regressors $\mathbf{x}'_i = (1, y_{i-1}, y_{i-2}, \dots, y_{i-p})$, $i = p + 1, \dots, n$, are lagged observations, t_1, \dots, t_m are unknown change points (with convention $t_0 = p$, $t_{m+1} = n$) and $\boldsymbol{\beta}_j = (\beta_{j0}, \beta_{j1}, \dots, \beta_{jp})'$ is a $(p + 1)$ -dimensional vector of regression coefficients in the j -th segment, $j = 1, \dots, m + 1$. The parameter β_{j0} is a constant in the model and it equals $\beta_{j0} = (1 - \sum_{l=1}^p \beta_{jl})\mu_j$, where $\mu_j = E y_i$ for $t_{j-1} < i \leq t_j$. The number of changes m is considered to be known. The errors e_1, \dots, e_n are independent and identically distributed (i.i.d.) random variables.

The problem is to test whether there are any changes present and to estimate time moments t_1, \dots, t_m when the coefficients change. We apply tests of no change ($m = 0$) versus the alternative hypothesis of $m = k$ changes. The test was proposed in [4] to detect multiple changes in linear regression models with non-trending regressors. We show that the test can be applied also to autoregressive models. The approximations to the critical values can be obtained either from the limit distribution of the test statistic under the null hypothesis, or by application of residual bootstrap. We propose the latter method here. Bootstrap methods to approximate critical values for a location model with at most one change were proposed in [2]. Since then there have been many different works on bootstrapping in the change point analysis, see a survey [10]. Recent work is [8], where bootstrapping was used

to approximate critical values of a CUSUM test in an autoregressive model with a single change.

Another recent work related to testing for changes in parameters of autoregressive models is [7]. The tests are based on the efficient score vector and can detect a single change in the mean, the error variance or the autoregressive parameters. The literature on testing for multiple changes in autoregressive models is rather sparse. Authors in [3] proposed likelihood-ratio-type tests for multiple structural changes in linear regression models, also allowing for autoregressive models. They also considered a sequential test procedure to determine the unknown number of changes. Another work on testing for multiple changes is [1], which deals with tests based on the Schwarz information criterion and uses the method of binary segmentation.

The remainder of the paper is organized as follows. In section 2 we introduce notation and formulate assumptions. In section 3 we describe the F type test for detection of changes. In section 4 critical values of the test are approximated using the bootstrap methods. The last section is devoted to simulations. The proofs of the main results will be published in [13].

2 Assumptions

For known number of changes $m = k$ we can estimate change points using the least squares approach. Let us call a vector of k change points $\mathbf{t}' = (t_1, \dots, t_k)$ a partition, which divides the whole sample into $k + 1$ subsamples. We define a set of all partitions such that the distance between the change points is at least $n\varepsilon$:

$$T_\varepsilon = \{(t_1, \dots, t_k) : t_{j+1} - t_j \geq n\varepsilon, \forall j = 0, \dots, k\}.$$

The value of the trimming parameter ε is often chosen from the interval $[0.05, 0.15]$. For a given partition $(t_1, \dots, t_k) \in T_\varepsilon$ we can calculate the optimal sum of squared residuals (SSR)

$$S_n(t_1, \dots, t_k) \equiv \sum_{j=1}^{k+1} \min_{\beta_j} \sum_{i=t_{j-1}+1}^{t_j} (y_i - \mathbf{x}'_i \beta_j)^2.$$

The estimated change points $(\hat{t}_1, \dots, \hat{t}_k)$ are such that

$$(\hat{t}_1, \dots, \hat{t}_k) = \arg \min_{t_1, \dots, t_k \in T_\varepsilon} S_n(t_1, \dots, t_k). \quad (2)$$

The algorithm for calculation of the best partition is described in [5] and is implemented in the R-package *strucchange*, see [15].

The residuals corresponding to the known partition $\mathbf{t}' = (t_1, \dots, t_k)$ and to the estimated partition $\hat{\mathbf{t}}' = (\hat{t}_1, \dots, \hat{t}_k)$ are

$$\tilde{e}_i(\mathbf{t}) = y_i - \mathbf{x}'_i \mathbf{C}_{t_{j-1}, t_j}^{-1} \sum_{i=t_{j-1}+1}^{t_j} \mathbf{x}_i y_i, \quad i = t_{j-1}+1, \dots, t_j, \quad j = 1, \dots, k+1$$

and $\tilde{e}_i = \tilde{e}_i(\hat{\mathbf{t}})$, respectively, where

$$\mathbf{C}_{k,l} = \sum_{i=k+1}^l \mathbf{x}_i \mathbf{x}_i', \quad \mathbf{C}_{p,l} = \mathbf{C}_l, \quad k = p, \dots, n-1, \quad l = p+1, \dots, n.$$

The residuals estimated from the model under the null hypothesis of no change are calculated

$$\hat{e}_i = y_i - \mathbf{x}_i' \mathbf{C}_n^{-1} \sum_{i=p+1}^n \mathbf{x}_i y_i, \quad i = p+1, \dots, n. \quad (3)$$

The true values of the parameters are denoted with a 0 superscript. Particularly, $\boldsymbol{\beta}_j^0 = (\beta_{j0}^0, \beta_{j1}^0, \dots, \beta_{jp}^0)'$, $j = 1, \dots, m+1$ and t_1^0, \dots, t_m^0 are the true values of the regression parameters and the change points, respectively. We use the following assumptions.

1. The observations y_{p+1}, \dots, y_n follow the model (1). The initial values y_1, \dots, y_p are independent of e_{p+1}, \dots, e_n . The roots of the polynomial $t^p - \beta_{j1}^0 t^{p-1} - \dots - \beta_{jp}^0$ corresponding to the j -th segment are less than 1 in absolute value, for each $j = 1, \dots, m+1$.
2. The initial values y_1, \dots, y_p satisfy assumption (2.1) in [9].
3. The errors e_i are i.i.d. with zero mean, nonzero finite variance σ^2 and finite moment $E|e_i|^4 < \infty$.
4. The change points satisfy

$$t_j^0 = \lfloor n\lambda_j^0 \rfloor, \quad 0 = \lambda_0^0 < \lambda_1^0 < \dots < \lambda_{m+1}^0 = 1.$$

The last assumption is needed for the asymptotic purposes. The change points have to be asymptotically distinct.

3 F type test

For a given partition $\mathbf{t}' = (t_1, \dots, t_k)$ the test for no change versus k changes is defined as

$$F_n(t_1, \dots, t_k; q) = \frac{1}{kq} \frac{SSR_0 - SSR_k}{\tilde{\sigma}_{n,k}^2(\mathbf{t})},$$

where $SSR_0 = \sum_{i=p+1}^n \hat{e}_i^2$ and $SSR_k = S_n(t_1, \dots, t_k)$ are the minimal sums of squared residuals under the null hypothesis and the alternative one, respectively and

$$\tilde{\sigma}_{n,k}^2(\mathbf{t}) = S_n(t_1, \dots, t_k)/(n-q).$$

The quantity q is the number of regressors in each segment. If we have an intercept in the model, then $q = p+1$, otherwise $q = p$, where p is the order of the autoregressive model.

For unknown change points we take

$$F_n^\varepsilon(k; q) = \max_{t_1, \dots, t_k \in T_\varepsilon} F_n(t_1, \dots, t_k; q). \quad (4)$$

The test (4) is equivalent to $F_n(\hat{t}_1, \dots, \hat{t}_k; q)$, where the estimated partition $\hat{\mathbf{t}}' = (\hat{t}_1, \dots, \hat{t}_k)$ is the best partition defined in (2). The estimator $\tilde{\sigma}_{n,k}^2 = \tilde{\sigma}_{n,k}^2(\hat{\mathbf{t}})$ in the test is a consistent estimator of σ^2 under both the null hypothesis and the alternative one, see Proposition 2.1 in [9] for a single change (it can be proved also for multiple changes). This equivalent form of the test is simpler for computations.

In practice the number of changes m is often unknown and we are more interested in testing the hypothesis of no change versus an unknown number of changes. If we know the upper bound M for the number of changes present, we can use the so called double maximum test defined as $DF_n^\varepsilon(M; q) = \max_{k=1, \dots, M} F_n^\varepsilon(k; q)$, see [4]. Here we will deal only with the case when the number of changes is known. The results of this paper can be easily applied also to the above test.

High values of the test statistic (4) indicate that the null hypothesis is violated. The critical values can be approximated either from the limit distribution of the test under the null hypothesis or by bootstrap methods.

To derive the limit distribution we need to rewrite the test statistic in a more comfortable form. The numerator of the test statistic can be expressed as a functional of partial weighted sums of residuals

$$SSR_0 - SSR_k = \sum_{j=1}^{k+1} \left(\sum_{i=t_{j-1}+1}^{t_j} \mathbf{x}_i \hat{e}_i \right)' \mathbf{C}_{t_{j-1}, t_j}^{-1} \left(\sum_{i=t_{j-1}+1}^{t_j} \mathbf{x}_i \hat{e}_i \right),$$

where residuals \hat{e}_i are defined in (3). Note that under the null hypothesis we have

$$SSR_0 - SSR_k \stackrel{H_0}{=} \left\{ - \left(\sum_{i=p+1}^n \mathbf{x}_i e_i \right)' \mathbf{C}_n^{-1} \left(\sum_{i=p+1}^n \mathbf{x}_i e_i \right) + \sum_{j=1}^{k+1} \left(\sum_{i=t_{j-1}+1}^{t_j} \mathbf{x}_i e_i \right)' \mathbf{C}_{t_{j-1}, t_j}^{-1} \left(\sum_{i=t_{j-1}+1}^{t_j} \mathbf{x}_i e_i \right) \right\}.$$

The following theorem holds.

Theorem 1. Let $\mathbf{W}(\cdot)$ be a vector of q independent standard Wiener processes on $[0, 1]$. Under assumptions A1 - A4 and $m = 0$ (model with no change),

$$F_n^\varepsilon(k; q) \xrightarrow{d} \sup_{\lambda_1, \dots, \lambda_k \in \Lambda_\varepsilon} F(\lambda_1, \dots, \lambda_k; q) \quad (5)$$

with

$$F(\lambda_1, \dots, \lambda_k; q) \stackrel{def}{=} \frac{1}{kq} \sum_{j=1}^k \frac{\|\lambda_j \mathbf{W}(\lambda_{j+1}) - \lambda_{j+1} \mathbf{W}(\lambda_j)\|^2}{\lambda_j \lambda_{j+1} (\lambda_{j+1} - \lambda_j)}$$

and the supremum is taken over the set

$$\Lambda_\varepsilon = \{(\lambda_1, \dots, \lambda_k) : \lambda_{j+1} - \lambda_j \geq \varepsilon, \forall j = 0, \dots, k\}.$$

Proof. We proceed as in Theorem 2.2 in [9] and modify the results to the F type test and multiple changes. We study convergence of the process $\mathbf{V}_n(t) = \sum_{i=p+1}^{\lfloor nt \rfloor} \mathbf{x}_i e_i / (\tilde{\sigma}_{n,k} \sqrt{n})$ together with the properties of the matrices $\mathbf{C}_{k,l}$. The test statistic can be written as a functional of the process $\mathbf{V}_n(\cdot)$. \square

Note that the limit distribution is the same as the limit distribution of the test when applied to linear regression models with non-trending regressors, see Proposition 6 in [4]. The asymptotic critical values are tabulated in [4, 6] for several different values of parameters ε, q, k .

The following theorem is about consistency of the tests.

Theorem 2. Given the assumptions A1 - A4 and the true number of changes $m > 0$ (model with at least one change), the test statistic $F_n^\varepsilon(k; q)$ diverges to infinity for all $k \geq 1$.

Proof. The consistency of the test $F_n^\varepsilon(1; q)$ against $m \geq 1$ changes is proved in Section 3.1 of [3]. This consistency implies the consistency of $F_n^\varepsilon(k; q)$ for $k \geq 2$ due to the inequality $F_n^\varepsilon(1; q) \leq 2F_n^\varepsilon(2; q) \leq 3F_n^\varepsilon(3; q) \leq \dots$. \square

4 Bootstrap methods

The bootstrap methods are explained, e.g., in [2, 11]. To save the space we only describe the testing procedure. Let $\hat{e}_1^*, \dots, \hat{e}_n^*$ be a bootstrap sample with replacement from the residuals $\hat{e}_1, \dots, \hat{e}_n$. We can also use the residuals $\tilde{e}_1, \dots, \tilde{e}_n$ estimated from the alternative model with k changes to have a little better power. Both bootstrap with or without replacement work well. If we don't have an intercept in the model, i.e., $q = p$, we apply the bootstrap method on centered residuals. A bootstrap test statistic is defined as the test statistic under the null hypothesis, where the errors e_1, \dots, e_n are replaced by the bootstrap sample $\hat{e}_1^*, \dots, \hat{e}_n^*$:

$$F_n^{\varepsilon*}(k; q) = \max_{t_1, \dots, t_k \in T_\varepsilon} \frac{1}{kq \tilde{\sigma}_{n,k}^2} \left\{ - \left(\sum_{i=p+1}^n \mathbf{x}_i \hat{e}_i^* \right)' \mathbf{C}_n^{-1} \left(\sum_{i=p+1}^n \mathbf{x}_i \hat{e}_i^* \right) + \sum_{j=1}^{k+1} \left(\sum_{i=t_{j-1}+1}^{t_j} \mathbf{x}_i \hat{e}_i^* \right)' \mathbf{C}_{t_{j-1}, t_j}^{-1} \left(\sum_{i=t_{j-1}+1}^{t_j} \mathbf{x}_i \hat{e}_i^* \right) \right\}. \quad (6)$$

The residuals depend on the original observations. We study the conditional limit distribution of $F_n^{\varepsilon*}(k; q)$, given y_1, \dots, y_n . We consider only randomness induced by the bootstrap sample. We show the quantiles obtained from this bootstrap distribution are asymptotically correct critical values to the original test. This is shown in the following theorem.

Theorem 3. Let the data $\mathbf{y} = (y_1, \dots, y_n)'$ follow the model (1) with $m \geq 0$ changes. Then under assumptions A1 - A4, as $n \rightarrow \infty$,

$$P(F_n^{\varepsilon*}(k; q) \leq x | \mathbf{y}) \xrightarrow{P} P \left(\sup_{\lambda_1, \dots, \lambda_k \in \Lambda_\varepsilon} \frac{1}{kq} \sum_{j=1}^k \frac{\|\lambda_j \mathbf{W}(\lambda_{j+1}) - \lambda_{j+1} \mathbf{W}(\lambda_j)\|^2}{\lambda_j \lambda_{j+1} (\lambda_{j+1} - \lambda_j)} \leq x \right).$$

Proof. The result follows from Theorem 3.2 in [8] and an easy application to the F type test and multiple changes. The main idea is that the partial sums of bootstrapped residuals $\sum_{i=p+1}^{\lfloor nt \rfloor} \mathbf{x}_i \hat{e}_i^*$ can be viewed as partial sums of weighted independent random vectors. We investigate the process $\mathbf{V}_n^*(t) = \sum_{i=p+1}^{\lfloor nt \rfloor} \mathbf{x}_i \hat{e}_i^* / (\tilde{\sigma}_{n,k} \sqrt{n})$ and limit properties of the matrices $\mathbf{C}_{k,l}$ under the null hypothesis and alternatives. \square

We see the conditional limit distribution of the bootstrap test statistic (6) coincides with the limit distribution of the original test under the null hypothesis. Note that the bootstrap approximations to critical values are correct not only when the data follow the null hypothesis ($m = 0$) but also when they follow alternatives.

5 Simulation results

For our simulation experiment we consider an autoregressive model of order $p = 1$, no intercept and with $m \in \{0, 1, 2, 3\}$ changes:

$$y_i = x_i \beta_j + e_i, \quad x_i = y_{i-1} \quad t_{j-1} < i \leq t_j, \quad j = 1, \dots, m+1.$$

The autoregression coefficient β_j has values between -0.2 and 0.7 . The sample size is set to $n = 200$, the number of changes we want to test under the alternative $k = 2$ and the trimming parameter $\varepsilon = 0.15$. In the single change-point model the change occurs after the first half of observations, in the two change-point model after the first and the second third of observations, etc.

To illustrate the size and power of the test (4), we plot size-power-curves plots (SPC-plots), see figure 1. SPC-plots were introduced in [12]. They show the empirical distribution function of p-values of the test $F_n^\varepsilon(k; q)$ for the null hypothesis or some alternative, with respect to the bootstrap distribution $F_n^{\varepsilon*}(k; q)$ (or to the limit distribution $\sup_{\lambda_1, \dots, \lambda_k \in \Lambda_\varepsilon} F(\lambda_1, \dots, \lambda_k; q)$) which was used to determine the critical values of the test. On the y-axis we see the actual size and power of the test for chosen quantiles on the x-axis. So under the null hypothesis the plot should be on a diagonal and under the

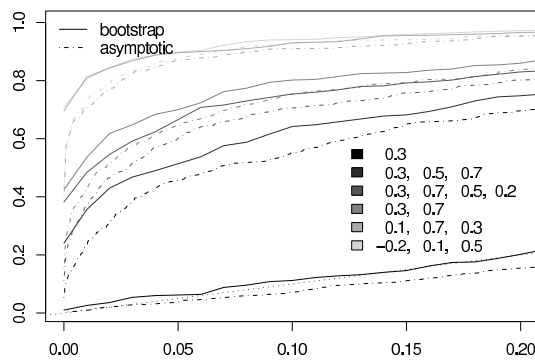


Figure 1: SPC-plots for the test $F_n^\varepsilon(k, q)$ (based on 500 replications) when critical values were approximated by bootstrap methods (solid lines) or obtained from limit distribution (dotdashed lines). Here $n = 200$, $k = 2$, $q = 1$, $\varepsilon = 0.15$. The bootstrap distribution (6) was calculated from 100 random bootstrap samples. The limit distribution (5) was obtained via simulations as described in [4] on page 57.

alternative it should be as steep as possible. We see the power of the test is better if we approximate critical values by bootstrap methods than if we use asymptotic critical values. All the simulations were done in program R, see [14].

6 Conclusion

We studied autoregressive models with multiple changes in coefficients occurring at unknown times. We applied the F type test to detect these changes. The corresponding critical values were approximated applying the bootstrap method. We showed this method provides asymptotically correct critical values. According to the simulation results the power of the test was better if we used approximations to the critical values obtained from the bootstrap method than if we used asymptotic critical values.

References

- [1] Al Ibrahim A.H., Ahmed M.A., BuHamra S.S. (2002). *Testing for multiple change-points in an autoregressive model using SIC criterion*. Journal of Applied Statistical Science, **11**, 57–71.
- [2] Antoch J., Hušková M. (2001). *Permutation tests in change point analysis*. Statistics & Probability Letters, **53**, 37–46.
- [3] Bai J. (1999). *Likelihood ratio tests for multiple structural changes*. Journal of Econometrics, **91**, 299–323.

- [4] Bai J., Perron P. (1998). *Estimating and testing linear models with multiple structural changes*. *Econometrica*, **66**, 47–78.
- [5] Bai J., Perron P. (2003). *Computation and analysis of multiple structural change models*. *Journal of Applied Econometrics*, **18**, 1–22.
- [6] Bai J., Perron P. (2003). *Critical values for multiple structural change tests*. *The Econometrics Journal*, **6**, 72–78.
- [7] Gombay E. (2008). *Change detection in autoregressive time series*. *Journal of Multivariate Analysis*, **99**, 451–464.
- [8] Hušková M., Kirch C., Prášková Z., Steinebach J. (2008) *On the detection of changes in autoregressive time series. II. Resampling procedures*. *Journal of Statistical Planning and Inference*, **138**, 1697–1721.
- [9] Hušková M., Prášková Z., Steinebach J. (2007). *On the detection of changes in autoregressive time series. I. Asymptotics*. *Journal of Statistical Planning and Inference*, **137**, 1243–1259.
- [10] Hušková M. (2004). *Permutation principle and bootstrap in change point analysis*. *Fields Institute Communications*, **44**, 273–291.
- [11] Hušková M., Antoch J. (2003). *Detection of structural changes in regression*. *Tatra Mountains Mathematical Publications*, **26**, 201–216.
- [12] Kirch C. (2006). *Resampling methods for the change analysis of dependent data*. Ph.D. thesis, Köln University.
- [13] Marušiaková M. (2009). Ph.D. thesis, Charles University in Prague, in preparation.
- [14] R Development Core Team (2007). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
- [15] Zeileis A., Leisch F., Hornik K., Kleiber C. (2002). *Strucchange: An R package for testing for structural change in linear regression models*. *Journal of Statistical Software*, **7**, 1–38.

Acknowledgement: The work was supported by grants GAČR 201/05/H007 and 201/09/0775. The author would like to thank Prof. Marie Hušková for fruitful discussions and Department of Mathematics, Uppsala University for a nice working place and friendly environment.

Address: M. Marušiaková, Faculty of Mathematics and Physics, Department of Probability and Mathematical Statistics, Sokolovská 83, 186 75 Prague 8 – Karlín

E-mail: maruskay@gmail.com

CHANGE POINT DETECTION BY BASIS PURSUIT

Jiří Neubauer, Vítězslav Veselý

Keywords: Change point detection, overcomplete model, sparse parameter estimation.

Abstract: The contribution deals with use of overcomplete models and sparse parameter estimation for change point detection in one-dimensional stochastic processes. These processes are estimated by 'Heaviside' functions. The BASIS PURSUIT algorithm is used to get sparse parameter estimation. The mentioned method of change point detection in stochastic processes is compared with standard methods by simulations.

Abstrakt: Příspěvek se zabývá možným využitím přeparametrizovaných modelů a hledání řídkých řešení pro detekci změn v jednorozměrných stochastických procesech. Tyto procesy jsou odhadovány pomocí „Heaviside“ funkcí. Pro hledání řídkých řešení je použit algoritmus BASIS PURSUIT. Uvedená metoda detekce změn v stochastických procesech je porovnávána se standardními metodami odhadu těchto změn pomocí simulací.

1 Introduction

Chen et al [3] proposed a new methodology based on Basis Pursuit for spectral representation of signals (vectors). Instead of just representing signals as superpositions of sinusoids (the traditional Fourier representation) they suggested alternate *dictionaries* — collections of parametrized waveforms — of which the wavelets dictionary is only the best known. Theoretical background for such systems (also called *frames*) can be found for example in [4, 8]. In traditional Fourier expansion a presence of jumps in the signal slows down the convergence rate preventing sparsity. The *heaviside* dictionary [3, 2.1.1] merged with the Fourier or wavelet dictionary can solve the problem quite satisfactorily.

A lot of other useful applications in a variety of problems can be found in [9, 7, 10].

In [10] kernel dictionaries showed to be an effective alternative to traditional kernel smoothing techniques. In this paper we are using heaviside dictionary in the same manner to denoise signal (univariate time series sample path) exhibiting jumps (discontinuities) in the mean (unlike kernel smoothing where the mean is supposed to be sufficiently smooth).

Consequently, the Basis Pursuit approach can be proposed as an alternative to conventional statistical techniques of change point detection [1, 2]. This paper presents results of an introductory empirical comparative study for the simplest case of detecting one change point buried in additive gaussian white noise.

2 Heaviside frame for change point detection

In this section we propose the method based on Basis Pursuit Algorithm (BPA) for the detection of the change point in the sample path $\{y_t\}$ in one dimensional stochastic process $\{Y_t\}$.

We assume a deterministic functional model on a bounded interval \mathcal{I} described by the dictionary $G = \{G_j\}_{j \in J}$ with atoms $G_j \in L^2(\mathcal{I})$ and with additive white noise e on a suitable finite discrete mesh $\mathcal{T} \subset \mathcal{I}$:

$$Y_t = x_t + e_t, \quad t \in \mathcal{T},$$

where $x \in \text{sp}(\{G_j\}_{j \in J})$, $\{e_t\}_{t \in \mathcal{T}} \sim WN(0, \sigma^2)$, $\sigma > 0$, and J is a big finite indexing set. Smoothed function $\hat{x} = \sum_{j \in J} \hat{\xi}_j G_j =: \mathbf{T}\hat{\xi}$ minimizes on \mathcal{T} ℓ^1 -penalized optimality measure $\frac{1}{2}\|\mathbf{y} - \mathbf{T}\xi\|^2$ as follows:

$$\hat{\xi} = \underset{\xi \in \ell^2(J)}{\text{argmin}} \frac{1}{2}\|\bar{y} - \mathbf{T}\xi\|^2 + \lambda\|\xi\|_1, \quad \|\xi\|_1 := \sum_{j \in J} \|G_j\|_2 \xi_j,$$

where $\lambda = \sigma\sqrt{2 \ln(\text{card } J)}$ is a smoothing parameter chosen according to the soft-thresholding rule commonly used in wavelet theory. This choice is natural because one can prove that with any orthonormal basis $G = \{G_j\}_{j \in J}$ the shrinkage via soft-thresholding produces the same smoothing result \hat{x} .

Solution of this minimization problem with λ close to zero may not be sparse enough: we are searching small $F \subset J$ such that $\hat{x} \approx \sum_{j \in F} \hat{\xi}_j G_j$ is a good approximation. That is why we apply the following four-step procedure described in [10] in more detail and implemented in [6].

- (A0) Choice of a raw initial estimate $\xi^{(0)}$ [10, 3.2.2(1)].
- (A1) We improve $\xi^{(0)}$ iteratively [10, 3.2.2(2)] by stopping at $\xi^{(1)}$ which satisfies optimality criterion. The solution $\xi^{(1)}$ is optimal but not sparse in general.
- (A2) Starting with $\xi^{(1)}$ we are looking for $\xi^{(2)}$ by BPA which tends to be nearly sparse and is optimal [10, 3.2.3].
- (A3) We construct a sparse and optimal solution ξ^* by removing negligible parameters and corresponding atoms from the model [10, 3.2.4], namely those satisfying $|\xi_j^{(2)}| < \alpha\|\xi^{(2)}\|_1$ where $0 < \alpha \ll 1$ is a suitable sparsity level, a typical choice being $\alpha = 0.05$ following an analogy with the statistical significance level.
- (A4) We repeat the step (A1) with the dictionary reduced according to the step (A3) and with a new initial estimate $\xi^{(0)} = \xi^*$ [10, 3.2.5]. We expect to obtain a possibly improved sparse estimate ξ^* .

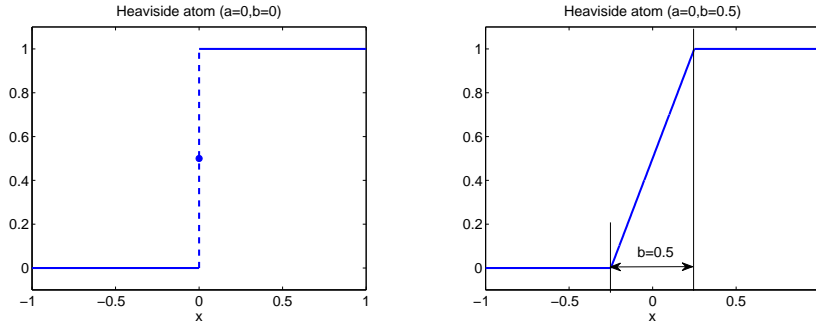


Figure 1: Heaviside atoms with parameters $a = 0, b = 0$ and $a = 0, b = 0.5$.

Hereafter we refer to this four-step algorithm as to BPA4. The steps (A1), (A2) and (A4) use Primal-Dual Barrier Method designed by M. Saunders [5]. This up-to-date sophisticated algorithm allows one to solve fairly general optimization problems minimizing convex objective subject to linear constraints. A lot of controls provide a flexible tool for adjusting the iteration process.

We build our dictionary from heaviside-shaped atoms on $L^2(\mathbb{R})$ derived from a fixed 'mother function' via shifting and scaling following the analogy with the construction of wavelet bases.

We construct an oversized shift-scale dictionary $G = \{G_{a,b}\}_{a \in \mathcal{A}, b \in \mathcal{B}}$ derived from the 'mother function' by varying the shift parameter a and the scale (width) parameter b between values from big finite sets $\mathcal{A} \subset \mathbb{R}$ and $\mathcal{B} \subset \mathbb{R}^+$, respectively ($J = \mathcal{A} \times \mathcal{B}$), on a bounded interval $\mathcal{I} \subset \mathbb{R}$ spanning the space $H = \text{sp}(\{1_{\mathcal{I}} G_{a,b}\}_{a \in \mathcal{A}, b \in \mathcal{B}})$, where

$$G_{a,b}(t) = \begin{cases} 1 & \text{for } t - a > b/2, \\ 1/2 + (t - a)/b & |t - a| \leq b/2, b \neq 0, \\ 1/2 & t = a, b = 0, \\ 0 & \text{otherwise.} \end{cases}$$

In the simulations below $\mathcal{I} = [0, 1]$, $\mathcal{T} = \{t/T\}_{t=1}^T$ (typically with mesh size $T = 100$), $\mathcal{A} = \{-1/T\} \cup \{t/T\}_{t=t_0}^{T-t_0}$ (t_0 is a boundary trimming) and scale b fixed to zero ($\mathcal{B} = \{0\}$). Some examples of Heaviside functions are displayed in the figure 1.

3 Change point detection

Consider the following model

$$Y_t = \begin{cases} \mu + \epsilon_t & t = 1, 2, \dots, c \\ \mu + \delta + \epsilon_t & t = c + 1, \dots, T, \end{cases} \quad (1)$$

where $\mu, \delta \neq 0, c < T$ are unknown parameters and ϵ_t are independent identically distributed random variables with zero mean and variance σ^2 and $E|\epsilon_t|^{2+\Delta} < \infty$ for some $\Delta > 0$. We assume that the change point satisfies $c = [T\gamma]$ for some $\gamma \in (0, 1)$, where $[x]$ denotes the integer part of x .

The unknown parameters c, μ and δ is possible to estimate by the least squared method. The least squares estimators $\hat{c}, \hat{\mu}$ and $\hat{\delta}$ of the parameters c, μ and δ are defined as solutions of the minimization problem

$$\min \left\{ \sum_{t=1}^k (Y_t - \mu)^2 + \sum_{t=k+1}^T (Y_t - \mu - \delta)^2; k \in \{1, \dots, T-1\}, \mu \in \mathbb{R}, \delta \in \mathbb{R} \right\}.$$

In other word, the unknown parameters are estimated in such a way that the sum of squares of residuals is minimal. The estimates of the parameters μ and δ are (see [1, 2])

$$\hat{\mu} = \bar{Y}_{\hat{c}} \quad \text{and} \quad \hat{\delta} = \bar{Y}_{\hat{c}}^0 - \bar{Y}_{\hat{c}},$$

where \hat{c} is a solution of the maximization problem

$$\hat{c} = \arg \max \left\{ \sqrt{\frac{T}{k(T-k)}} \cdot |S_k|; k \in \{1, \dots, T-1\} \right\}, \quad (2)$$

where $S_k = \sum_{t=1}^k (Y_t - \bar{Y}_T)$, $\bar{Y}_T = \frac{1}{T} \sum_{t=1}^T Y_t$, $\bar{Y}_{\hat{c}} = \frac{1}{\hat{c}} \sum_{t=1}^{\hat{c}} Y_t$ and $\bar{Y}_{\hat{c}}^0 = \frac{1}{T-\hat{c}} \sum_{t=\hat{c}+1}^T Y_t$.

If we solve given maximization problem for $[\beta T] \leq k \leq [(1-\beta)T]$, where β is a small positive real constant less than 1 (usually $\beta \in (0.01, 0.1)$), then we get so called trimmed estimation of the change point.

M-estimators of change point We assume the model (1) with the error terms having a common distribution F that is symmetric about zero. The M -estimator $\hat{\mu}_M(\psi)$ of μ generated by score function ψ is defined as a solution of the equation $\sum_{t=1}^T \psi(Y_t - u) = 0$. The score function ψ is assumed to be monotone and skew symmetric. The M -estimator $\hat{c}_M(\psi)$ of c is defined as a solution of (2) where the partial sums S_k are substituted by $S_{k,M}(\psi) = \sum_{t=1}^k \psi(Y_t - \hat{\mu}_T(\psi))$, $k = 1, \dots, T$.

R-estimators of change point We assume the model (1) with the error terms having absolutely continuous density f . The rank-based estimator of c is based on the partial sums $S_{k,R} = \sum_{t=1}^k (a(R_t) - \bar{a}_T)$, $k = 1, \dots, T$, where R_1, \dots, R_T are the ranks corresponding to Y_1, \dots, Y_T , $a(1), \dots, a(T)$ are scores and $\bar{a}_T = \frac{1}{T} \sum_{t=1}^T a(t)$. The R -estimator of c is defined as a solution of (2), where the partial sums S_k are substituted by the partial sums $S_{k,R}$.

4 Simulation study

The methods mentioned above were used to compare several estimates of the change point: BP method, max-type and trimmed max-type method, M and R -estimators. We consider the model (1). For the purpose of simulations we put $\mu = 0, \delta = 1, T = 100$, the error terms are independent normally distributed with zero mean and the standard deviation $\sigma = 0.5, 1, 1.5$ and 2 . We calculate simulation of this model with the change point $c = 20, 50$ and 80 (500 simulations for each combination of the parameters). The results for $\sigma = 1.5$ and 2 are summarised in the tables 1 and 2. (\bar{x} denotes the sample mean, s is the sample standard deviation, x_{\min} and x_{\max} are minimum and maximum of the sample, $x_{0.25}$ and $x_{0.75}$ are the lower and the upper sample quartiles, $x_{0.50}$ is the median, R_Q denotes the interquartile range $x_{0.75} - x_{0.25}$). The parameter β in the trimmed max-type method is 0.05 . BP estimates were calculated by the following procedure: we constructed the dictionary (Heaviside functions with parameters $a = 0, 1, 2, \dots, 99$ and $b = 0$), the significant atoms were detected by BP algorithm firstly, then we got the smoothed process using these significant atoms and finally max-type estimation were applied. To get M -estimators of the change point the Huber ψ -function where $\psi(x) = x$ for $|x| \leq B$ and $\psi(x) = B \text{sign}(x)$ for $|x| > B$ were chosen. The Wilcoxon scores $a(t) = t/(T+1), t = 1, \dots, T$ were used in calculation of R -estimators.

5 Conclusion

If the variance of the error terms ϵ_t is small ($\sigma = 0.5$ or 1), there is no significant difference between these 5 methods of the change point estimation. In case that the variance is larger, BP method seems to be slightly better in term of the sample standard deviations and the interquartile ranges of BP estimates (see tables 1 and 2, the estimates of the parameter c). All methods give us comparable estimates of parameter δ .

The Basis Pursuit approach proposes an alternative technique of the change point detection. This paper deals with the introductory empirical comparative study for the simplest case of detecting one change point. The mentioned method can be used on detection of two or more change points, or another sort of change point which would required another atoms and dictionary G .

$c = 20$	BP		Max-type		Trimmed		M -estim.		R -estim.	
	\hat{c}	$\hat{\delta}$	\hat{c}	$\hat{\delta}$	\hat{c}	$\hat{\delta}$	\hat{c}	$\hat{\delta}$	\hat{c}	$\hat{\delta}$
\bar{x}	26.26	1.14	29.21	1.19	28.18	1.14	29.80	1.09	29.21	1.19
s	15.29	0.36	23.77	0.91	20.69	0.56	22.60	0.79	23.77	0.91
x_{\min}	1.00	0.39	1.00	-4.14	5.00	-2.36	1.00	-3.16	1.00	-4.14
x_{\max}	98.00	3.11	99.00	4.83	95.00	2.79	99.00	4.51	99.00	4.83
$x_{0.25}$	19.00	0.87	17.00	0.98	18.00	0.94	18.00	0.93	17.00	0.98
$x_{0.50}$	21.00	1.12	20.00	1.23	21.00	1.19	21.00	1.19	20.00	1.23
$x_{0.75}$	28.00	1.37	30.00	1.51	28.50	1.46	32.00	1.45	30.00	1.51
R_Q	9.00	0.50	13.00	0.53	10.50	0.52	14.00	0.53	13.00	0.53

$c = 50$	BP		Max-type		Trimmed		M -estim.		R -estim.	
	\hat{c}	$\hat{\delta}$	\hat{c}	$\hat{\delta}$	\hat{c}	$\hat{\delta}$	\hat{c}	$\hat{\delta}$	\hat{c}	$\hat{\delta}$
\bar{x}	51.59	1.15	49.77	1.15	50.26	1.15	50.07	1.13	49.77	1.15
s	11.23	0.31	14.04	0.59	12.76	0.36	13.44	0.50	14.04	0.59
x_{\min}	1.00	-0.19	1.00	-4.45	5.00	-2.05	1.00	-2.90	1.00	-4.45
x_{\max}	95.00	2.94	99.00	4.53	95.00	2.35	98.00	4.21	99.00	4.53
$x_{0.25}$	48.00	0.95	46.00	0.97	46.00	0.96	46.00	0.96	46.00	0.97
$x_{0.50}$	50.00	1.14	50.00	1.17	50.00	1.17	50.00	1.16	50.00	1.17
$x_{0.75}$	55.00	1.35	54.00	1.36	54.00	1.36	54.00	1.36	54.00	1.36
R_Q	7.00	0.40	8.00	0.40	8.00	0.40	8.00	0.40	8.00	0.40

$c = 80$	BP		Max-type		Trimmed		M -estim.		R -estim.	
	\hat{c}	$\hat{\delta}$	\hat{c}	$\hat{\delta}$	\hat{c}	$\hat{\delta}$	\hat{c}	$\hat{\delta}$	\hat{c}	$\hat{\delta}$
\bar{x}	77.80	1.25	72.76	1.20	73.36	1.20	72.57	1.18	72.76	1.20
s	14.21	0.52	22.17	0.87	19.91	0.53	20.35	0.64	22.17	0.87
x_{\min}	2.00	-0.52	1.00	-4.75	5.00	-1.73	1.00	-2.97	1.00	-4.74
x_{\max}	98.00	3.53	99.00	4.81	95.00	2.56	99.00	4.23	99.00	4.81
$x_{0.25}$	75.50	0.93	73.00	0.97	73.00	0.96	71.00	0.94	73.00	0.97
$x_{0.50}$	80.00	1.21	80.00	1.24	80.00	1.21	79.50	1.19	80.00	1.24
$x_{0.75}$	84.00	1.52	83.00	1.53	83.00	1.47	83.00	1.45	83.00	1.53
R_Q	8.50	0.59	10.00	0.56	10.00	0.51	12.00	0.51	10.00	0.56

Table 1: $\sigma = 1.5$, $\delta = 1$, $T = 100$.

$c = 20$	BP		Max-type		Trimmed		M -estim.		R -estim.	
	\hat{c}	$\hat{\delta}$	\hat{c}	$\hat{\delta}$	\hat{c}	$\hat{\delta}$	\hat{c}	$\hat{\delta}$	\hat{c}	$\hat{\delta}$
\bar{x}	28.66	1.32	33.24	1.39	32.24	1.28	33.73	1.31	33.24	1.39
s	19.27	0.63	27.41	1.37	23.95	0.76	26.82	1.01	27.41	1.37
x_{\min}	1.00	-3.65	1.00	-4.49	5.00	-2.10	1.00	-4.42	1.00	-4.49
x_{\max}	98.00	4.89	99.00	7.98	95.00	3.48	99.00	6.37	99.00	7.98
$x_{0.25}$	18.50	0.97	17.00	1.07	18.00	1.00	17.00	1.03	17.00	1.07
$x_{0.50}$	22.50	1.29	22.00	1.44	22.00	1.39	22.00	1.42	22.00	1.44
$x_{0.75}$	32.00	1.59	41.00	1.79	37.50	1.69	41.50	1.75	41.00	1.79
R_Q	13.50	0.62	24.00	0.72	19.50	0.69	24.50	0.72	24.00	0.72

$c = 50$	BP		Max-type		Trimmed		M -estim.		R -estim.	
	\hat{c}	$\hat{\delta}$	\hat{c}	$\hat{\delta}$	\hat{c}	$\hat{\delta}$	\hat{c}	$\hat{\delta}$	\hat{c}	$\hat{\delta}$
\bar{x}	53.48	1.29	51.09	1.35	49.65	1.29	49.12	1.28	51.09	1.35
s	16.78	0.63	21.03	1.07	18.58	0.56	20.36	0.84	21.03	1.07
x_{\min}	1.00	-4.11	1.00	-4.79	5.00	-1.77	1.00	-4.11	1.00	-4.79
x_{\max}	98.00	4.74	99.00	7.67	95.00	3.10	99.00	4.80	99.00	7.67
$x_{0.25}$	47.00	1.04	42.00	1.10	42.00	1.08	41.00	1.09	42.00	1.10
$x_{0.50}$	51.00	1.28	50.00	1.35	50.00	1.32	50.00	1.34	50.00	1.35
$x_{0.75}$	60.00	1.54	59.00	1.62	58.00	1.57	57.00	1.59	59.00	1.62
R_Q	13.00	0.50	17.00	0.52	16.00	0.49	16.00	0.50	17.00	0.52

$c = 80$	BP		Max-type		Trimmed		M -estim.		R -estim.	
	\hat{c}	$\hat{\delta}$	\hat{c}	$\hat{\delta}$	\hat{c}	$\hat{\delta}$	\hat{c}	$\hat{\delta}$	\hat{c}	$\hat{\delta}$
\bar{x}	74.23	1.37	68.46	1.42	69.17	1.31	67.96	1.34	68.46	1.42
s	20.93	0.92	26.82	1.35	23.45	0.72	26.56	1.04	26.82	1.35
x_{\min}	1.00	-3.81	1.00	-5.09	5.00	-1.77	1.00	-3.81	1.00	-5.09
x_{\max}	98.00	5.04	99.00	7.37	95.00	3.23	99.00	5.04	99.00	7.37
$x_{0.25}$	72.00	0.94	61.00	1.07	62.00	1.05	60.00	1.05	61.00	1.07
$x_{0.50}$	80.00	1.32	79.00	1.43	79.00	1.36	79.00	1.40	79.00	1.43
$x_{0.75}$	86.00	1.81	84.50	1.84	83.00	1.68	84.00	1.82	84.50	1.84
R_Q	14.00	0.87	23.50	0.77	21.00	0.63	24.00	0.77	23.50	0.77

Table 2: $\sigma = 2$, $\delta = 1$, $T = 100$.

References

- [1] Antoch J., Hušková M., Jarušková D. (2000). *Change point detection*. In *5th ERS IASC Summer School*, IASC 2000.
- [2] Antoch J., Hušková M., Jarušková D. (1998). *Change point problem po deseti letech*. In *Sborník prací desáté zimní školy JČMF ROBUST 98*, JČMF 1998, 1–42, ISBN 80-7015-624-4.
- [3] Chen S.S., Donoho D.L., Saunders M.A. (1998). *Atomic decomposition by basis pursuit*. *SIAM J. Sci. Comput.* **20** (1), 33–61 (2001 reprinted in *SIAM Review* **43** (1), 129–159).
- [4] Christensen O. (2003). *An introduction to frames and Riesz bases*. Birkhäuser: Boston-Basel-Berlin.
- [5] Saunders M.A. (1997–2001). *pdsco.m: MATLAB code for minimizing convex separable objective functions subject to $Ax = b, x \geq 0$* .
- [6] Veselý V. (2001–2008). *framebox: MATLAB toolbox for overcomplete modeling and sparse parameter estimation*.
- [7] Veselý V., Tonner J., Hrdličková Z., Michálek J., Kolář M. (2008–2009). *Analysis of PM_{10} air pollution in Brno based on generalized linear model with strongly rank-deficient design matrix*. *Environmetrics*, to appear.
- [8] Veselý V. (2002). *Hilbert-space techniques for spectral representation in terms of overcomplete bases*. In *Proceedings of the Summer School DATASTAT'2001*, Čihák near Žamberk, Horová I. (ed.); *Folia Fac. Sci. Nat. Univ. Masaryk. Brunensis, Mathematica* **11**: Masaryk University, Brno, Czech Rep., 259–273¹.
- [9] Veselý V., Tonner J. (2005). *Sparse parameter estimation in overcomplete time series models*. *Austrian Journal of Statistics*, **35** (2&3), 371–378¹.
- [10] Zelinka J., Veselý V., Horová I. (2004). *Comparative study of two kernel smoothing techniques*. In *Proceedings of the Summer School DATASTAT'2003*, Svratka, Horová I (ed.); *Folia Fac. Sci. Nat. Univ. Masaryk. Brunensis, Mathematica* **15**: Masaryk University, Brno, Czech Rep., 419–436¹.

Acknowledgement: The research was supported by the grant MO0FEM000401 and by MŠMT ČR (Ministry of Education of the Czech Republic) under research contract MSM0021622418.

Address: J. Neubauer, Department of Econometrics, Faculty of Economics and Management, University of Defence, Kounicova 65, 612 00 Brno
 V. Veselý, Department of Applied Mathematics and Computer Science, Faculty of Economics and Administration, Masaryk University, Lipová 41a, 602 00 Brno

E-mail: Jiri.Neubauer@unob.cz, vesely@econ.muni.cz

¹download available at http://www.math.muni.cz/~vesely/papers_cz.html

AKO POROZUMIEŤ NELINEÁRNEJ METÓDE NAJMENŠÍCH ŠTVORCOV POMOCOU GEOMETRIE

Andrej Pázman

Klíčové slová: Nelineárna regresia, krivosti, oblasti spoľahlivosti, distribúcia odhadov.

Abstrakt: Článok je stručnou odozvou skúseností z viacročnej prednášky o nelineárnych regresných modeloch a z predchádzajúceho autorovho výskumu v tejto oblasti. Z obsahu: reparametrizácia modelu, interpretácia Gaussovej-Newtonovej metódy, vplyvy vnútornej a parametrickej krivosti modelu na vlastnosti odhadov, zjednodušený výklad asymptotickej normality odhadov, geometrické odvodenie hustoty pravdepodobnosti odhadov, oblasti spoľahlivosti pre parametre, otázky jednoznačnosti odhadu. Verím, že predkladané, čiastočne pedagogické poznatky, budú pre čitateľa dostatočne zaujímavé a dobre pochopiteľné.

Abstract: The paper reflects briefly the experience from a many-years lecture on nonlinear regression and from a preceding research on this topic. Contents: reparametrization, interpretation of the Gauss-Newton method, influence of the intrinsic and parameter curvatures on the properties of estimators, a simplified exposition of the asymptotics of estimators, a geometric derivation of the density of estimators, confidence regions, problems with the uniqueness of estimators. I hope that the presented, partly pedagogical exposition will be sufficiently clear and interesting to the reader.

1 Ukážka zložitosti istého modelu z aplikácií

Začneme výklad poukázaním na to, že aj v takej rýdzo aplikačnej oblasti, akou je konzervačný priemysel, sú potrebné výrazne nelineárne parametrické modely. Uvádzame preto príklad Baranyiho a Robertsovho modelu rastu počtu baktérií v potravine [1]. Podľa osobného vyjadrenia pracovníkov biometrického laboratória v INRA vo Francúzsku, ide o veľmi uznávaný model v praxi. Počet baktérií y_i v čase x_i je tu modelovaný vzťahom

$$E(y_i) \approx \ln(\theta_1) + \theta_2 x_i + \ln[A(x_i, \theta)] - \ln[B(x_i, \theta)], \quad i = 1, \dots, N,$$

kde

$$\begin{aligned} A(x, \theta) &= e^{-\theta_2 x} + e^{-\theta_2 \theta_3} - e^{(-\theta_2 x - \theta_2 \theta_3)}, \\ B(x, \theta) &= 1 + \left[e^{\theta_2 A(x, \theta)} - 1 \right] / \left(\frac{\theta_4}{\theta_1} \right). \end{aligned}$$

Tu N je počet meraní, θ_1 je počiatkový počet baktérii, θ_2 je maximálna rýchlosť nárastu počtu baktérii počas celej doby kazivosti, θ_3 je doba zanedbateľnej kazivosti potravy, θ_4 je počet baktérii v úplne pokazenej potravy. Komplikovanosť týchto vzťahov nevyžaduje ďalší komentár, ale je aj pre teoretika prekvapením.

2 Nelineárny regresný model

Tento dobre známy model je daný vzťahmi

$$\begin{aligned} y_i &= \eta(x_i, \theta) + \varepsilon_i; \quad \theta \in \Theta \subset \mathbb{R}^m \\ E(\varepsilon_i) &= 0, \quad Cov(\varepsilon_i, \varepsilon_j) = \sigma^2 \delta_{ij}; \quad i, j = 1, \dots, N \end{aligned}$$

alebo vo vektorovom zápise

$$\begin{aligned} y &= \eta(\theta) + \varepsilon \\ E(\varepsilon) &= 0, \quad Var(\varepsilon) = \sigma^2 I, \end{aligned}$$

kde $y = (y_1, \dots, y_N)^T$ je vektor pozorovaných veličín, $\eta(\theta)$ je vektor ich stredných hodnôt, $\eta(\theta) = (\eta(x_1, \theta), \dots, \eta(x_N, \theta))^T$, pri danej hodnote $\theta \in \Theta$, $\Theta \in \mathbb{R}^m$ je parametrický priestor a $\varepsilon = (\varepsilon_1, \dots, \varepsilon_N)^T$ je vektor chýb. Odhad neznámych parametrov $\theta = (\theta_1, \dots, \theta_m)^T$ (nelineárnou) metódou najmenších štvorcov (MNS) je

$$\hat{\theta} = \arg \min_{\theta \in \Theta} \sum_{i=1}^N [y_i - \eta(x_i, \theta)]^2 = \arg \min_{\theta \in \Theta} \|y - \eta(\theta)\|^2$$

Poznamenávame, že v tomto modeli predpokladáme, že Θ má neprázdne vnútro t.j. $int(\Theta) \neq \emptyset$, že Θ nemá izolované body, že funkcia $\theta \rightarrow \eta(\theta)$ má spojité druhé parciálne derivácie na $int(\Theta)$ a že matica prvých derivácií $\frac{\partial \eta(\theta)}{\partial \theta^T}$ má plnú hodnotu v každom bode $\theta \in int(\Theta)$ (t.j. experiment obsahuje dostatočný počet pozorovaní).

Vo všeobecnosti symbolom $\bar{\theta}$ budeme označovať skutočnú hodnotu parametra θ .

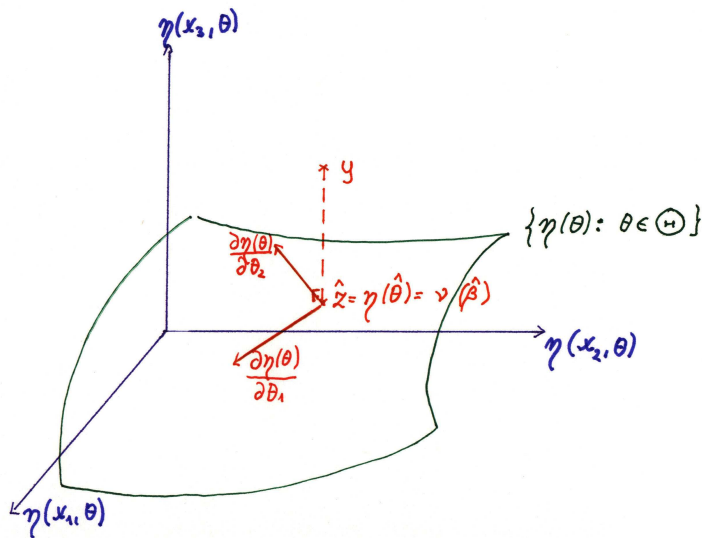
3 Plocha stredných hodnôt a reparametrizácia modelu

Opäť začneme príkladom a to známym Michaelisovým-Mentenovým modelom

$$y_i = \frac{\theta_1 x_i}{\theta_2 + x_i} + \varepsilon_i,$$

kde y_i je napr. produkt katalytickej chemickej reakcie v čase x_i . Môžeme použiť aj inú parametrizáciu modelu

$$\beta_1 = \theta_1, \quad \beta_2 = \frac{\theta_1}{\theta_2},$$



Obrázok 1: Plocha stredných hodnôt, dotykové vektory a projekcia bodu y na túto plochu.

pričom β_1 je maximálna úroveň krivky $x \rightarrow E_x(y)$ a β_2 je rýchlosť narastania krivky blízko počiatku, teda nové parametre sú dobre interpretovateľné. Potom „nový model“ je

$$y_i = \frac{\beta_1 \beta_2 x_i}{\beta_1 + \beta_2 x_i} + \varepsilon_i$$

Jednoduchá otázka (študentská): Ako súvisia MNŠ odhady v oboch modeloch? Formulujme otázku trochu všeobecnejšie. Máme model (I)

$$\begin{aligned} y &= \eta(\theta) + \varepsilon, \\ \hat{\theta} &= \arg \min_{\theta \in \Theta} \|y - \eta(\theta)\|^2, \end{aligned}$$

a model (II)

$$\begin{aligned} y &= \nu(\beta) + \varepsilon, \\ \hat{\beta} &= \arg \min_{\beta \in \psi(\Theta)} \|y - \nu(\beta)\|^2, \end{aligned}$$

pričom ide o reparametrizáciu modelu (I), podľa vzťahov $\theta = \phi(\beta)$, $\beta = \psi(\theta)$, $\nu(\beta) = \eta(\phi(\beta))$.

Tvrdenie: Platí $\hat{\theta} = \phi(\hat{\beta})$, $\hat{\beta} = \psi(\hat{\theta})$.

Tvrdenie zdôvodníme geometricky. Zavedieme *plochu stredných hodnôt*, t.j. množinu možných stredných hodnôt vektora y

$$\mathcal{E} = \{\eta(\theta) : \theta \in \Theta\} = \{\nu(\beta) : \beta \in \psi(\Theta)\}.$$

Určenie MNSŕ odhadu vlastne znamená projektovať bod y na plochu \mathcal{E} , teda nájsť najprv vektor (pozri obr. 1)

$$\hat{z} = \arg \min_{z \in \mathcal{E}} \|y - z\|^2$$

a potom parametrizovať tento bod v rámci modelu (I) resp. modelu (II)

$$\hat{z} = \eta(\hat{\theta}) \text{ resp. } \hat{z} = \nu(\hat{\beta}).$$

Teda $\hat{\theta}$ a $\hat{\beta}$ sú len dve rôzne parametrizácie toho istého bodu plochy \mathcal{E} , odkiaľ vyplýva $\hat{\theta} = \phi(\hat{\beta})$, $\hat{\beta} = \psi(\hat{\theta})$.

4 Ako pochopiť geometricky Gaussov-Newtonov algoritmus výpočtu odhadov

V tomto iteračnom algoritme sa využíva linearizácia nelineárneho modelu pomocou Taylorovho radu v každom iteračnom bode $\theta^{(n)} \in \text{int}(\Theta)$. To znamená že model $y = \eta(\theta) + \varepsilon$ sa aproximuje lineárnym modelom

$$y = \eta(\theta^{(n)}) + \left[\frac{\partial \eta(\theta)}{\partial \theta^T} \right]_{\theta=\theta^{(n)}} (\theta - \theta^{(n)}) + \varepsilon.$$

Plocha stredných hodnôt tohto modelu

$$\tau = \left\{ \eta(\theta^{(n)}) + \left[\frac{\partial \eta(\theta)}{\partial \theta^T} \right]_{\theta=\theta^{(n)}} \alpha : \alpha \in \mathbb{R}^m \right\}$$

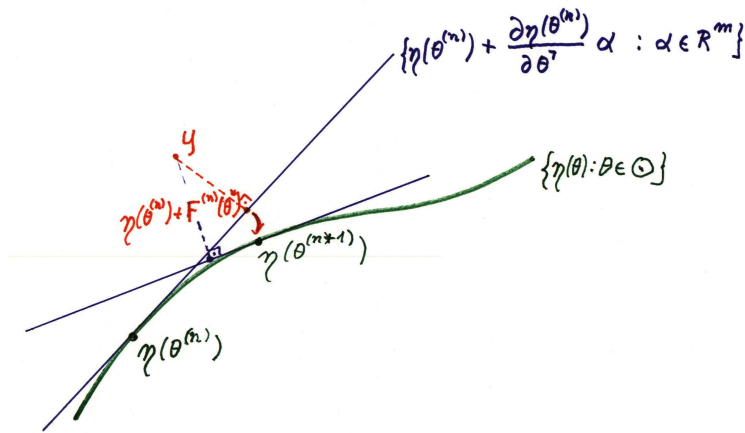
je geometricky interpretovateľná ako *dotyková rovina* ku ploche \mathcal{E} . (Pozri obr. 2 pre prípad $\dim(\theta) = 1$. Na obr. 1 dotyková rovina je určená dotykovými vektormi $\frac{\partial \eta(\theta)}{\partial \theta_1}$, $\frac{\partial \eta(\theta)}{\partial \theta_2}$). Gaussov-Newtonov algoritmus (iteračný) je určený iteračným krokom

$$\theta^{(n+1)} = \theta^{(n)} + \lambda_n v^{(n)},$$

kde λ_n je dĺžka iteračného kroku (tú tu nebudeme rozvádzať), kdežto *smier* iteračného kroku $v^{(n)}$ je určený odhadom $\tilde{\theta}$ vektora θ v linearizovanom modeli, t.j.

$$\begin{aligned} v^{(n)} &= (\tilde{\theta} - \theta^{(n)}) \\ &= \left[\frac{\partial \eta^T(\theta)}{\partial \theta} \frac{\partial \eta(\theta)}{\partial \theta^T} \right]_{\theta=\theta^{(n)}}^{-1} \left[\frac{\partial \eta^T(\theta)}{\partial \theta} \right]_{\theta=\theta^{(n)}} [y - \eta(\theta^{(n)})]. \end{aligned}$$

Podobne postupujeme v ďalšom iteračnom bode $\theta^{(n+1)}$ atď. Aj z názoru z obr. 2 cítíme, že postupnosť $\{\hat{\theta}^{(n)}\}_{n=1}^{\infty}$ sa postupne približuje ku hľadanému odhadu $\hat{\theta}$.



Obrázok 2: Dotykové priamky a geometrická interpretácia linearizácie modelu, ako aj Gaussovej-Newtonovej metódy.

Poznamenávame, že ide o špeciálny prípad tzv. „Fisher scoring“ metódy používané vo všeobecnejších modeloch pre MLE odhady, pretože v prípade normálne rozdelených chýb

$$M(\theta) = \frac{\partial \eta^T(\theta)}{\partial \theta} \frac{\partial \eta(\theta)}{\partial \theta^T} = \sigma^2 E_{\theta} \left[-\frac{\partial^2 \ln f(y | \theta)}{\partial \theta \partial \theta^T} \right]$$

je Fisherova informačná matica pre $\sigma = 1$, a vektor

$$\sigma^{-2} \frac{\partial \eta^T(\theta)}{\partial \theta} [y - \eta(\theta)] = \frac{\partial \ln f(y | \theta)}{\partial \theta}$$

je „scoring“ vektor v bode θ . Vo všeobecnosti možno povedať, že nelineárny regresný model je „pravzorom“ iných nelineárne parametrizovaných modelov, čo tiež motivuje jeho podrobnejšie štúdium.

5 Krivosť plochy \mathcal{E} a možnosť zlyhania MNŠ

Ak odhad

$$\hat{\theta} = \arg \min_{\theta \in \Theta} \|y - \eta(\theta)\|^2$$

leží na $\text{int}(\Theta)$, tak spĺňa podmienku $\frac{\partial}{\partial \theta_i} [\|y - \eta(\theta)\|^2]_{\hat{\theta}} = 0$, $i = 1, \dots, m$, t.j.

$$[\eta(\hat{\theta}) - y]^T \left[\frac{\partial \eta(\theta)}{\partial \theta_i} \right]_{\hat{\theta}} = 0; \quad i = 1, \dots, m.$$

Teda reziduálny vektor $y - \eta(\hat{\theta})$ je kolmý na dotykový priestor ku \mathcal{E} v bode $\hat{\theta}$ (pozri obr. 1). Tato vlastnosť ešte nestačí na to, aby $\hat{\theta}$ bol aspoň bodom lokálneho minima sumy štvorcov $\|y - \eta(\theta)\|^2$. Potrebujeme ešte aj maticu druhých derivácií

$$\begin{aligned} \frac{\partial^2}{\partial\theta_i\partial\theta_j} \left[\frac{1}{2} \|y - \eta(\theta)\|^2 \right]_{\hat{\theta}} &= \\ &= \left[\frac{\partial\eta^T(\theta)}{\partial\theta_i} \frac{\partial\eta(\theta)}{\partial\theta_j} \right]_{\hat{\theta}} + [\eta(\hat{\theta}) - y]^T \left(\frac{\partial^2\eta(\theta)}{\partial\theta_i\partial\theta_j} \right)_{\hat{\theta}} \\ &= M(\hat{\theta}) + [\eta(\hat{\theta}) - y]^T [I - P(\hat{\theta})] \left(\frac{\partial^2\eta(\theta)}{\partial\theta_i\partial\theta_j} \right)_{\hat{\theta}}, \end{aligned}$$

ktorá musí byť pozitívne definitná. Tu $P(\theta)$ označuje ortogonálny projektor na dotykový priestor (generovaný dotykovými vektormi $\frac{\partial\eta(\theta)}{\partial\theta_i}$, pozri obr. 1)

$$P(\theta) = \sum_{i,j} \frac{\partial\eta(\theta)}{\partial\theta_i} [M^{-1}(\theta)]_{i,j} \frac{\partial\eta^T(\theta)}{\partial\theta_j}$$

a tento sme mohli „vpašovať“ do posledného riadku, pretože

$$[I - P(\hat{\theta})] [\eta(\hat{\theta}) - y] = \eta(\hat{\theta}) - y.$$

Ako súvisí táto pozitívna definitnosť s geometriou? Najprv trochu algebry. Pre každé $v \in \mathbb{R}^m$, $v \neq 0$ pozitívna definitnosť vlastne znamená že

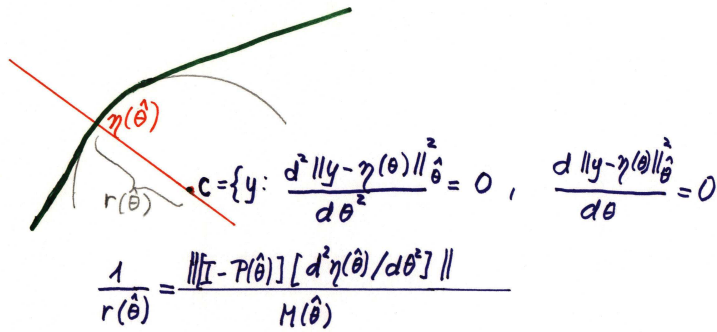
$$\begin{aligned} 0 &< v^T M(\hat{\theta}) v + [\eta(\hat{\theta}) - y]^T [I - P(\hat{\theta})] \left\{ \sum_{i,j} v_i \left(\frac{\partial^2\eta(\theta)}{\partial\theta_i\partial\theta_j} \right) v_j \right\}_{\hat{\theta}} \\ &= v^T M(\hat{\theta}) v \left[1 - \frac{[\eta(\hat{\theta}) - y]^T [I - P(\hat{\theta})] \{\cdot\}}{v^T M(\hat{\theta}) v} \right] \end{aligned}$$

Teda

$$\frac{[\eta(\hat{\theta}) - y]^T [I - P(\hat{\theta})] \{\cdot\}}{v^T M(\hat{\theta}) v} < 1$$

Táto nerovnosť platí triviálne ak $[\eta(\hat{\theta}) - y]^T [I - P(\hat{\theta})] \{\cdot\} < 0$. Na obr. 3 je to prípad, keď bod y leží na opačnej strane krivky \mathcal{E} než stred krivosti C . V opačnom prípade, ak nerovnosť neplatí, môžeme použiť Schwarzovu nerovnosť

$$[\eta(\hat{\theta}) - y]^T [I - P(\hat{\theta})] \{\cdot\} \leq \|y - \eta(\hat{\theta})\| \| [I - P(\hat{\theta})] \{\cdot\} \|$$



Obrázok 3: Stred a polomer krivosti krivky stredných hodnôt modelu. Stred krivosti a jeho okolie je oblasťou nestability MNS.

Teda postačujúcou podmienkou na pozitívnu definitnosť je

$$\|y - \eta(\hat{\theta})\| \left(\frac{\| [I - P(\hat{\theta})] \left\{ \sum_{i,j} v_i \left(\frac{\partial^2 \eta(\theta)}{\partial \theta_i \partial \theta_j} \right) v_j \right\}_{\hat{\theta}} \|}{v^T M(\hat{\theta}) v} \right) < 1$$

pre každé $v \in \mathbb{R}^m$, $v \neq 0$. Možno pomerne jednoducho dokázať (pozri [11], kap. 3.1 a 4.2), že výraz v guľatej zátvorke (\cdot) je obrátený polomer krivosti $r_v(\hat{\theta})$ geodetickej krivky idúcej cez bod $\hat{\theta}$ v smere v . Na obr. 3 je $r_v(\hat{\theta})$ polomer kružnice „najtesnejšie“ sa dotykajúcej sa krivky \mathcal{E} v bode a majúcej svoj stred v bode C . To znamená, že postačujúcou podmienkou na pozitívnu definitnosť je platnosť nerovnosti $\|y - \eta(\hat{\theta})\| < r_v(\hat{\theta})$ pre každú geodeticú krivku na \mathcal{E} idúcu cez bod $\eta(\hat{\theta})$.

Na obr. 3 je ilustrovaný prípad, keď $\dim(\theta) = 1$. Ak $\|y - \eta(\hat{\theta})\| > r_v(\hat{\theta})$ a súčasne bod y leží na tej istej strane krivky stredných hodnôt ako stred krivosti C , tak v bode $\hat{\theta}$ je lokálne maximum sumy štvorcov (namiesto minima). Vo všeobecnosti, ak pravdepodobnosť

$$Pr \left\{ y \in \mathbb{R}^N : \|y - \eta(\hat{\theta}(y))\| > r_v(\hat{\theta}(y)) \right\}$$

je veľká, potom odhad je s veľkou pravdepodobnosťou zmätočný (veľmi vzdialený od skutočnej hodnoty θ , alebo nestabilný). Príslušný experiment treba doplniť ďalšími pozorovaniami, alebo znovu naplánovať.

Použitý výraz

$$K_{int}(\theta, v) = \frac{\| [I - P(\theta)] \left\{ \sum_{i,j} v_i \left(\frac{\partial^2 \eta(\theta)}{\partial \theta_i \partial \theta_j} \right) v_j \right\} \|}{v^T M(\theta) v}$$

sa nazýva *vnútorná krivosť* (alebo *vnútorná miera nelinearity*) v bode θ a v smere v . Do nelineárnej regresie ju zaviedli v [2], ale v inom tvare a s inou motiváciou. Priamy súvis s tu uvedeným vzorcom možno najst' v [11], kap. 5.5.

6 Oblasť spoľahlivosti pre θ hlavne v modeloch s malou vnútornou krivosťou

Ak model je vnútorne lineárny ($K_{int}(\theta, v) = 0$ pre každé θ, v), tak plocha \mathcal{E} je nezakrivená, rovinná. To ešte nemusí znamenať, že regresný model je lineárny, ale znamená to, že reparametrizáciou ho možno zmeniť na lineárny model v nových parametroch β

$$y = F\beta + \varepsilon.$$

Pre takýto model je dobre známy elipsoid spoľahlivosti (pre normálne chyby ε)

$$\mathcal{O} = \left\{ \beta \in \mathbb{R}^m : (\beta - \hat{\beta})^T M (\beta - \hat{\beta}) \leq ms^2 F_{m, N-m} (1 - \alpha) \right\}.$$

Tu $F_{m, N-m} (1 - \alpha)$ je kvantil F-rozdelenia, $s^2 = \frac{\|y - F\hat{\beta}\|^2}{N - m}$, $M = F^T F$. Ak si uvedomíme, že vďaka reparametrizácii, symboly $F\beta$ a $\eta(\theta)$ sú len rôzne označenia toho istého bodu na ploche stredných hodnôt \mathcal{E} , dostávame

$$(\beta - \hat{\beta})^T M (\beta - \hat{\beta}) = \|F\beta - F\hat{\beta}\|^2 = \|\eta(\theta) - \eta(\hat{\theta})\|^2$$

a teda oblasť spoľahlivosti pre pôvodnú (nelineárnu) parametrizáciu modelu je

$$\left\{ \theta : \|\eta(\theta) - \eta(\hat{\theta})\|^2 \leq m \frac{\|y - \eta(\hat{\theta})\|^2}{N - m} F_{m, N-m} (1 - \alpha) \right\}.$$

Spoľahlivosť je presne $(1 - \alpha)$ ak chyby pozorovania sú normálne. Je pozoruhodné ako „lacno“ sme pomocou geometrie získali túto oblasť spoľahlivosti.

Prirodzenou dopĺňujúcou otázkou je, aká oblasť spoľahlivosti sa má použiť ak model nie je vnútorne lineárny. Je možné niekoľko zovšeobecnení uvedenej oblasti, niektoré dávajú síce presne požadovanú spoľahlivosť $1 - \alpha$, ale sú v silne nelineárnych modeloch nepoužiteľné, pretože zahrnujú veľmi výrazne do oblasti spoľahlivosti okolia relatívnych minimí sumy štvorcov. Takouto oblasťou je napr.

$$\left\{ \theta : \|P(\theta)[y - \eta(\theta)]\|^2 \leq m \frac{\|[I - P(\theta)][y - \eta(\theta)]\|^2}{N - m} F_{m, N-m} (1 - \alpha) \right\}$$

Tento nedostatok nemá najčastejšie používaná oblasť spoľahlivosti, ktorá je založená na pomere vierohodnosti a jeho asymptotickom F rozdelení

$$\left\{ \theta : \|y - \eta(\theta)\|^2 - \|y - \eta(\hat{\theta})\|^2 \leq m \frac{\|y - \eta(\hat{\theta})\|^2}{N - m} F_{m, N-m}(1 - \alpha) \right\}.$$

Pri menších súborech meraní a vo vnútorne výrazne nelineárnych modeloch jej spoľahlivosť nie je presne $(1 - \alpha)$, ani pri normálnych chybách pozorovania, napriek tomu sa považuje za omnoho lepšiu než elipsoid spoľahlivosti, vychádzajúci z asymptotickej normality odhadu $\hat{\theta}$ (pozri [2] alebo [12]).

Poznamenávame, že vo vnútorne lineárnych modeloch sa všetky uvedené oblasti zhodujú.

7 Geometrické ohraňovania možnosti lineárizácie nelineárneho regresného modelu vyjadrené pomocou parametrickej krivosti

Analyzujeme oprávnenosť štandardnej linearizácie modelu

$$\begin{aligned} y &= \eta(\theta) + \varepsilon; & \theta &\in \Theta \subset \mathbb{R}^m \\ E(\varepsilon) &= 0, & \text{Var}(\varepsilon) &= \sigma^2 I, \end{aligned}$$

t.j. nahradenie modelom

$$\begin{aligned} y &= \eta(\theta^*) + \left[\frac{\partial \eta(\theta)}{\partial \theta^T} \right]_{\theta=\theta^*} (\theta - \theta^*) + \varepsilon \\ E(\varepsilon) &= 0, & \text{Var}(\varepsilon) &= \sigma^2 I, & \theta^* &= \text{bod linearizácie} \end{aligned}$$

MNS odhad v linearizovanom modeli je dobre známy (pozri tiež časť 4)

$$\begin{aligned} \tilde{\theta} - \theta^* &= \left[\frac{\partial \eta^T(\theta)}{\partial \theta} \frac{\partial \eta(\theta)}{\partial \theta^T} \right]_{\theta^*}^{-1} \left[\frac{\partial \eta^T(\theta)}{\partial \theta} \right]_{\theta^*} [y - \eta(\theta^*)] \\ &= M^{-1} F^T [y - \eta(\theta^*)], \end{aligned}$$

kde zmysel označení M a F vyplýva z predchádzajúcej rovnosti. Podstatnou je výchyľka odhadu $\tilde{\theta}$ posudzovaná v pôvodnom modeli, ktorá sa rovná

$$E_{\tilde{\theta}}(\tilde{\theta}) - \bar{\theta} = (\theta^* - \bar{\theta}) + M^{-1} F^T [\eta(\bar{\theta}) - \eta(\theta^*)].$$

Pripomínáme, že $\bar{\theta}$ je skutočná hodnota parametra θ . Približne platí (kvadratická Taylorova formula)

$$[\eta(\bar{\theta}) - \eta(\theta^*)] \doteq F(\bar{\theta} - \theta^*) + \left\{ \sum_{i,j} (\bar{\theta} - \theta^*)_i \left[\frac{\partial^2 \eta(\theta)}{\partial \theta_i \partial \theta_j} \right]_{\theta^*} (\bar{\theta} - \theta^*)_j \right\}.$$

Teda po dosadení dostaneme pre výchylku $E_{\bar{\theta}}(\tilde{\theta}) - \bar{\theta} \doteq M^{-1}F^T\{\cdot\}$. Variančná matica odhadu $\tilde{\theta}$ je evidentne $Var(\tilde{\theta}) = \sigma^2 M^{-1}$. Teda porovnanie výchylky s varianciou dáva

$$\begin{aligned} & \left[E_{\bar{\theta}}(\tilde{\theta}) - \bar{\theta} \right]^T Var^{-1}(\tilde{\theta}) \left[E_{\bar{\theta}}(\tilde{\theta}) - \bar{\theta} \right] \doteq \sigma^{-2} \{\cdot\}^T F M^{-1} F^T \{\cdot\} \\ &= \sigma^{-2} \|P(\theta^*)\{\cdot\}\|^2 \\ &\leq \left[(\bar{\theta} - \theta^*)^T [\sigma^{-2} M] (\bar{\theta} - \theta^*) \right]^2 \times \\ &\quad \times \sigma^2 \frac{\left\| P(\theta^*) \left\{ \sum_{i,j} (\bar{\theta} - \theta^*)_i \frac{\partial^2 \eta(\theta)}{\partial \theta_i \partial \theta_j} (\bar{\theta} - \theta^*)_j \right\}_{\theta^*} \right\|^2}{\left\{ (\bar{\theta} - \theta^*)^T M (\bar{\theta} - \theta^*) \right\}^2} \\ &= [\text{Mahalanobisova vzdial.}]^2 \times \sigma^2 \times [\text{parametrická miera nelinearity}]^2. \end{aligned}$$

Tu me využili, že

$$K_{par}(\theta, v) = \frac{\left\| P(\theta) \left\{ \sum_{i,j} v_i \left(\frac{\partial^2 \eta(\theta)}{\partial \theta_i \partial \theta_j} \right) v_j \right\} \right\|}{v^T M(\theta) v}$$

je parametrickou krivosťou (mierou nelinearity) v bode θ a v smere v . Linearizácia modelu je teda dobrá, ak alebo $\bar{\theta}$ je „Mahalanobisovsky“ blízko ku θ^* = referenčnému bodu linearizácie, alebo ak σ^2 (= disperzia chyby pozorovania) je malá, alebo ak parametrická krivosť v bode θ^* a v smere $(\bar{\theta} - \theta^*)$ je malá.

Parametrická krivosť, hoci je formálne podobná vnútornej krivosti, má značne odlišnú interpretáciu. Dobré sa dá vysvetliť v prípade $\dim(\theta) = 1$ (pozri [2], kde táto krivosť bola prvýkrát uvedená, alebo [11] kap. 3.1). V takom prípade plocha stredných hodnôt \mathcal{E} je proste krivkou. Označme symbolom $\beta = \beta(\theta)$ jej dĺžku. Algebraicky táto je určená diferenciálnou rovnicou $\frac{d\beta(\theta)}{d\theta} = \left\| \frac{d\eta(\theta)}{d\theta} \right\|$ a práve pomocou nej sa dá ukázať, že platí

$$K_{par}(\theta) = \frac{\left\| P(\theta) \left\{ \left(\frac{d^2 \eta(\theta)}{d\theta^2} \right) \right\} \right\|}{M(\theta)} = \frac{\left| \frac{d^2 \beta(\theta)}{d\theta^2} \right|}{\left| \frac{d\beta(\theta)}{d\theta} \right|^2}.$$

(O nejakom smere v sa v prípade $\dim(\theta) = 1$ nedá hovoriť a „matica“ $M(\theta)$ je proste číslo). Teda parametrická krivosť meria zrýchlenie dĺžky β voči parametru θ . V prípade ľubovolnej $\dim(\theta)$ veľká hodnota $K_{par}(\theta, v)$ signalizuje nerovnomernosť parametrizácie θ na ploche stredných hodnôt \mathcal{E} . Práve táto nerovnomernosť spôsobuje, že aj v prípade normálnych chýb pozorovania, odhad $\hat{\theta}$ nemá normálne rozdelenie. Taktiež, asymptotická normalita odhadu $\hat{\theta}$ sa dosahuje rýchlejšie, ak parametrická krivosť je malá.

8 Ako možno geometricky vysvetliť asymptotickú normalitu MNŠ odhadov

Predpoklady

$$\begin{aligned} y_i &= \eta(x_i, \theta) + \varepsilon_i; \quad i = 1, 2, \dots \\ \text{Var}(\varepsilon_i) &= \sigma^2, \varepsilon_i \text{ nezávislé, rovnako rozdelené,} \\ x_i &\in X = \text{konečná množina} \end{aligned}$$

Nech $\frac{N(x)}{N}$ je relatívna početnosť bodu x v postupnosti x_1, x_2, \dots, x_N a nech pre každé $x \in X$ existuje

$$\lim_{N \rightarrow \infty} \frac{N(x)}{N} = \xi(x)$$

kde ξ je nejaká pravdepodobnostná miera na X (= design experimentu, pozri referát R. Harmana v zborníku). Označme

$$\hat{\theta}^{(N)} = \arg \min_{\theta \in \Theta} \sum_{i=1}^N [y_i - \eta(x_i, \theta)]^2$$

Tvrdenie o asymptotickej normalite: *Pre veľké N približne platí*

$$\hat{\theta}^{(N)} \sim \mathcal{N}(\bar{\theta}, \sigma^2 [NM(\xi, \bar{\theta})]^{-1})$$

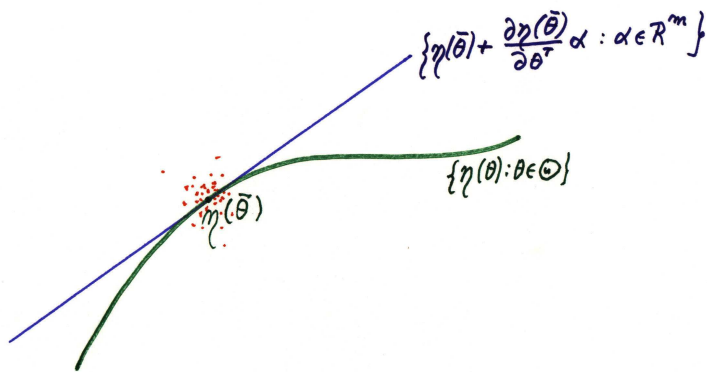
kde $\bar{\theta} = \text{skutočné } \theta$,

$$M(\xi, \theta) = \sum_{x \in X} \frac{\partial \eta(x, \theta)}{\partial \theta} \frac{\partial \eta(x, \theta)}{\partial \theta^T} \xi(x)$$

Toto tvrdenie vyzerá komplikované, už aj preto, že odhad $\hat{\theta}^{(N)}$ nevieme vyjadriť nejakým vzorcom. Skutočne, presné dôkazy sú veľmi komplikované, pravda obsahujú aj výroky o rýchlosti a spôsobe konvergencie. Ovšem akákoľvek intuícia je v nich utopená v záplave pomocných tvrdení. V skutočnosti je situácia pomerne jednoduchá, o čom nás presvedčuje geometria. Tá navyše hneď naznačuje, kam sa majú uberať snahy o spresnenia rozdelenia odhadu $\hat{\theta}^{(N)}$ (pre nie príliš veľké N).

Zdôvodnenie tvrdenia: Trochu algebry:

$$\begin{aligned} \sum_{i=1}^N [y_i - \eta(x_i, \theta)]^2 &= \sum_{i=1}^N y_i^2 - 2 \sum_{x \in X} \sum_{\substack{i=1 \\ x_i=x}}^N y_i \eta(x, \theta) + \sum_{x \in X} N(x) [\eta(x, \theta)]^2 \\ &= \sum_{i=1}^N y_i^2 - \sum_{x \in X} N(x) [\bar{y}_N(x)]^2 + \\ &\quad + \sum_{x \in X} N(x) [\bar{y}_N(x) - \eta(x, \theta)]^2, \end{aligned}$$



Obrázok 4: Asymptotická koncentrácia simulovaných vektorov y v okolí bodu $\eta(\bar{\theta})$.

kde

$$\bar{y}_N(x) = \frac{1}{N(x)} \sum_{\substack{i=1 \\ x_i=x}}^N y_i, \quad E_{\theta}[\bar{y}_N(x)] = \eta(x, \theta), \quad \text{Var}[\bar{y}_N(x)] = \frac{\sigma^2}{N(x)}.$$

Teda pre veľké N

$$\begin{aligned} \hat{\theta}^{(N)} &= \arg \min_{\theta \in \Theta} \sum_{x \in X} \frac{N(x)}{N} [\bar{y}_N(x) - \eta(x, \theta)]^2 \\ &\doteq \arg \min_{\theta \in \Theta} \sum_{x \in X} \xi(x) [\bar{y}_N(x) - \eta(x, \theta)]^2. \end{aligned}$$

Teda namiesto vektora pozorovaných dát y stačí uvažovať vektor priemerov

$$(\bar{y}_N(x) : x \in X, \xi(x) > 0)^T.$$

Naviac, z centrálnej limitnej vety vyplýva, že pre veľké N platí

$$\bar{y}_N(x) \sim \mathcal{N}\left(\eta(x, \bar{\theta}), \frac{\sigma^2}{N\xi(x)}\right)$$

kde $\bar{y}_N(x)$ sú nezávislé pre rôzne $x \in X$. Tieto teda vyhovujú nelineárnemu regresnému modelu

$$\begin{aligned} \bar{y}_N(x) &= \eta(x, \theta) + \varepsilon(x); \quad x \in X, \xi(x) > 0 \\ \text{Var}(\varepsilon(x)) &= \frac{\sigma^2}{N\xi(x)}. \end{aligned}$$

Pri veľkom N , body $\bar{y}_N(x)$ sú koncentrované okolo $\eta(x, \bar{\theta})$, teda plochu stredných hodnôt posledného modelu môžeme aproximovať jej dotykovou rovinou (pozri obr. 4). Ináč povedané, pre veľké N je oprávnený lineárny model s normálnymi chybami

$$\bar{y}_N(x) = \eta(x, \bar{\theta}) + \left[\frac{\partial \eta(x, \theta)}{\partial \theta^T} \right]_{\bar{\theta}} (\theta - \bar{\theta}) + \varepsilon(x); \quad x \in X,$$

a MNS odhad je približne

$$\begin{aligned} \hat{\theta}^{(N)} - \bar{\theta} &\doteq \left\{ \sum_{x \in X} \left[\frac{\partial \eta(x, \theta)}{\partial \theta} \frac{\partial \eta(x, \theta)}{\partial \theta^T} \right]_{\bar{\theta}} \xi(x) \right\}^{-1} \times \\ &\times \sum_{x \in X} \left[\frac{\partial \eta(x, \theta)}{\partial \theta} \right]_{\bar{\theta}} [\bar{y}_N(x) - \eta(x, \bar{\theta})] \xi(x) \\ &\sim \mathcal{N} \left(0, \sigma^2 [NM(\xi, \bar{\theta})]^{-1} \right) \end{aligned}$$

čo sme chceli ukázať.

9 Geometrické odvodenie neasymptotickej hustoty odhadu $\hat{\theta}$

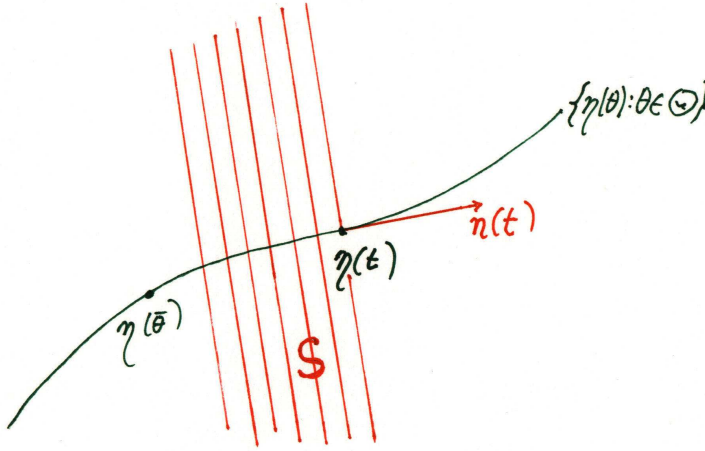
Existuje aj iná, podstatne presnejšia aproximácia hustoty pravdepodobnosti odhadu $\hat{\theta}$ než jej normálna asymptotická hustota. Táto má tvar

$$\begin{aligned} q(\hat{\theta} | \bar{\theta}) &= \frac{\det [Q(\hat{\theta}, \bar{\theta})]}{[2\pi]^{m/2} \sigma^m \det^{1/2} [M(\hat{\theta})]} \times \\ &\times \exp \left\{ \left\| P(\hat{\theta}) [\eta(\hat{\theta}) - \eta(\theta)] \right\|^2 \right\} \end{aligned}$$

$$\text{kde } Q(\hat{\theta}, \bar{\theta}) = M(\hat{\theta}) + [\eta(\hat{\theta}) - \eta(\theta)]^T [I - P(\hat{\theta})] \left[\frac{\partial^2 \eta(\theta)}{\partial \theta \partial \theta^T} \right]_{\hat{\theta}}.$$

Odvodenie, ak $\dim(\theta) = 1$: Distribučná funkcia v nejakom bode $t \in \Theta$ sa približne rovná (pozri obr. 5)

$$\begin{aligned} F(t) &\doteq \Pr(y \in S) = \Pr \left\{ y : [y - \eta(t)]^T \frac{\partial \eta(t)}{\partial t} < 0 \right\} \\ &= \Pr \left\{ y : [y - \eta(\bar{\theta})]^T n(t) < [\eta(t) - \eta(\bar{\theta})]^T n(t) \right\} \\ &= \int_{-\infty}^{[\eta(t) - \eta(\bar{\theta})]^T n(t)} \frac{1}{[2\pi]^{1/2} \sigma} \exp \left\{ -\frac{u^2}{2\sigma^2} \right\} du \end{aligned}$$



Obrázok 5: Odvodenie hustoty pravdepodobnosti v bode t ak $\dim(\theta) = 1$.

kde $n(t) = \frac{\partial \eta(t)}{\partial t} / \left\| \frac{\partial \eta(t)}{\partial t} \right\|$ je jednotkový vektor. Derivovaním d.f. $F(t)$ dostaneme požadovanú hustotu,

$$q(\hat{\theta} | \bar{\theta}) = \frac{dF(\hat{\theta})}{d\hat{\theta}}$$

pravda po pilnom upravovaní výsledných vzorcov. Je geometricky jasné, v čom spočíva aproximácia pri odvodení $q(\hat{\theta} | \bar{\theta})$. V tom, že pre rôzne body $t \in \Theta$ sa polroviny S prekrývajú. Ale tento prekryt nastáva až pre body veľmi vzdialené od \mathcal{E} , presnejšie až v tej oblasti, ktorá je viac vzdialená od plochy \mathcal{E} , než sú polomery krivosti geodetických čiar na \mathcal{E} . Ale z časti 5 vieme, že táto oblasť musí mať takmer nulovú pravdepodobnosť, ak MNS odhad $\hat{\theta}$ má vôbec mať nejaký zmysel. Vzhľadom na tento fakt, možno považovať hustotu $q(\hat{\theta} | \bar{\theta})$ za „takmer presnú“.

Odvodenie hustoty $q(\hat{\theta} | \bar{\theta})$ pre $\dim(\theta) > 1$ je podstatne zložitejšie (pozri [9] alebo [11]). Krátko po [9] (a nezávisle) bola hustota $q(\hat{\theta} | \bar{\theta})$ odvodená aj asymptoticky, tzv. saddlepoint metódou [6]. Nie je prekvapujúce, že rýchlosť konverencie je tu omnoho väčšia ako pri asymptotickej normalite. Geometrické metódy umožňujú však tu nahradiť asymptotické metódy, s nejasnou voľbou počtu pozorovaní N v aplikáciach, inými metódami, ktoré pri znalosti geometrie plochy \mathcal{E} umožňujú určiť, či aproximácia pomocou $q(\hat{\theta} | \bar{\theta})$ je dobrá alebo nie.

Vyšetrovanie hustoty pre $\hat{\theta}$ ukázalo ešte jeden zaujímavý fakt: pre $\dim(\theta) > 1$ je $q(\hat{\theta} | \bar{\theta})$ takmer presnou aproximáciou len ak Riemannov tenzor krivosti s komponentmi

$$R_{ikjl}(\theta) \equiv \frac{\partial^2 \eta^T(\theta)}{\partial \theta_i \partial \theta_j} [I - P(\theta)] \frac{\partial^2 \eta(\theta)}{\partial \theta_k \partial \theta_l} - \frac{\partial^2 \eta^T(\theta)}{\partial \theta_i \partial \theta_l} [I - P(\theta)] \frac{\partial^2 \eta(\theta)}{\partial \theta_k \partial \theta_j}$$

je nulový (pre $\dim(\theta) = 1$ to platí triviálne). V opačnom prípade aj on vystupuje vo výraze pre takmer presnú hustotu. Napr. ak $\dim(\theta) = 2$, tak „takmer presná“ hustota je [11]

$$q(\hat{\theta} | \bar{\theta}) + \text{const} \frac{R_{1221}(\hat{\theta})}{\det^{1/2}[M(\hat{\theta})]} \exp \left\{ \left\| P(\hat{\theta}) [\eta(\hat{\theta}) - \eta(\theta)] \right\|^2 \right\}$$

a situácie je ešte komplikovanejšia pre väčšie dimenzie parametra θ .

Je zaujímavé, že technika geometrického odvodenia $q(\hat{\theta} | \bar{\theta})$ sa dá veľmi dobre použiť aj pri odvodzovaní hustoty penalizovaného MNŠ odhadu, čo napríklad bolo nevyhnutné pre algoritmus optimálneho navrhovania experimentov [5]. Alebo ak vektor chýb ε nie je normálne rozdelený ale má iba sféricky (resp. elipticky) symetrické rozdelenie [10].

10 Problémy jednoznačnosti MNŠ odhadov

Platí nasledujúca veta (pozri [8] alebo [11], kap. 4.4):

Veta. *Nech rozdelenie pravdepodobnosti vektora chýb ε je absolútne spojité vzhľadom na Lebesgueovu mieru na \mathbb{R}^N , kde N je počet pozorovaní. Potom*

$$\begin{aligned} \Pr \left\{ y \in \mathbb{R}^N : \exists_{\eta(\bar{\theta}) \neq \eta(\hat{\theta})} \left\| y - \eta(\bar{\theta}) \right\|^2 = \right. \\ \left. \left\| y - \eta(\hat{\theta}) \right\|^2 = \arg \min_{\theta \in \Theta} \left\| y - \eta(\theta) \right\|^2 \right\} \\ = 0. \end{aligned}$$

Ináč povedané, MNŠ odhad $\eta(\hat{\theta})$ je určený jednoznačne skoro iste.

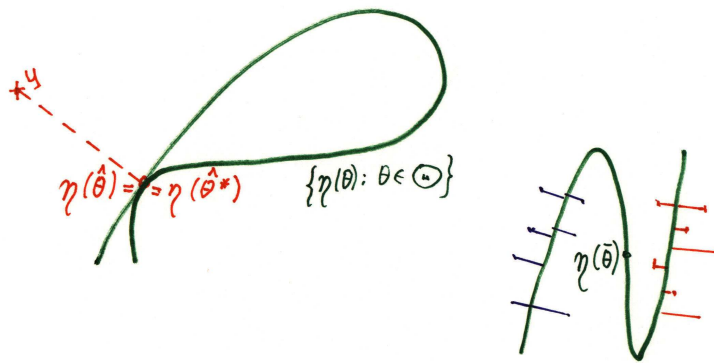
Poznamednávame, že nepredpokladáme normalitu chýb pozorovania. Dôkaz vyžaduje tvrdenia z diferenciálnej geometrie. V prvom rade nasledujúcu vetu:

Veta Sardova: *Nech zobrazenie*

$$\psi : U \subset \mathbb{R}^s \rightarrow {}^{do} \mathbb{R}^s$$

je spojitě diferencovateľné, potom

$$\text{Leb. miera} \left\{ \psi(x) : \det \left[\frac{\partial \psi^T(x)}{\partial x} \right] = 0 \right\} = 0.$$



Obrázok 6: Krivka stredných hodnôt v prípade nejednoznačnosti zobrazenia $\theta \in \Theta \rightarrow \eta(\theta)$, resp. v prípade výrazných lokálnych miním.

Pomocou tejto vety možno ukázať, že množina stredov krivosti geodetických kriviek na ploche stredných hodnôt \mathcal{E} má nulovú Lebesgueovu mieru. A my už vieme z časti 5, že stredy krivosti sú bodmi nejednoznačnosti odhadov $\eta(\hat{\theta})$. Ďalší postup vyžaduje vylúčiť ďalšie nejednoznačnosti. V podstate sa určitým, nie celkom triviálnym trikom dokáť že, že množina bodov $\hat{\theta}$, ktoré dávajú nejednoznačné $\eta(\hat{\theta})$, sa dá regulárne zobraziť na $m-1$ dimenzionálnu podmnožinu priestoru \mathbb{R}^m , a teda má nulovú Lebesgueovu mieru. Nulová pravdepodobnosť potom vyplýva z predpokladanej absolutnej spojitosti.

Tento geometrický postup bol zovšeobecnený aj na odhad maxima vierohodnosti v tzv. zakrivených exponenciálnych triedach, a celkom nedávno istý nemecký matematik [7] ho ďalej zovšeobecnil na triedu tzv. kontrastných odhadov.

Na druhej strane, jednoznačnosť určenia $\eta(\hat{\theta})$ neznamená, že nie sú ťažkosti s jednoznačnosťou odhadu $\hat{\theta}$. Tieto môžu byť spôsobené jednak singularitou modelu, jednak nejednoznačnosťou zobrazenia $\theta \in \Theta \rightarrow \eta(\theta)$, a nakoniec existenciou viacerých výrazných lokálnych miním funkcie $\theta \in \Theta \rightarrow \|y - \eta(\theta)\|^2$. V kontraste s hore uvedenou vetou je dokázané, že pokiaľ model nie je vnútorne lineárny, pravdepodobnosť nastatia aspoň dvoch lokálnych miním, je vždy kladná (pozri [4]).

Tieto otázky nejednoznačnosti sú komplikované a zatiaľ sa dajú čiastočne zodpovedať len fisherovskými alebo bayesovskými úvahami o funkcii vierohodnosti. Dôležité je však, že ich kvalitatívne geometrické pochopenie je vlastne jednoduché, aspoň v prípade $\dim(\theta) = 1$. Situácia keď zobrazenie $\theta \in \Theta \rightarrow \eta(\theta)$ nie je jednoznačné, alebo keď funkcia $\theta \in \Theta \rightarrow \|y - \eta(\theta)\|^2$

má s veľkou pravdepodobnosťou niekoľko význačných lokálnych miním, je ilustrovaná na obr. 6.

11 Citát klasika

Joseph Louis Lagrange: *Pokiaľ algebra a geometria išli po oddelených cestách, ich pokrok bol pomalý a aplikácie obmedzené. Ale keď sa tieto vedy združili, získali jedna od druhej sviežu vitalitu a napredovali rýchlo ku dokonalosti.*

Literatúra

- [1] Baranyi J., Roberts T.A. (1994). *A dynamic approach to predicting bacterial growth in food*. Internat. J. Food Microbiol. **23**, 277–294.
- [2] Bates D.M., Watts D.G. (1980). *Relative curvature measures of nonlinearity* (with discussion). J. Roy. Statist. Soc., Ser. B. **42**. 1–25.
- [3] Bates D.M., Watts D.G. (1988). *Nonlinear Regression Analysis and its Applications*. Wiley, New York.
- [4] Demidenko E. (2000). *Is this the least squares estimate?* Biometrika **87**, 437–452.
- [5] Gauchi J.-P., Pázman A. (2005). *Designs in nonlinear regression by stochastic minimization of functionals of the mean square error matrix*. J. Statist. Planning and Inference **236**, 1135–1152.
- [6] Hougaard, P. (1985). *Saddlepoint approximations for curved exponential families*. Statistics & Prob. Letters **3**, 161–166.
- [7] Kratschmer V. (2007). *The uniqueness of extremum estimation*. Statist. & Prob. Letters **77**, 942–991.
- [8] Pázman A. (1984). *Nonlinear least squares - uniqueness versus ambiguity*. Math. Operationsforschung & Stat., Serie Statistics **15**, 323–336.
- [9] Pázman A. (1984). *Probability distribution of the multivariate nonlinear least squares estimates*. Kybernetika **20**, 209–230.
- [10] Pázman A. (1988). *Distribution of the weighted L.S. estimates in nonlinear models with symmetrical errors*. Kybernetika **24**, 413–427
- [11] Pázman A. (1993). *Nonlinear statistical models*. Kluwer, Dordrecht.
- [12] Seber and Wild, (1989). *Nonlinear Regression*. Wiley, New York.

Adresa: A. Pázman, Katedra aplikovanej matematiky a štatistiky, Fakulta matematiky, fyziky a informatiky, Univerzita Komenského, 842 48 Bratislava
E-mail: pazman@fmph.uniba.sk

CHANGE DETECTION IN AR PROCESSES

Jakub Pečánka

Keywords: change point detection, autoregressive processes, asymptotics.

Abstract: Autoregressive time series models (*AR*) of given order p have $p+2$ parameters; the mean μ , the variance σ^2 and the p autoregressive parameters $\varphi_1, \dots, \varphi_p \in \mathbb{R}$. Detection of change in the value of any of these parameters is an important problem. This article presents some results from the theory of change point detection in *AR* processes. The attention is given to statistics based on the sums of weighted residuals, which are in the setting of the *AR* model closely related to the maximum likelihood ratio statistics. We also present results of our computer simulations.

Abstrakt: Autoregressní modely časových řad (*AR*) řádu p mají $p+2$ parametrů. Jsou jimi střední hodnota μ , rozptyl σ^2 and p autoregressních parametrů $\varphi_1, \dots, \varphi_p$. Detekce změn hodnot těchto parametrů je důležitý problém. Tento článek obsahuje některé výsledky z teorie detekce těchto změn. Pozornost je věnována asymptotice statistik založených na částečných součtech vážených reziduí. Článek obsahuje výsledky počítačových simulací.

1 Introduction

In [3] we studied the asymptotic distribution of several statistics for detecting change point. Those statistics are based on partial sums of weighted residuals and the results concerning their asymptotic distribution were given in [2]. The article [2] also contained several lemmas with their proofs, but the proofs of the main theorems were only sketchy. In [3] we formulated several additional lemmas which together with the results by [2] constitute a rigorous proof of the Theorem 2 of the present article. We formulate the main theorems and two general lemmas and provide the proofs of the lemmas.

2 Theoretical results

We assume to have a set of observations Y_1, \dots, Y_n from an $AR(p)$ process, $p \in \mathbb{N}$ fixed. We define a recursive equation, which represents the null hypothesis, by

$$Y_t - \mu = \varphi_1(Y_{t-1} - \mu) + \dots + \varphi_p(Y_{t-p} - \mu) + \varepsilon_t, \quad p < t, \quad (1)$$

where $\mu, \sigma^2, \varphi_1, \dots, \varphi_p$ are fixed constants, $\{\varepsilon_t\}$ is a white noise sequence with variance σ^2 and Y_1, \dots, Y_p are some initial values. The equation corresponding to the alternative is

$$Y_t = \begin{cases} \varphi_0 + \varphi_1 Y_{t-1} + \dots + \varphi_p Y_{t-p} + \varepsilon_t, & p < t \leq k^*, \\ \psi_0 + \psi_1 Y_{t-1} + \dots + \psi_p Y_{t-p} + \varepsilon_t, & k^* < t, \end{cases} \quad (2)$$

with $p < k^* \leq n$. Thus, the pair of hypotheses in question is $H_0 : k^* = n$ and $H_1 : k^* < n$, or more precisely, we define the *null hypothesis condition*

(H.1) *The observations Y_{p+1}, \dots, Y_n follow the model (1) with $k^* = n$; The observations Y_1, \dots, Y_p are independent of $\varepsilon_{p+1}, \dots, \varepsilon_n$; The characteristic polynomial $\phi(z) = 1 - \varphi_1 z - \dots - \varphi_p z^p$ has all roots outside the unit ball (causality, stationary solution).*

and the *alternative hypothesis condition* we formulate as

(H.2) *The observations Y_{p+1}, \dots, Y_n follow the model (2) with $k^* = \lceil \tau n \rceil$, for some fixed $\tau \in (0, 1)$; The observations Y_1, \dots, Y_p are independent of the innovations $\varepsilon_{p+1}, \dots, \varepsilon_n$; The polynomials $\phi_1(z) = 1 - \varphi_1 z - \dots - \varphi_p z^p$ and $\phi_2(z) = 1 - (\varphi_1 + \delta_1)z - \dots - (\varphi_p + \delta_p)z^p$ have all roots outside the unit ball (causality).*

In both cases the initial observations Y_1, \dots, Y_p in the recursive equation (1) are generated from $\{\varepsilon_t\}$ according to

(H.3) *The vector of observations $\mathbf{x}_{p+1} = (Y_p, \dots, Y_1)'$ satisfies $\mathbf{x}_{p+1} - \boldsymbol{\mu} = \sum_{j=0}^{\infty} \mathbf{B}^j \boldsymbol{\varepsilon}_{p-j}$, where*

$$\mathbf{B} = \begin{pmatrix} \varphi_1, \dots, \varphi_p \\ \mathbf{I}_{p-1} & \mathbf{0} \end{pmatrix} \quad \text{and} \quad \boldsymbol{\varepsilon}_k = (\varepsilon_k, 0, \dots, 0)',$$

with \mathbf{I}_{p-1} denoting the $(p-1)$ -dimensional unit matrix and $\mathbf{0}$ denoting the $(p-1)$ -dimensional zero vector.

In the following we assume that the AR process is centered, i.e. $\boldsymbol{\mu} = \mathbf{0}$, and we denote $\mathbf{x}_{t+1} = (Y_t, \dots, Y_{t-p})'$. Additionally, we denote by $\|\cdot\|$ the p -dimensional Euclidean norm and $\{\mathbf{B}(t), t \in [0, 1]\}$ is a p -dimensional standard Brownian bridge process with independent components, i.e. $\mathbf{B}(t) = (B_1(t), \dots, B_p(t))'$ with $\{B_i(t), t \in [0, 1]\}$ being a standard Brownian bridge for all $i = 1, \dots, p$.

2.1 Test statistics

In [2] the asymptotic behavior of several test statistics based on *partial sums of weighted residuals* were investigated. We define

$$\begin{aligned} T_n &= \frac{1}{\widehat{\sigma}_n^2} \max_{p < k < n} \{ \mathbf{S}'_k \mathbf{C}_k^{-1} \mathbf{C}_n (\mathbf{C}_k^0)^{-1} \mathbf{S}_k \}, \\ T_n(\epsilon) &= \frac{1}{\widehat{\sigma}_n^2} \max_{n\epsilon \leq k \leq n(1-\epsilon)} \{ \mathbf{S}'_k \mathbf{C}_k^{-1} \mathbf{C}_n (\mathbf{C}_k^0)^{-1} \mathbf{S}_k \}, \\ T_n(q, \epsilon) &= \frac{1}{\widehat{\sigma}_n^2} \sup_{t \in [\epsilon, 1-\epsilon]} \left\{ \frac{\mathbf{S}'_{\lfloor (n+1)t \rfloor} \mathbf{C}_n^{-1} \mathbf{S}_{\lfloor (n+1)t \rfloor}}{q^2(t)} \right\}, \end{aligned}$$

$$T_n(q) = \frac{1}{\widehat{\sigma}_n^2} \sup_{t \in (0,1)} \left\{ \frac{\mathbf{S}'_{\lfloor (n+1)t \rfloor} \mathbf{C}_n^{-1} \mathbf{S}_{\lfloor (n+1)t \rfloor}}{q^2(t)} \right\},$$

where $\epsilon \in (0, \frac{1}{2})$, $q(t)$ is a positive weight function on $(0, 1)$ and $\widehat{\sigma}_n^2$ is a suitable estimator¹ of the variance parameter σ^2 and for $k = p + 1, \dots, n$

$$\begin{aligned} \mathbf{S}_k &= \sum_{t=p+1}^k \mathbf{x}_t (Y_t - \mathbf{x}'_t \boldsymbol{\varphi}_n), & \boldsymbol{\varphi}_n &= \mathbf{C}_n^{-1} \sum_{t=p+1}^n \mathbf{x}_t Y_t, \\ \mathbf{C}_k &= \sum_{t=p+1}^k \mathbf{x}_t \mathbf{x}'_t, & \mathbf{C}_k^0 &= \sum_{t=k+1}^n \mathbf{x}_t \mathbf{x}'_t = \mathbf{C}_n - \mathbf{C}_k. \end{aligned}$$

The statistics T_n is derived as maximum likelihood ratio statistics for testing for change in $\varphi_1, \dots, \varphi_p$ under the assumption of known variance σ^2 . The other three statistics are various modifications of T_n .

2.2 Asymptotics

The following two theorems give the asymptotic distribution of the statistics T_n , $T_n(\epsilon)$, $T_n(q, \epsilon)$ and $T_n(q)$. The formulation of the first theorem originates in [1] and that of the second one in [2].

Theorem 1 (Asymptotics of T_n). *Let assumptions (H.1) and (H.3) be satisfied and let $\widehat{\sigma}^2$ be an estimator of σ^2 such that $\widehat{\sigma}^2 - \sigma = o_{\mathbb{P}}((\log \log n)^{-1})$ as $n \rightarrow \infty$. Furthermore, let $\{\varepsilon_t, t \in \mathbb{Z}\}$ be i.i.d. centered random variables with positive variance σ^2 and finite moment $\mathbb{E}|\varepsilon_t|^{4+\nu}$, for some $\nu > 0$, with density f such that*

$$\sup_{a \in \mathbb{R}} \frac{1}{|a|} \int_{\mathbb{R}} |f(x+a) - f(x)| dx < \infty.$$

Then, for all $t \in \mathbb{R}$ it holds

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P} \left(A(\log n) T_n^{1/2} - D_p(\log n) \leq t \right) &= \exp \{ -2e^{-t} \}, \\ \lim_{n \rightarrow \infty} \mathbb{P} \left([A^2(\log n) T_n - D_p^2(\log n)] / D_p(\log n) \leq t \right) &= \exp \{ -2e^{-t/2} \}. \end{aligned}$$

Theorem 2 (Asymptotics of $T_n(\epsilon)$, $T_n(q, \epsilon)$ and $T_n(q)$). *Let assumptions (H.1) and (H.3) be satisfied and let $\widehat{\sigma}^2$ be a consistent estimator of σ^2 . Furthermore, let $\{\varepsilon_t, t \in \mathbb{Z}\}$ be i.i.d. centered random variables with positive variance σ^2 and finite fourth moment $m_4 = \mathbb{E}|\varepsilon_t|^4$. Then, for any $\epsilon \in (0, \frac{1}{2})$, it holds*

¹For details on suitable estimator of variance as well as an estimator of the change point see for example [2] or [3].

$$\sqrt{T_n(\epsilon)} \rightarrow_{\mathcal{D}} \sup_{t \in [\epsilon, 1-\epsilon]} \frac{\|\mathbf{B}(t)\|}{\sqrt{t(1-t)}}.$$

If, in addition, $q(t)$ is a function that satisfies $\inf\{q(t); t \in (\epsilon, 1 - \epsilon)\} > 0$, then for any $\epsilon \in (0, 1)$ it holds

$$\sqrt{T_n(q, \epsilon)} \rightarrow_{\mathcal{D}} \sup_{t \in [\epsilon, 1-\epsilon]} \frac{\|\mathbf{B}(t)\|}{q(t)}.$$

If, in addition, a function $q_\beta(t)$ is defined as $q_\beta(t) = (t(1-t))^\beta$, $t \in (0, 1)$, with $\beta \in [0, \frac{1}{2})$, then also

$$\sqrt{T_n(q_\beta)} \rightarrow_{\mathcal{D}} \sup_{t \in (0,1)} \frac{\|\mathbf{B}(t)\|}{q_\beta(t)}.$$

For the lack of available space we do not provide the proofs of the previous two theorems. We refer the reader to [1] for the proof of Theorem 1. For the proof of Theorem 2 see [3] and [2]. The proof in [3] utilizes several general lemmas which have formulations independent of the the setting of Theorem 2. In this article we give formulations and proofs of two of these lemmas.

Lemma 3. Let $S_n(\epsilon)$, $S(\epsilon)$, S_n and S be random variables such that for all $\epsilon > 0$ it holds

i) $S_n(\epsilon) \rightarrow_{\mathcal{D}} S(\epsilon)$, as $n \rightarrow \infty$,

ii) $\lim_{\epsilon \rightarrow 0+} S(\epsilon) = S$, almost surely,

iii) for all $\delta > 0$

$$\lim_{\epsilon \rightarrow 0+} \limsup_{n \rightarrow \infty} \mathbb{P}(|S_n(\epsilon) - S_n| > \delta) = 0.$$

Then, it also holds that $S_n \rightarrow_{\mathcal{D}} S$, as $n \rightarrow \infty$.

Proof. Under the assumptions of the lemma, for any $x \in \mathbb{R}$ at which the distribution function $\mathbb{P}(S \leq x)$ is continuous, it holds

$$\begin{aligned} \mathbb{P}(S_n \leq x) &= \mathbb{P}(S_n \leq x, |S_n(\epsilon) - S_n| \leq \delta) + \mathbb{P}(S_n \leq x, |S_n(\epsilon) - S_n| > \delta) \\ &\leq \mathbb{P}(S_n \leq x, |S_n(\epsilon) - S_n| \leq \delta) + \mathbb{P}(|S_n(\epsilon) - S_n| > \delta) \\ &\leq \mathbb{P}(S_n(\epsilon) \leq x + \delta) + \mathbb{P}(|S_n(\epsilon) - S_n| > \delta), \end{aligned} \quad (3)$$

and

$$\begin{aligned} \mathbb{P}(S_n > x) &= \mathbb{P}(S_n > x, |S_n(\epsilon) - S_n| \leq \delta) + \mathbb{P}(S_n > x, |S_n(\epsilon) - S_n| > \delta) \\ &\leq \mathbb{P}(S_n > x, |S_n(\epsilon) - S_n| \leq \delta) + \mathbb{P}(|S_n(\epsilon) - S_n| > \delta) \\ &\leq \mathbb{P}(S_n(\epsilon) > x - \delta) + \mathbb{P}(|S_n(\epsilon) - S_n| > \delta). \end{aligned} \quad (4)$$

From (3) it follows, as $n \rightarrow \infty$, followed by $\epsilon \rightarrow 0_+$, that

$$\begin{aligned} & \lim_{n \rightarrow \infty} \mathbf{P}(S_n \leq x) \\ & \leq \lim_{\epsilon \rightarrow 0_+} \lim_{n \rightarrow \infty} \mathbf{P}(S_n(\epsilon) \leq x + \delta) + \lim_{\epsilon \rightarrow 0_+} \lim_{n \rightarrow \infty} \mathbf{P}(|S_n(\epsilon) - S_n| > \delta) \\ & = \lim_{\epsilon \rightarrow 0_+} \lim_{n \rightarrow \infty} \mathbf{P}(S_n(\epsilon) \leq x + \delta) = \lim_{\epsilon \rightarrow 0_+} \mathbf{P}(S(\epsilon) \leq x + \delta) = \mathbf{P}(S \leq x + \delta), \end{aligned}$$

for any $\delta > 0$, which implies

$$\lim_{n \rightarrow \infty} \mathbf{P}(S_n \leq x) \leq \mathbf{P}(S \leq x). \quad (5)$$

Analogously, from (4) we get

$$\begin{aligned} & \lim_{n \rightarrow \infty} \mathbf{P}(S_n > x) \\ & \leq \lim_{\epsilon \rightarrow 0_+} \lim_{n \rightarrow \infty} \mathbf{P}(S_n(\epsilon) > x - \delta) + \lim_{\epsilon \rightarrow 0_+} \lim_{n \rightarrow \infty} \mathbf{P}(|S_n(\epsilon) - S_n| > \delta) \\ & = \lim_{\epsilon \rightarrow 0_+} \lim_{n \rightarrow \infty} \mathbf{P}(S_n(\epsilon) > x - \delta) = \lim_{\epsilon \rightarrow 0_+} \mathbf{P}(S(\epsilon) > x - \delta) = \mathbf{P}(S > x - \delta), \end{aligned}$$

for any $\delta > 0$, which implies $\lim_{n \rightarrow \infty} \mathbf{P}(S_n > x) \leq \mathbf{P}(S \geq x)$, and thus

$$\lim_{n \rightarrow \infty} \mathbf{P}(S_n \leq x) \geq \mathbf{P}(S < x). \quad (6)$$

Continuity of the distribution function of S at x and the relations (5) and (6) imply $\lim_{n \rightarrow \infty} \mathbf{P}(S_n \leq x) = \mathbf{P}(S \leq x)$, which concludes the proof. \square

Lemma 4. *Let $\{\mathbf{R}(t), t \in (0, 1)\}$ be a p -dimensional continuous stochastic process, $p \in \mathbb{N}$, and let $h : \mathbb{R}^p \rightarrow (0, 1)$ be a function. Then, it holds*

$$\lim_{\epsilon \rightarrow 0_+} \sup_{t \in [\epsilon, 1-\epsilon]} h(\mathbf{R}(t)) = \sup_{t \in (0, 1)} h(\mathbf{R}(t)), \quad \text{almost surely.}$$

Proof. We denote

$$M(\epsilon) = \sup_{t \in [\epsilon, 1-\epsilon]} h(\mathbf{R}(t)), \quad M = \sup_{t \in (0, 1)} h(\mathbf{R}(t)).$$

First, we will show that $\lim_{\epsilon \rightarrow 0_+} M(\epsilon) \leq M$, almost surely. In order to do that, we realize that a supremum of a function over a set A , such that a set B is its subset, i.e. $B \subset A$, is always at least as big as a supremum of the same function over the set B . This property yields that $M \geq M(\epsilon_1) \geq M(\epsilon_2)$ for any $0 < \epsilon_1 < \epsilon_2 < 1$, since $[\epsilon, 1 - \epsilon] \subset (0, 1)$ for any $\epsilon \in (0, 1)$, which implies $M \geq M(\epsilon)$ for any $\epsilon \in (0, 1)$, and from that the relationship $\lim_{\epsilon \rightarrow 0_+} M(\epsilon) \leq M$, almost surely, follows.

Next, we show $\lim_{\epsilon \rightarrow 0_+} M(\epsilon) \geq M$, almost surely.

CASE 1: Assume there exists $t_0 \in (0, 1)$ such that $h(\mathbf{R}(t_0)) = M$. Then, there exists $\epsilon_0 \in (0, 1)$ such that $t_0 \in [\epsilon_0, 1 - \epsilon_0]$. Then $t(\epsilon) = t_0$ for all $0 <$

$\epsilon < \epsilon_0$, which implies $M(\epsilon) \geq M$ for $0 < \epsilon < \epsilon_0$ and thus $\lim_{\epsilon \rightarrow 0^+} M(\epsilon) \geq M$, almost surely.

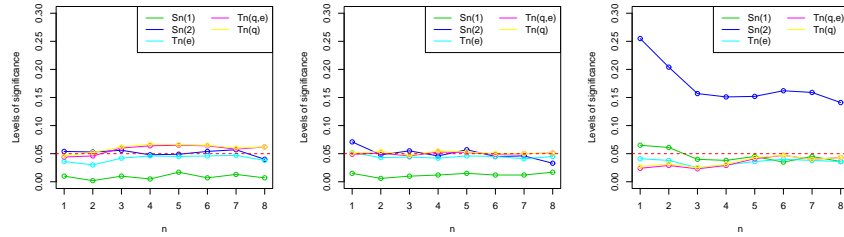
CASE 2: If there does not exist $t_0 \in (0, 1)$ at which $h(\mathbf{R}(t))$ attains its maximum on $(0, 1)$, then for all $t \in (0, 1)$ it holds $h(\mathbf{R}(t)) < M$. However, there exist a sequence $\{t_n\}_{n \in \mathbb{N}} \subset (0, 1)$ such that $\lim_{n \rightarrow \infty} h(\mathbf{R}(t_n)) = M$. There also exists a sequence $\{\epsilon_n\}$ such that $t_n \in [\epsilon_n, 1 - \epsilon_n]$ for each $n \in \mathbb{N}$ and $\epsilon_n \rightarrow 0_+$ as $n \rightarrow \infty$. Since it holds $h(\mathbf{R}(t_n)) \leq h(\mathbf{R}(t(\epsilon_n)))$ for all $n \in \mathbb{N}$ from the definition of $t(\epsilon_n)$, it also holds $\lim_{\epsilon \rightarrow 0^+} M(\epsilon) = \lim_{n \rightarrow \infty} h(\mathbf{R}(t(\epsilon_n))) \geq \lim_{n \rightarrow \infty} h(\mathbf{R}(t_n)) = M$, almost surely. \square

3 Simulations

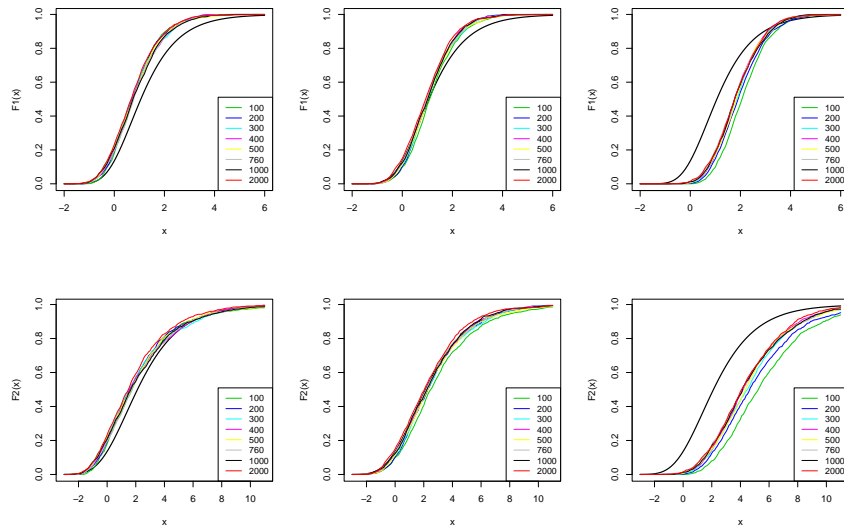
In [3] we simulated large number of AR processes of 3 different orders of auto-regression $p_1 = 1$, $p_2 = 3$ and $p_3 = 5$. For each order we simulated 1000 repetitions of AR processes of lengths $n = 100, 200, 300, 400, 500, 760, 1000, 2000$. We performed these simulations under three different distributions of the innovations $\{\varepsilon_t\}$: the standard normal distribution, t_6 -distribution and the uniform distribution on $[-1, 1]$. For the lack of available space we present here only the normal case. We should like to mention that the results were very similar for all three distributions. We also simulated processes with fixed and randomly chosen coefficients. We only present the case of fixed coefficients, which were fixed at $\varphi_1 = 0.5$ for $p = 1$, and $\varphi_1 = 0.5, \varphi_2 = -0.4, \varphi_3 = 0.1$ for $p = 3$, and $\varphi_1 = 1, \varphi_2 = -0.1, \varphi_3 = 0.3, \varphi_4 = -0.4, \varphi_5 = 0.1$ for $p = 5$. These coefficients were used to generate AR with no change point and the before change portion of AR processes with change point. On the other hand, the the coefficients used to simulate the after change point portion of the data were fixed at the values $\psi_1 = 0.8$ for $p = 1$, and $\psi_1 = 1, \psi_2 = -1, \psi_3 = 0.1$ for $p = 3$, and $\psi_1 = -0.5, \psi_2 = 0.5, \psi_3 = 0.3, \psi_4 = -0.4, \psi_5 = 0.1$ for $p = 5$. We fixed the location of the change point at values corresponding to the choice of $\tau = \frac{1}{2}$, i.e. $k^* = \lfloor \frac{n}{2} \rfloor$.

3.1 Behavior of statistics under null hypothesis

From the following graphs it is apparent that while most of the statistics behaved in accordance with theory, the attained levels of significance were not always near the chosen value of $\alpha = 5\%$ even for large sample sizes. For example, a lower level of significance was attained by the test statistic $S_n(1) = A_n(\log n)T_n^{1/2} - D_p(\log n)$ for the two smaller choices of order $p = 1, 3$, with the results for $p = 3$ being slightly more favorable. On the other hand, already for the case of $p = 5$ the attained level of significance for a test based on $S_n(1)$ is quite close to the asymptotic level of 5%. However, in the case of $S_n(2) = (A_n^2(\log n)T_n - D_p^2(\log n))/D_p(\log n)$ for $p = 5$ the convergence is particularly slow and the attained level of significance even for $n = 2000$ is much higher than 5%. We also conclude that the tests based on $T_n(\epsilon)$, $T_n(q_0, \epsilon)$ and $T_n(q_0)$ have the attained level quite close to the theoretical value of 5%.

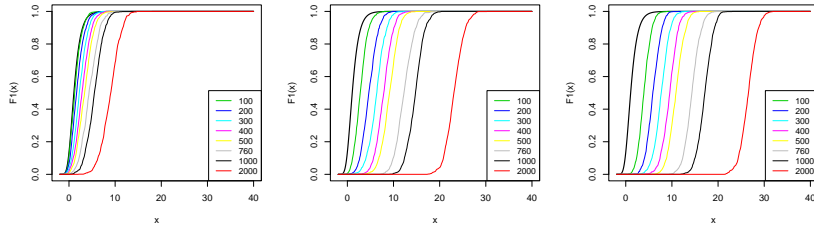


Next, we illustrate the behavior of $S_n(1)$ and $S_n(2)$ by the convergence of their distribution functions. The upper row corresponds to $S_n(1)$, the lower row corresponds to $S_n(2)$, with orders increasing from left to right ($p = 1, 3, 5$).

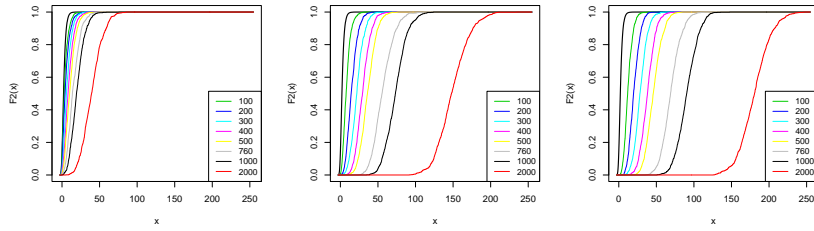


3.2 Detection of change point under alternatives

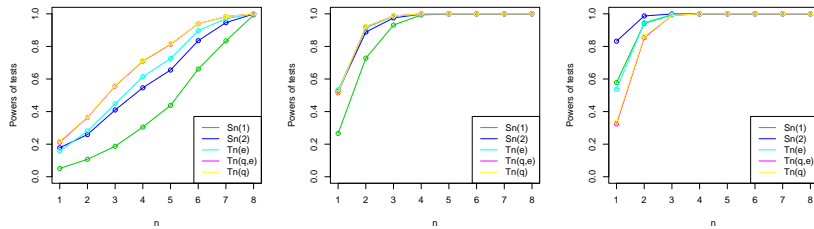
The tests based on the statistics T_n , $T_n(\epsilon)$, $T_n(q_0, \epsilon)$ and $T_n(q)$ with $q = (t(1-t))^\beta$, $t \in (0, 1)$, $\beta \in [0, \frac{1}{2}]$ are consistent (see [2] or [3]). To illustrate this we again simulated $N = 1000$ realizations of AR processes with a change of autoregression coefficients at time $k^* = \lfloor \frac{n}{2} \rfloor$. We present here the results for $S_n(1)$ and $S_n(2)$. The distribution of $S_n(1)$ behaved as shown in the following three graphs for increasing orders $p = 1, 3, 5$ from left to right.



An analogous illustration of the behavior $S_n(2)$ for $p = 1, 3, 5$ increasing from left to right is shown in the next three graphs.



In our graphs we can observe an apparent shifting of the empirical distribution functions of $S_n(1)$ and $S_n(2)$ towards infinity for all orders $p = 1, 3, 5$ as n increases. Note, that the convergence seems to be faster for higher orders of autoregression. Based on the results in [3] we can also conclude that the convergence of $S_n(1)$ and $S_n(2)$ to infinity appears faster for $\tau = 0.5$ than for either of other two choices $\tau = 0.25$ and $\tau = 0.75$. Finally, we illustrate consistency of all of the presented test statistics $S_n(1)$, $S_n(2)$, $T_n(\epsilon)$, $T_n(q, \epsilon)$ and $T_n(q)$ in terms of empirical power for $p = 1, p = 3$ and $p = 5$ with $\tau = 0.5$ and $\epsilon = 0.1$. The simulated results seem to be quite satisfactory.



References

[1] Davis, R., Huang, D., and Yao, Y.C. (1995). Testing for a change in the parameter values and order of an autoregressive model. *Annals of Statistics*, **23**(11):282–304.

- [2] Hušková, M., Prášková, Z., and Steinebach, J. (2007). On the detection of changes in autoregressive time series, i. asymptotics. *Journal of Statistical Planning and Inference*, **137**(4):1243–1259.
- [3] Pečánka, J. (2008). Change Detection in Autoregressive Time Series *Master Thesis*, VU University Amsterdam.

Acknowledgement: This work is based on my master thesis defended at VU University Amsterdam in 2008 under the supervision of Aad van der Vaart. The participation to the conference was supported by ČSOB and Median.

Address: J. Pečánka, MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8 & VU University Amsterdam, De Boelelaan 1081a, 1081 HV, Amsterdam, The Netherlands

E-mail: j.pecanka@few.vu.nl

ZKUŠENOSTI SE STATISTICKÝM MODELOVÁNÍM SPOTŘEBY ZEMNÍHO PLYNU

Emil Pelikán, Marek Brabec, Marek Malý,
Ondřej Konár, Jiří Vondráček

Klíčová slova: Typové diagramy dodávky, nelineární efekty, GAM.

Abstrakt: Úlohou statistického modelování spotřeby zemního plynu se kolektiv pracovníků Ústavu informatiky AV ČR začal intenzivně zabývat v souvislosti s návrhem modelu pro předpověď celkové spotřeby plynu v určitém regionu. Později byl ve spolupráci se Západočeskou plynárenskou, a.s. vyvinut model pro odhad tzv. nevyfakturované dodávky zemního plynu. Práce na tomto tématu se prohloubily s účastí v programu „Informační společnost“ (za podpory grantu GA AV ČR č. 1ET400300513) a na projektu „Tvorba typových diagramů dodávky zemního plynu“ (řešeném ve spolupráci s partnery v České republice a na Slovensku).

V příspěvku jsou shrnuty zkušenosti řešitelů získané při modelování, a to zejména v průběhu přípravy dat a návrhu, implementace a vlastního testování statistických modelů spotřeby zemního plynu. Příspěvek je rozdělen do čtyř hlavních částí, popisujících obecný úvod do problematiky, stav řešení ve světě, příklady použitých statistických modelů, možnosti jejich dalšího rozvoje a faktory ovlivňující kvalitu modelů. V závěru jsou shrnuty význam a další perspektivy statistického modelování spotřeby zemního plynu.

Abstract: This paper summarizes authors' experiences in mathematical modeling of natural gas consumption. Some issues related to the development, testing and implementation of several mathematical models are described. The paper is divided into four parts dealing with general questions of natural gas consumption modeling, examples of some statistical techniques, possibilities of their further development and factors affecting the quality of estimates.

1 Úvod

První projekty spojené s tematikou modelování spotřeby zemního plynu byly řešeny v Ústavu informatiky AV ČR od r. 1999 v souvislosti s potřebou předpovědi spotřeby zemního plynu daného regionu na krátkodobý, střednědobý i dlouhodobý predikční horizont. Některé z navržených modelů byly implementovány do informačních systémů plynárenských distribučních společností v ČR i na Slovensku, kde byly nebo dodnes jsou používány v rutinním provozu. Ze statistického pohledu byly implementované modely založeny zejména na nelineární regresi a jejich modifikacích. Cenné bylo získání zkušeností se zpracováním „plynárenských“ dat a seznámení se s problematikou nasazení matematických modelů do procesu rozhodování a řízení v reálné praxi. Některé nově navržené postupy a metody pak byly publikovány např.

v [12, 29, 32, 33]. Práce na tématu modelování spotřeby se intenzivněji rozebíhly v souvislosti s nárůstem požadavků na kvalitu odhadu spotřeby zemního plynu zákazníků s malou a střední spotřebou (tj. zákazníků, jejichž roční spotřeba se pohybuje pod hranicí 420000 m^3) a zejména pak s udělením grantové podpory GA AV ČR v rámci programu „Informační společnost“ č. 1ET400300513 na léta 2005–2009 a dále s účastí autorů na projektu „Tvorba typových denních diagramů dodávky zemního plynu“, který je řešen ve spolupráci s partnery v České republice a na Slovensku.

Motivací pro modelování spotřeby zákazníků s malou a střední spotřebou je fakt, že tento typ zákazníků je měřen (odečítán) pouze jednou za poměrně dlouhé období, obvykle jeden rok. Jsou tak vlastně známy jen stavy plynoměru udávající kumulativní spotřebu ve zhruba ročních intervalech. Přitom informace o chování zákazníků mezi dvěma odečty je velmi důležitá pro potřeby bilancování, stanovení dodávky nevyfakturované složky, marketingu, plánování kapacit apod. [11]. Zákazníci jsou odečítáni cyklicky v různých časových intervalech v různých obdobích kalendářního roku. Každý měsíc je odečtena přibližně $\frac{1}{12}$ z celého zákaznického kmene. Podíl spotřeby tohoto typu zákazníků na celkové spotřebě je výrazný, v ČR přesahuje hodnotu 40 % [12]. Zákazníci s velkou spotřebou jsou měřeni každodenně v hodinovém rozlišení. Zavedení obdobného přístupu u zákazníků s malou a střední spotřebou je však ekonomicky nereálné, neboť počty těchto zákazníků se pohybují v milionech. Osazení takovýchto odběrných míst měřícími přístroji s pamětí anebo s možností dálkového odečtu by při ceně přístroje v řádu desetitisíců Kč představovalo v rámci celé ČR náklady řádově v desítkách miliard korun. Z tohoto důvodu má matematické modelování chování zákazníků veliký praktický význam. Protože modelování individuálního chování by bylo velmi obtížné, přistupuje se k statistickému modelování chování určité skupiny zákazníků, u nichž lze předpokládat, že spotřebovávají zemní plyn k podobným účelům (např. k otopu, ohřevu teplé vody, vaření).

Další text je uspořádán následovně. V kapitole 2 jsou prezentovány některé literární poznatky a odkazy na práce jiných autorů zabývajících se podobnou problematikou. Jádrem práce je kapitola 3, v níž jsou podrobněji popsány projekty, na jejichž řešení se autoři podíleli. Jde o

- projekt ELVÍRA zaměřený na predikční modely pro předpověď celkové spotřeby plynu v určitém regionu (oddíl 3.1),
- projekt GAMMA zaměřený na odhad objemu odebraného, ale dosud nevyfakturovaného plynu (oddíl 3.2) a
- projekt TDD zaměřený na konstrukci tzv. typových diagramů dodávky zemního plynu (oddíl 3.3).

Detailně budou uvedeny především zkušenosti s modelem založeným na typových diagramech spotřeby zemního plynu [7]. V následujících oddílech 3.4 a 3.5 jsou popsány některé možné směry dalšího rozvoje prezentovaných modelů a otázky posuzování jejich výkonnosti. V kapitole 4 jsou stručně v širším kontextu naznačeny další perspektivy statistického modelování spotřeby zemního plynu.

2 Modelování spotřeby v literatuře

V archívech se nacházejí 50 i více let staré články zabývající se z různých pohledů a v různé míře detailu modelováním spotřeby plynu či elektřiny. Protože lze obě tyto komodity poměrně obtížně skladovat, je co nejpřesnější znalost denního i hodinového průběhu poptávky pro distributory klíčová. Proto se mnoho textů věnuje úloze předpovědi spotřeby energie. Deregulace trhu s plynem a elektřinou přinesla požadavek rozpočtu určitého celku spotřeby mezi jednotlivé obchodní subjekty. Pro rozpočet spotřeby v rámci časového úseku byla dříve používána tzv. otopová křivka, což je 12 koeficientů vymezujících pro každý měsíc jeho procentuální podíl na roční spotřebě [49]. Formalizovat lze tuto úlohu vytvořením tzv. typových diagramů dodávky (TDD). Na nich je založen náhradní postup pro stanovení denního či hodinového odběru skupiny konečných zákazníků, kteří nejsou průběhově měřeni, viz např. [17].

Podle míry detailu lze rozlišit úlohy na modelování celkové spotřeby (buďto celostátní či za distribuční společnost), modelování spotřeby menších celků (určitých skupin zákazníků) a modelování individuální spotřeby. Z pohledu časového rozlišení jde zpravidla o modelování spotřeby za měsíc, den či hodinu. Většina dostupné literatury se týká modelování celku. Hlavní modely představené níže v oddílech 3.2 a 3.3 jsou konstruovány pro skupiny zákazníků. Je potřeba, aby skupiny byly relativně homogenní průběhem spotřeby plynu během roku a přitom navzájem zřetelně odlišné. Počet skupin by měl být relativně malý (v jednotkách) a přiřazení zákazníků do skupin musí být jasné, jednoznačné a v praxi proveditelné. Část publikovaných prací se věnuje hlavně vytvoření relevantních skupin a návazné modelování má pak jen velmi jednoduchou podobu, např. prezentuje empirické průměry. Modelováním na individuální úrovni jsme se zabývali v textu [6].

Modelování průběhu spotřeby může být založeno na různých typech podkladových dat, což zároveň určuje nákladnost a přesnost postupu [1]. Nejpřesnější, ale i nejnákladnější je modelování založené na průběhových měřeních vzorku zákazníků. Na dalším místě je matematické modelování využívající korelace spotřeby a externích veličin, typicky veličin charakterizujících počasí. Jde často o regresní modely, které mohou být využity v kombinaci s průběhovými měřeními. Další je metoda zástupného dne, při níž je v historických datech nalezen den, jehož charakteristiky co nejvíce odpovídají odhadovanému dni. Jeho spotřeba poslouží jako odhad spotřeby v aktuálním dni. Jiné, podstatně méně přesné postupy představují tzv. statické profily, které jsou pro každou kombinaci období a typu dne pevně dány bez ohledu na konkrétní podmínky (počasí), či kalendářní rotace, kdy se v historii vybírají odpovídající dny dle kalendářního data.

Záhy byla rozpoznána zásadní závislost spotřeby na počasí, které bývá charakterizováno především aktuální teplotou a teplotním normálem a dále např. silou a směrem větru, vlhkostí vzduchu. Již v článku [5] z roku 1965 je regresně modelována závislost na teplotě. Jako další důležité určující prediktory se v literatuře ukazují jevy spojené s kalendářem, tedy poloha dne

v týdnu (resp. rozlišení na pracovní a víkendový den), svátky, délka slunečního svitu, a dále indikátor sezóny, teplotní pásmo (resp. nadmořská výška), charakteristika zákazníka a typu jeho spotřeby, charakteristika domu, případně i cena energie. Někdy se používá jedna kompozitní proměnná, která může být např. funkcí zejména aktuální teploty, rychlosti větru, efektivní teploty (půl dnešní + půl včerejší teploty) a sezónního normálu efektivní teploty.

Regresní závislost spotřeby Y_t na teplotě T_t ve dni t bývá modelována lineárně, polynomiálně a nejčastěji nelineárně, např. sigmoidální křivkou

$$Y_t = \frac{\alpha}{1 + \left(\frac{\beta}{T_t - 40}\right)^\gamma} + \delta \quad (1)$$

s parametry $\alpha, \beta, \gamma, \delta$ [17].

V článku [15] je popsán regresní model použitý pro předpověď spotřeby ve společnosti British Gas NW. Předpověď poptávky Y_{t+1} na následující den je modelována na základě spotřeby Y_t aktuálního dne ve tvaru

$$Y_{t+1} = (Y_t - A_t)(1 + (T_{t+1}^{\text{eff}} - T_t^{\text{eff}})\beta + (W_{t+1} - W_t)\gamma) + A_{t+1}, \quad (2)$$

kde W_t a W_{t+1} představují průměrnou rychlost větru v aktuální den a její předpověď pro den následující, A_t je korekční člen postihující vliv dne v týdnu (zejména odstávky průmyslu v pátek, sobotu a neděli) a β a γ jsou koeficienty charakterizující citlivost vůči teplotě a větru. Efektivní teplota je stanovena vztahem $T_t^{\text{eff}} = \alpha T_t + \alpha(1 - \alpha)T_{t-1} + \alpha(1 - \alpha)^2 T_{t-2} + \dots$ jako exponenciálně vážený průměr teplot předchozích dní s doporučenou volbou $\alpha = 0,5$, T_{t+1}^{eff} je předpověď na následující den.

Do modelu byla ještě začleněna tzv. nehomogenní teplotní citlivost v podobě

$$\beta = \begin{cases} \beta_1 & \text{pokud } T_t^{\text{eff}} \geq T^* \\ \beta_2 & \text{pokud } T_t^{\text{eff}} \leq T^* \end{cases},$$

která zohledňuje proměnlivou citlivost zákazníků vůči teplotě v závislosti na aktuální průměrné teplotě a fakt, že při stejné teplotě bývá spotřeba nižší v létě než v zimě. Tento přístup lze nalézt ve více publikacích. T^* je vhodně zvolená teplota zpravidla odpovídající teplotě v přechodových měsících jara a podzimu, nejčastěji v rozmezí 14–18°C. Podobná koncepce potlačující zcela závislost na teplotě v teplých dnech je často označována jako „heating degree day (HDD)“ [18]. Např. v práci [16], modelující spotřebu plynu v Palermu, vstupuje do modelů veličina $HDD_t = \max(18 - T_t; 0)$. Pro ilustraci uvádíme ukázkou dvou modelů z této práce. Autoři uvažovali kromě vlivu teploty rovněž vlhkost vzduchu (H_t pro den t), atmosférický tlak (P_t) a srážky (R_t) a dále indikátor svátku (F_t), dne v týdnu ($W_{it} = 1$, pokud t patří i -tému dni v týdnu, jinak $W_{it} = 0$) a měsíce ($M_{jt} = 1$, pokud t patří j -tému měsíci v roce), ε_t je náhodná chybová složka. Při výpočtech na základě modelu

$$\begin{aligned}
Y_{t+1} = & c + \alpha_0 HDD_t + \alpha_1 HDD_{t-1} + \alpha_2 HDD_{t-2} \\
& + \beta_0 H_t + \beta_1 H_{t-1} + \beta_2 H_{t-2} \\
& + \gamma_0 R_t + \gamma_1 R_{t-1} + \gamma_2 R_{t-2} \\
& + \delta_0 P_t + \delta_1 P_{t-1} + \delta_2 P_{t-2} \\
& + F_t + \sum_{i=2}^7 \vartheta_i W_{it} + \varepsilon_t
\end{aligned} \tag{3}$$

nevykazovaly charakteristiky počasí s výjimkou teploty zásadní vliv na poptávku po plynu. Pro zohlednění autokorelace náhodné složky modelu autoři začlenili do modelu autoregresní strukturu druhého řádu v chybové složce, takže finální model měl podobu

$$Y_{t+1} = \omega + \alpha_0 HDD_t + \alpha_1 HDD_{t-1} + \alpha_2 HDD_{t-2} + F_t + W_t + \varepsilon_t, \tag{4}$$

kde

$$(1 - \varphi_1 L - \varphi_2 L^2) \varepsilon_t = \zeta_t, \tag{5}$$

L je operátor zpoždění, $\varphi_1, \varphi_2, \omega$ jsou parametry a ε_t, ζ_t jsou nekorelované chybové složky. Místo indikátoru dne v týdnu zde bylo použito rozlišení na pracovní dny a víkend ($W_t = 1$, pokud den t je víkendový).

Další, převážně nelineární, regresní přístupy k modelování spotřeby plynu na různém stupni komplexity jsou popsány např. v [46, 2, 17, 18, 35, 36]. Regresní a ekonometrické modely mohou obsahovat ARMA členy pro zachycení efektů v čase se měnících neměřitelných veličin [48]. Při modelování se dále používají především časové řady a speciálně Boxova-Jenkinsova metodologie ARIMA [27, 42], Bayesovské modely [3] a neuronové sítě [44, 21].

Podobné úlohy modelování a predikce jsou řešeny i v rámci elektroenergetiky. V literatuře jsou zastoupeny více než úlohy týkající se spotřeby plynu a byly jim věnovány i monografie či jejich části [8, 14]. Regresní model pro spotřebu elektřiny je popsán např. v článcích [40, 43], dále lze jmenovat semiparametrický přístup s využitím časových řad [26], Bayesovský model [10], časově proměnlivé periodické spliny [19], Holtovo-Wintersovo exponenciální vyhlazování [45], či regresní stromy [9]. Dynamické přístupy založené na časových řadách (např. [30]) jsou často modelovány na velkých celcích a navíc mají k dispozici pro praktické použití na celém zákaznickém kmene data v jemnějším rozlišení, než je v našich podmínkách dosažitelné. Na stránkách Operátora trhu s elektřinou, a.s. (<http://www.ote-cr.cz/>) lze nalézt informace o použití TDD v české elektroenergetice.

Problémy diskutované v článcích v souvislosti s aplikací jednotlivých metod zahrnují mj. nepřesnosti modelování, vliv výběrové chyby, postupné proměňování profilu spotřeby (např. dlouhodobý pomalý trend ke snižování spotřeby). Dále jde o chyby v přiřazování zákazníků do skupin, chyby na

vstupu do systému distribuce (např. v důsledku ztrát), nepřesnosti v algoritmizaci postupu či existenci zákazníků, nespádajících do populací reprezentovaných stávajícími skupinami.

3 Modelování spotřeby zemního plynu v ČR a SR

V této kapitole představíme tři základní projekty, kterými jsme se při modelování spotřeby zemního plynu zabývali.

3.1 Projekt ELVÍRA

V projektu, který byl pod názvem ELVÍRA řešen v letech 1999 – 2007, se řešitelská skupina podílela na tvorbě predikčních modelů pro předpověď celkové spotřeby zemního plynu v daném regionu na krátkodobý, střednědobý i dlouhodobý horizont.

V krátkodobém horizontu šlo o úlohu předpovědi denních (resp. hodinových) hodnot očekávané spotřeby zemního plynu na 7 dní (resp. na 168 hodin) dopředu. Jako vstupní údaje sloužily historické hodnoty spotřeby, teploty a dále předpovědi teploty na 7 dní dopředu. Předpověď na následující den byla dále korigována s využitím předpovězených hodnot počtu hodin slunečního svitu a průměrné rychlosti větru. V predikčním systému, který byl realizován na základě navržených modelů, byly implementovány i algoritmy pro výpočet změny akumulace zemního plynu v potrubí a algoritmy pro detekci a korekci výpadků měření a odlehlých hodnot.

V úloze předpovědi spotřeby na střednědobý horizont (cca na jeden rok dopředu) se predikovaly měsíční spotřeby pro různé scénáře vývoje počasí (teplý, studený resp. normální rok).

Pro dlouhodobou předpověď na období několika let do modelu navíc vstupovaly údaje o skutečném a předpokládaném vývoji ekonomiky a o cenách zemního plynu i o cenách alternativních zdrojů energie (elektřiny, uhlí, dřeva apod). Bližší podrobnosti projektu ELVÍRA popisují publikace [31, 32, 34].

Navržený predikční systém byl implementován a byl (anebo stále je) úspěšně používán v celkem 5 regionálních distribučních plynárenských společnostech v ČR a v SPP a.s. na Slovensku. V současné době probíhá v ČR restrukturalizace (slučování) jednotlivých distribučních společností a předpokládá se vytvoření jediného centrálního dispečinku s novým predikčním systémem.

3.2 Projekt GAMMA

Ve spolupráci se Západočeskou plynárenskou, a.s. je vyvíjen model pro odhad spotřeby zemního plynu, nazvaný GAMMA. O tomto modelu jsme již referovali na ROBUSTu 2006 [22]. Jeho základním účelem je odhad objemu nevyfakturovaného plynu, tedy plynu, který již byl zákazníky odebrán, ale ještě nebyl vyfakturován. Tento objem nelze zjistit přímo z důvodu výše zmíněných

zhruba ročních rozestupů mezi odečty spotřeby zákazníka. Jde o nelineární regresní model pro spotřebu Y_{ikt} zákazníka i ze třídy k ve dni t ve tvaru

$$Y_{ikt} = \mu_{ik}(\Psi_t e^{-\gamma_k f(T_t, N_t)} + p_k) + \varepsilon_{ikt}, \quad (6)$$

kde μ_{ik} je efekt konkrétního zákazníka, Ψ_t sezónní složka modelu, γ_k parametr určující míru závislosti spotřeby ve skupině zákazníků k na teplotě, p_k stálá složka spotřeby zákazníků ze skupiny k a T_t , resp. N_t představují skutečnou, resp. normálovou teplotu ve dni t a ε_{ikt} je náhodná složka.

Vzhledem k nestabilnímu chování modelu (6) v letním období (neúměrné kolísání modelové spotřeby) bylo přistoupeno k úpravě modelu tzv. „přepínačem“. Spotřeba je pak v zimním období (definovaném průměrnou teplotou za poslední tři dny) modelována modelem (6) a v letním období (doplňk k zimnímu) konstantou ($Y_{ikt} = \mu_{ik}q_k + \varepsilon_{ikt}$).

Parametry modelu jsou z provozních důvodů odhadovány dvoufázově. V pravidelných intervalech (v současné době půl roku) jsou předodhadovány parametry p_k , q_k , γ_k a parametr μ_{ik} je (mj. vzhledem k proměnlivému složení zákaznického kmene) odhadován znovu při každém běhu modelu.

Podrobnější informace k modelu GAMMA lze nalézt například v publikacích [47, 22, 11, 4].

3.3 Projekt TDD

V dalších oddílech podrobněji představíme náš přístup k řešení projektu konstrukce typových diagramů dodávky (TDD).

3.3.1 Charakteristika dat Tvorba typových diagramů dodávky zemního plynu (ale i elektřiny) má dva aspekty dané zadáním a požadavky budoucích uživatelů. Prvním z nich je konstrukce hladké TDD křivky jako „typického“ chodu spotřeb za „standardních“ podmínek, druhým pak výpočet korekcí na odchylky od standardu/normálu. Důležitým předpokladem je, že určité skupiny zákazníků spotřebovávají plyn relativně podobným způsobem ve smyslu podobné dynamiky spotřeby v průběhu roku, podobné odpovědi na změny v teplotě apod. V elektroenergetice jsou takové skupiny přirozeně dány příslušností k tarifu [25], avšak v plynárenství bohužel žádné přímočaré přiřazení neexistuje. Je proto třeba na základě analýzy dostupných dat takové skupiny nejprve najít a poté pro každou z nich odhadnout jak typickou TDD křivku, tak parametry pro korekce na odchylky od normálu. Model je založen na datech z průběhového měření spotřeb náhodně vybraných zákazníků a jejich segmentaci.

Segmentace zákazníků Je jasné, že segmentaci zákazníků dle způsobů spotřeby lze provést velmi rozdílnými způsoby. K tomu, aby výsledná klasifikace byla prakticky použitelná v reálných podmínkách, je nutné se při vytváření kategorií omezit na znaky, jejichž hodnoty jsou běžně dostupné

pro všechny zákazníky v provozních databázích. Jejich nabídka není zdaleka ideální a nelze ji navíc snadno měnit bez velkých dodatečných nákladů. Zákazníci jsou tedy pro účely modelování spotřeby tříděni dle příslušnosti k

- kategorii zákazníka – domácnost, maloodběratel, střední odběratel,
- charakteru spotřeby – vytápění, ohřev vody, vaření, technologie,
- typu podniku – výrobní prostory, škola, nemocnice, kotelna, atd.

Kombinací jednotlivých kategorií vytvořených na základě tří třídících kritérií je ovšem příliš mnoho, a proto tato segmentace bohužel poskytuje příliš mnoho tříd s malým zastoupením. Navíc se ve zpracovávaných datech vyskytují evidentně chybné zatřídění zákazníci, zejména proto, že skutečný charakter spotřeby neodpovídá způsobu zákazníkem deklarovanému. Pro dosažení vyšší stability výsledků i robustnosti vůči drobným chybám v klasifikaci je nutné sloučení tříd s podobným charakterem spotřeby, zejména s podobným obrazem její dynamiky v průběhu roku. Tento problém byl řešen pomocí shlukové analýzy aplikované na týdenní data průběhově měřených zákazníků (metody k-průměrů, Wardova) [23, 28]. Na základě získaných výsledků pak bylo vytvořeno 8 shluků (segmentů). Jejich struktura je poněkud odlišná pro ČR a SR. V ČR jsou 4 segmenty pro domácnosti (označované „domo“), 4 pro maloodběratele a střední odběratele (označované „moso“). Na Slovensku pak jsou 2 domo segmenty a 6 moso segmentů.

Při konstrukci TDD modelu používáme stratifikaci dle segmentu. V důsledku toho jsou všechny parametry modelu stejně jako jeho hladká část segmentově specifické.

Data Pro statistické modelování lze využít údaje, které jsou k dispozici v informačních systémech jednotlivých distribučních společností. Jde zejména o hodnoty odečtu spotřeby jednotlivých zákazníků za několik posledních let. Data z těchto odečtů odpovídají součtu denních spotřeb daného zákazníka za období přibližně jednoho roku. Umístění intervalu pokrytého odečtem je však u různých zákazníků obecně různé. Dále lze z typu uzavřené smlouvy odvodit informace o očekávaném charakteru spotřeby, kategorii zákazníka a jeho lokalizaci.

Pro návrh statistického modelu a odhad jeho parametrů je zapotřebí získat ještě údaje o chování zákazníků mezi řádnými odečty. Jinými slovy, je nutno získat informaci o sezónní složce spotřeby, jež v agregovaných přibližně ročních součtech takřka není přítomna, ale je potřebná k mnoha odhadům. To je zajištěno pomocí vzorku průběhově měřených zákazníků, jejichž spotřeba je zaznamenávána po hodinách, v takřka spojitým čase. Protože každé prováděné měření je poměrně nákladné, je velikost vzorku průběhově měřených zákazníků velmi limitována. V rámci projektu TDD máme k dispozici cca 1000 průběhově měřených zákazníků v ČR a cca 550 na Slovensku. Formálně vzato, jde o vzorek funkcionálních dat [38]. Požadovaný stupeň

časového rozlišení je však podstatně nižší (hodiny, eventuálně dny), takže modelujeme příslušné součty (integrály) „v čase spojitých“ měření.

Konečný výběr zákazníků do vzorku prováděly plynárenské společnosti, neboť z důvodu ochrany dat nemohly poskytnout seznam všech svých zákazníků, jenž by posloužil jako opora výběru. Statistikům byly k dispozici pouze souhrnné údaje o počtu zákazníků a celkové výši spotřeby pro určité skupiny odběratelů. Údaje byly stratifikovány podle kategorie zákazníka, pásma spotřeby (tarifního pásma), distribučního závodu (jednotlivé závody v ČR zhruba pokrývají území původních krajů) a nadmořské výšky (dvě kategorie), resp. teplotního pásma (tři kategorie). U malých a středních odběratelů byli zákazníci ještě klasifikováni podle typu podniku. Po vyhodnocení dostupných dat byl pro každou relevantní kombinaci třídících znaků stanoven požadovaný počet průběhově měřených zákazníků s cílem dosáhnout co nejlepší reprezentativity vzorku vůči celému zákaznickému kmeni. Při tom se vycházelo hlavně z poměru spotřeby dané podskupiny k celkové spotřebě a z posouzení nehomogenity průběhu spotřeb uvnitř každé kategorie. Analýza ukázala, že variabilita odběru ve skupině je zhruba úměrná spotřebě skupiny. V případech variability vyšší, než odpovídá průměrné spotřebě, byl rozsah vzorku v dané skupině expertně zvýšen, jelikož u zákazníků s velkou spotřebou a velkým rozptylem hodnot je účelné mít četnější zastoupení. Požadavek na zastoupení nadmořské výšky, resp. teplotního pásma byl formulován pouze v podobě doporučeného procentuálního zastoupení jednotlivých kategorií. Dále bylo třeba, aby byli vybráni zákazníci s dlouhodobě stabilním chováním vzhledem k odběru zemního plynu.

Kromě výše zmíněných průběhových dat jsou v projektu k dispozici i agregovaná data, jež jsou sbírána rutinně, a nevyžadují tedy dodatečná speciální měření. Jedná se o denní spotřeby za ČR, resp. SR. Ty nejsou ve skutečnosti změřeny přímo, vznikají jako rozdíl mezi vstupem do distribuční sítě a součtem veličin, jež nejsou pro modelování relevantní (spotřeby zákazníků s velkými odběry měřené denně, odhadované ztráty, apod.). Dále jsou k dispozici obdobná agregovaná data za vybrané menší celky řádově stovek zákazníků (např. uzavřené rozvodné sítě). Stejně jako v případě celků zde platí, že hodnotu relevantní pro modelování lze získat pouze jako rozdíl chybami zatížených veličin. Z tohoto důvodu jsou zmíněné agregované údaje využívány v procesu návrhu modelu jen jako testovací data.

V neposlední řadě jsou pro konstrukci modelu zapotřebí další „exogenní“ vysvětlující proměnné. Jde zejména o teplotu, ale také o indikátory různých jevů spojených s kalendářem (den v týdnu, svátek, změna času ze zimního na letní apod.). Soubor těchto indikátorových proměnných by teoreticky mohl být širší, jenže snaha o větší detail naráží na limitace ze strany dat. Máme k dispozici sledování za několik málo let, v nichž je počet opakování řídce se vyskytujících jevů velmi malý. Na rozdíl od efektu teploty, který lze sledovat v průběhu celé datové řady, nelze jejich efekt tudíž zatím dobře postihnout. Příkladem takového jevu jsou tzv. „bridging days“, tj. pracovní

dnů mezi víkendem a státním svátkem připadajícím na úterý či čtvrtek. Část podniků odstává v příslušné pondělí či pátek výrobu a podniky tak sníží spotřebu plynu, zatímco u domácností oproti běžným pracovním dnům spotřeba vzroste.

3.3.2 Struktura TDD modelu Model TDD je sestaven ze dvou komponent. Nejprve je vytvořen denní submodel a následně hodinový submodel, který denní součty rozpočítává do hodin. Obě komponenty jsou pak sloučeny do výsledného celkového modelu [7].

Spotřeby s_{it} zákazníka i ve dni t jsou transformovány pomocí funkce

$$\log\left(\frac{s_{it} + 0,5}{p_i}\right), \quad (7)$$

kde p_i je průměr dlouhodobé historické spotřeby zákazníka i vztahovaný na den. Symbol \log značí přirozený logaritmus. Konstanta 0,5 se přičítá kvůli možnému výskytu nulových denních spotřeb. Po transformaci probíhá analýza založená na předpokladu (přibližné) normality. Efektivně tedy pracujeme s logaritnicko-normálním rozdělením původních dat. Tento přístup přináší nejen potlačení váhy odlehklých pozorování, tedy robustnost, ale i symetrizaci rozdělení transformovaných hodnot. Standardizací individuální denní spotřebou p_i odstraníme značné (i několikařádkové) rozdíly v průměrné spotřebě mezi jednotlivými zákazníky. Získáme tak srovnatelná relativní data namísto absolutních.

V modelu i v předpisu pro použití TDD používáme pro vyjádření hodinové spotřeby s_{ikth} následující indexy: i pro zákazníka, $i = 1, \dots, n_k$; k pro segment, $k = 1, \dots, K$; t pro den, $t = 1, \dots, 365$ (resp. 366 v přestupném roce) a h pro hodinu, $h \in t$, $h = 1, \dots, 24$ (resp. 23 nebo 25 ve dnech se změnou času).

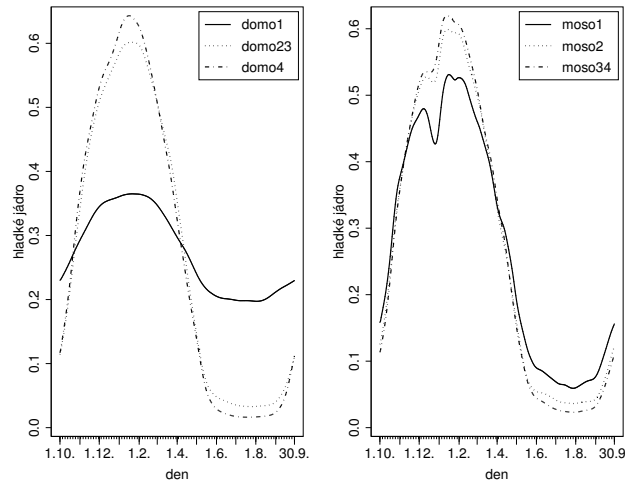
Denní submodel Pro spotřebu $s_{ikt\bullet}$ zákazníka i ze segmentu k ve dni t máme následující model:

$$\log\left(\frac{s_{ikt\bullet} + 0,5}{p_{ik}}\right) = \log(D_{kt}) + \log(C_{kt\bullet}) + \log(M_k) + \eta_{ikt}, \quad (8)$$

kde

D_{kt} je denní korekce, která zahrnuje korekci na den v týdnu, na státem uznaný svátek, speciální efekty Vánoc a Velikonoc, korekci na průměrnou denní teplotu T_t a její odchylky od dlouhodobě typické hodnoty N_t ,

$C_{kt\bullet}$ je hladké jádro denního odhadu spotřeby; odpovídá neparametrickému odhadu roční periodicity spotřeb, korigovaných na teplotu a kalendářové jevy, příklad jeho průběhu pro několik českých segmentů je v obrázku 1,



Obrázek 1: Odhad hladkého jádra $C_{kt\bullet}$ pro několik českých segmentů.

p_{ik} je denní průměr historické spotřeby daného zákazníka, který individualizuje odhad na konkrétního zákazníka,

M_k je korekce na vychýlení při zpětné transformaci odhadu pořízeného na logaritmické škále na škálu původních hodnot (souvisí se skutečností, že $Ef(X) \neq f(EX)$ pro nelineární funkci f a s tím, že pro plynárenské účely je třeba odhadovat střední hodnoty, nikoli mediány),

η_{ikt} je náhodná normálně rozdělená chyba s nulovou střední hodnotou a rozptylem σ^2 .

Po odlogaritmování získáme model

$$s_{ikt\bullet} = p_{ik} D_{kt} C_{kt\bullet} M_k e^{\eta_{ikt}} - 0,5. \quad (9)$$

Rovnice (9) určuje multiplikatívni logaritmicko-normální model, resp. plně aditivní model na logaritmické škále. Logaritmicko-normální model implikuje heteroskedasticitu na původní škále. Kromě toho zde máme vlivem členu p_{ik} rozptyl individuálně-specifický. Individuální efekt p_{ik} („výškové usazení“) je zde separován od časového průběhu („tvaru“). Jinými slovy, model nedovoluje interakci mezi jedincem a tvarem spotřeby. To je ve shodě s chováním podstatné části zákazníků i s tím, že individuální data o tvaru spotřeby obecně nejsou v provozních podmínkách k dispozici.

Pro zjednodušení implementace se na přání zadavatele zanedbává konstanta 0,5 (neboť v daném kontextu představuje velmi malý objem plynu) a pro odhad spotřeby s_{ikt} se tedy v praxi používá model

$$\hat{s}_{ikt\bullet} = p_{ik} D_{kt} C_{kt\bullet} M_k. \quad (10)$$

Jedná se o semiparametrický zobecněný aditivní model (GAM; generalized additive model) s flexibilní hladkou částí (hladké jádro — křivka $C_{kt\bullet}$) a s parametrickou částí (pro korekce). Odhad parametrů lze založit například na „backfitting“ algoritmu [20], zejména pokud použijeme vyhlazení LOESS založené na lokálně vážené regresní funkci [13]. Jako jeden z produktů takové implementace dostáváme odhad sezónní složky (tj. hladkého jádra $C_{kt\bullet}$) očištěné od teplotních a kalendářních vlivů. Členy $C_{kt\bullet}$ jsou normalizovány tak, že udávají procentuální podíly jednotlivých dnů na roční spotřebě za standardních teplotních i kalendářových podmínek. Standardní specifikace (např. v S-plus či R, [37, 39]) přináší nevýhodu v tom, že není jednoduché požadovat periodickou variantu LOESSu. Přitom periodičnost je přitom samozřejmým požadavkem při odhadu sezónní složky. V zásadě zde lze použít různých ad hoc přístupů, např. prokládání periodicky prodloužených dat, dvoukrokového postupu s penalizací i na „švu“ apod. Elegantnější je ale alternativní postup založený např. na lineárním penalizovaném splinu [41]. Při implementaci v podobě lineárního modelu se smíšenými efekty (LME) lze zapsat periodickou podmínku na hladkou část již relativně snadno.

Člen pro denní korekci lze rozepsat do tvaru

$$D_{kt} = e^{kor_den_{kt}} e^{kor_svatek_{kt}} e^{kor_vanoce_{kt}} e^{kor_velikonoce_{kt}} e^{kor_teplota_{kt}}. \quad (11)$$

Vidíme tedy, že denní korekce D_{kt} se v skládá z multiplikativních efektů dvou typů:

- i) neteplotní korekce na kalendářové efekty (na logaritmické škále jde o model analogický vícefaktorové analýze rozptylu, kvůli požadované interpretaci hladkého jádra $C_{kt\bullet}$ volíme „sum“ parametrizaci pro den v týdnu a „baseline“ parametrizaci pro efekty svátků, Vánoc a Velikonoce),
- ii) teplotní korekce (tj. korekce na odchylku teploty od dlouhodobého normálu).

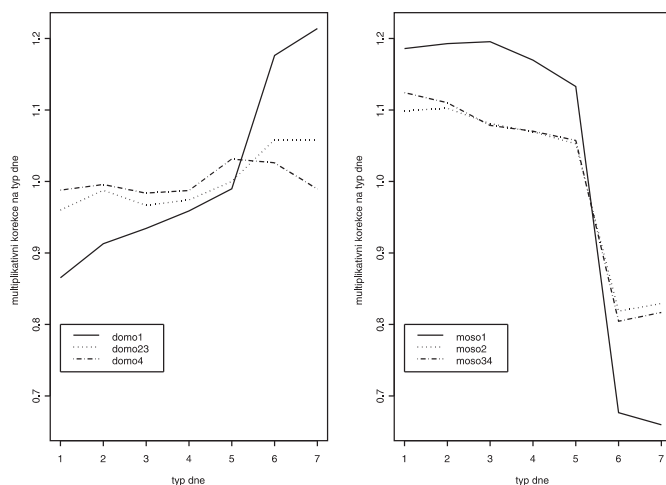
Například na obrázku 2 vidíme multiplikativní korekce $e^{kor_den_{kt}}$ na jednotlivé dny v týdnu, které vznikají transformací efektů jednotlivých typů dne.

Teplotní korekční člen $kor_teplota_{kt}$ počítáme podle vzorce

$$\begin{aligned} kor_teplota_{kt} = & \gamma_{0k} \varphi\left(\frac{T_t - M_k}{U_k}\right) \varphi\left(\frac{T_t - N_t - m_k}{u_k}\right) - \gamma_{0k} \varphi\left(\frac{N_t - M_k}{U_k}\right) \varphi\left(\frac{-m_k}{u_k}\right) \\ & + \gamma_{1k} \varphi\left(\frac{T_{t-1} - M_k}{U_k}\right) \varphi\left(\frac{T_{t-1} - N_{t-1} - m_k}{u_k}\right) - \gamma_{1k} \varphi\left(\frac{N_{t-1} - M_k}{U_k}\right) \varphi\left(\frac{-m_k}{u_k}\right), \end{aligned} \quad (12)$$

kde

$$\varphi(x) = 1 - \frac{1}{1 + e^{-x}} \quad (13)$$



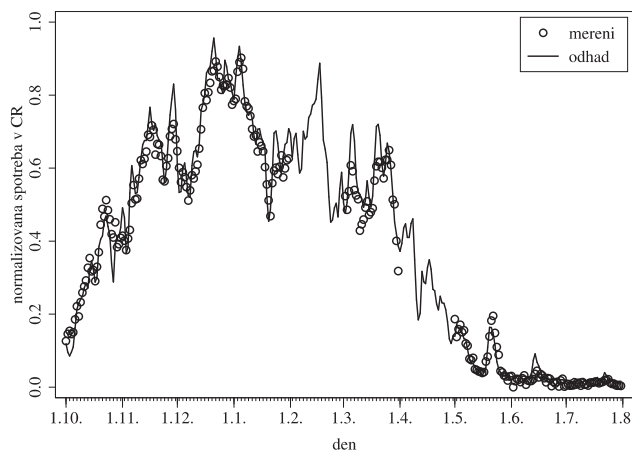
Obrázek 2: Týdenní profil pro české segmenty (pondělí=1, neděle=7).

je doplněk distribuční funkce logistického rozdělení, T_t je průměrná denní teplota za den t , N_t je dlouhodobě typická teplota pro den t a γ_{0k} , γ_{1k} , m_k , u_k , M_k , U_k jsou odhadované parametry. Funkce $\varphi(\cdot)$, která je klesající a má dolní i horní asymptotu, zohledňuje podstatné rysy chování spotřeb vzhledem k teplotě. Teplotní korekce sestává z nelineárně transformované diference mezi skutečnou a dlouhodobě typickou teplotou $T_t - N_t$ a nelineární modulace regresního koeficientu v závislosti na transformované teplotě. Výsledný součin je vyhodnocen v hodnotách současných (pro t) a o jeden den zpožděných (pro $t-1$). Členy se zápornými znaménky ve vzorci (12) jsou přidány jen kvůli požadované pohodlné „normálové“ interpretaci hladkého jádra C_{kt} . Kladné členy představují lineární časově invariantní filtr v nelineárně transformované teplotě. Jde o jednoduchý případ modelu ze třídy „distributed lags“, vyskytující se relativně často např. v ekonometrii [24], zde ale v nelineární podobě. Intuitivně je zřejmé, že zahrnutí časově zpožděných proměnných do modelu je poměrně důležité, neboť zohledňuje zpoždění v reakci spotřebitele na teplotu. To může být způsobeno jak subjektivitou jeho rozhodování, tak např. nezanedbatelnou tepelnou kapacitou vytápěných budov, dnem v týdnu i dalšími faktory. Použitá forma teplotní korekce byla určena po rozsáhlém testování na dostupných datech. Zejména poznamenáme, že přidávání vyšších zpoždění (o 2 a více dní) kvalitu modelu z pohledu přibližných testů poměrem věrohodnosti významně nezlepšovalo.

V praxi se model používá pro odhad spotřeby $S_{ik,\mathcal{T}} = \sum_{t \in \mathcal{T}} s_{ikt}$ (či ještě častěji $S_{\bullet k,\mathcal{T}} = \sum_{i \in k} \sum_{t \in \mathcal{T}} s_{ikt}$) ve dnech z intervalu \mathcal{T} , a to ve dvou typech situací. Buď jde o čas budoucí nebo o minulost. Do budoucna se počítá celkem standardně predikce spotřeby za den \hat{s}_{ikt} či interval $\hat{S}_{ik,\mathcal{T}} = \sum_{t \in \mathcal{T}} \hat{s}_{ikt}$. Pokud je však interval \mathcal{T} celý v minulosti a jde o odhad spotřeby pro $t \in \mathcal{T}$, je dobré

využít toho, že $S_{ik,\mathcal{T}}$ je kompletně známa. Podmíněné odhady využívající znalosti $S_{ik,\mathcal{T}}$ lze snadno získat jako $S_{ik,\mathcal{T}} \cdot (\hat{s}_{ikt} / \sum_{t' \in \mathcal{T}} \hat{s}_{ikt'})$. Model v tomto případě udává jen tvar, nikoli výškové usazení. Toho díky podmiňování znalostí součtové spotřeby není zapotřebí. Podmíněné odhady přicházejí ke slovu při typickém rutinním použití TDD modelu, kdy je požadován „rozpočet“ známé spotřeby (odečítané v cca ročních intervalech na plynoměrech) do kratších intervalů, resp. „rozpočet“ celkové spotřeby mnoha zákazníků (případně celého zákaznického kmene) na skupiny.

Praktické ověřování kvality TDD modelu nebývá vždy úplně standardní. Kromě obvyklé „kvality proložení“ na vzorku průběhově měřených zákazníků, či kvality predikce v čase na témže vzorku se posuzuje i kvalita modelu při použití na zákazníky nezahrnuté do průběhového měření. Takových je samozřejmě naprostá většina. Problémem je, že u nich individuální informace o průběhu spotřeby, proti které by se odhady z modelu daly porovnávat, z principu chybí. Individuálně jsou dostupné jen přibližně roční součty, které nám samy o sobě k detailnímu průběhu spotřeby mnoho neřeknou. Lze ale využít agregovaných dat v denním rozlišení. Ta jsou tvořena součty spotřeb relevantních odběratelů za jednotlivý den v nějakém větším celku, např. v celé národní distribuční síti, jak to ukazuje obrázek 3, nebo v jejích uzavřených částech kontrolovaných jednotlivými rozvodnami. Vzhledem k poznámkám z oddílu 3.3.1 jde i v tomto případě ve skutečnosti jen o odhad, byť prováděný v plynárenském prostředí tak často, že je běžně považován za „pravdu“. Formálně tedy lze pro posouzení kvality modelu porovnat sečtená pozorování $S_{\mathcal{K},\mathcal{T}} = \sum_{k \in \mathcal{K}} \sum_{i \in k} \sum_{t \in \mathcal{T}} s_{ikt}$ se sečtenými odhady



Obrázek 3: Celková spotřeba (normalizovaná z důvodu ochrany dat) za ČR. Chybějící data jsou důsledkem výpadků v měření i jiných problémů.

$$\hat{S}_{\mathcal{K},\mathcal{T}} = \sum_{k \in \mathcal{K}} \sum_{i \in k} p_{ik} \sum_{t \in \mathcal{T}} \frac{\hat{s}_{ikt}}{p_{ik}} = \sum_{k \in \mathcal{K}} \sum_{i \in k} p_{ik} \sum_{t \in \mathcal{T}} \hat{D}_{kt} \hat{C}_{kt} \hat{M}_k. \quad (14)$$

Obecně mohou pozorování $S_{\mathcal{K},\mathcal{T}}$ odpovídat integraci pouze přes čas (pak se jedná o odhad spotřeby daného jedince za delší období) nebo přes jedince (pak se většinou jedná o spotřebu skupiny složené z jedinců jednoho nebo více segmentů; \mathcal{K} značí množinu zákaznických segmentů) či přes čas i jedince současně.

Specifikem plynoucím z výše uvedené „rozpočtové“ úlohy je skutečnost, že pro některé účely se kontroluje jen tvarová shoda modelu s daty, nikoli „výškové usazení“, zatímco pro jiné účely jde o shodu mezi \hat{s}_{ikt} a s_{ikt} . I zde ale přichází ke slovu praktická hodnocení jak dle absolutních, tak dle relativních odchylek. Díky různé integraci v čase jde o kritéria v různém časovém rozlišení.

Hodinový submodel Hodinový submodel slouží k rozpočtu denní spotřeby do hodin. Modelují se podíly jednotlivých hodin na denní spotřebě

$$\pi_{kth} = \frac{\sum_i s_{ikth}}{\sum_{i'} \sum_{j \in t} s_{i'ktj}}, \quad (15)$$

a to po logitové transformaci. Pro logitově transformovaný podíl hodiny h na celkové spotřebě za den t je navržen model

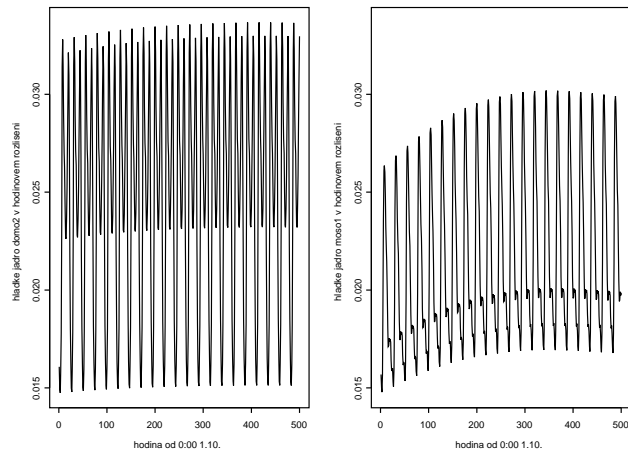
$$\log \left(\frac{\pi_{kth}}{1 - \pi_{kth}} \right) = a_k + b_{kh} + c_{kth} + d_{kth} + \varepsilon_{kth}, \quad (16)$$

kde a_k je absolutní člen, b_{kh} je „obecný“ efekt hodiny h uvnitř dne t , c_{kth} je změna hodinového profilu spojená s pracovním/nepracovním dnem, d_{kth} je hladká změna hodinového profilu v průběhu roku a ε_{kth} je náhodný, přibližně normálně rozdělený člen. Parametry a hladká část hodinového submodelu jsou pak, podobně jako v případě denního submodelu, odhadovány v pracovním gaussovském GAM modelu.

Celkový model Syntézou denního (10) a hodinového (16) modelu byl vytvořen celkový model TDD pro výpočet spotřeby v jednotlivých hodinách. Označíme-li systematickou část $a_k + b_{kh} + c_{kth} + d_{kth}$ modelu (16) jako L_{kth} , počítáme odhad spotřeby i -tého zákazníka k -tého segmentu v h -té hodině dne t jako $\hat{s}_{ikt} \bullet H_{kth}$. Odhad spotřeby za celý den t , tedy $\hat{s}_{ikt\bullet}$, je dán vzorcem (10). Hodinová korekce H_{kth}

$$H_{kth} = \frac{1/(1 + \exp(-L_{kth}))}{\sum_{j=1}^{24} (1/(1 + \exp(-L_{ktj})))} \quad (17)$$

je normalizována tak, aby se celkový model v hodinovém rozlišení sčítal do denního submodelu. Ve dnech s přechodem na letní či zimní čas 24 v součtu



Obrázek 4: Hladké jádro v hodinovém rozlišení pro segmenty domo2 a moso1 v Čechách, prvních 500 hodin od 1. října.

nahrazujeme příslušným počtem hodin (23, respektive 25). Obrázek 4 ukazuje hladké jádro v hodinovém rozlišení. Jde o prvních 500 hodin od 1. října, tedy necelých 21 dní. Vidíme, že chování sezónní části modelu je poměrně komplikované, mj. se mění amplituda denního cyklu. Dny jsou poměrně jasně rozeznatelné pohledem na minima, jež jsou dosahována v nočních hodinách.

Ve „dvoukrokovém“ přístupu založeném na denním a hodinovém submodelu používáme následující soustavy rovnic odhadu tvaru

$$f_D((s_{ikt\bullet})_{i \in k, t=1, \dots, T}, \theta_D) = 0 \quad \text{a} \quad f_H\left(\left(\frac{s_{\bullet kth}}{s_{\bullet kt\bullet}}\right)_{h \in t, t=1, \dots, T}, \theta_H\right) = 0$$

pro denní (θ_D), respektive hodinové (θ_H) parametry namísto simultánních, plně propojených rovnic

$$f((s_{ikth})_{i \in k, h \in t, t=1, \dots, T}, \theta_D, \theta_H) = 0.$$

V důsledku toho sice ztrácíme něco na efektivitě, to však vzhledem k nemalému rozsahu dat není zcela kritické, zvláště v souvislosti s faktem, že hodinové a denní parametry jsou vzájemně téměř ortogonální. Naopak ale získáváme na jednoduchosti odhadu a robustnosti vůči různým nepřesnostem v datech s jemným (hodinovým) časovým rozlišením.

3.4 Možné směry dalšího rozvoje modelů

3.4.1 Modifikace struktury modelu

Teplotní odpověď Odhady (10) používají logistického tvaru teplotní korekce (12), jehož prostřednictvím zachycují základní charakteristiky chování spotřeb vůči teplotě. Můžeme ale uvažovat o případné změně detailního tvaru teplotní odpovědi, např. o záměně funkce $\varphi(\cdot)$ za nějakou $\varphi_{ALT}(\cdot)$. Vždy půjde o klesající křivku s dolní i horní asymptotou, tedy o doplněk distribuční funkce nějakého spojitého rozdělení. Dle detailních analýz provedených nedávno na nově přístupných rozsáhlých datech se zdá, že by mohlo být výhodné připustit i asymetrická rozdělení. Jako zvláště slibný kandidát se jeví doplněk distribuční funkce extrémálního rozdělení (minim), tedy $\varphi_{ALT}(x) = \exp(-\exp(x))$. Nárůst spotřeby z minima bývá totiž leckdy strmější než pokles od maxima. Samozřejmě lze uvažovat i o dalších možnostech včetně například neparametrického odhadu. Je však třeba si uvědomit, že praktický dopad změn zejména pro levý chvost $\varphi_{ALT}(\cdot)$ je relativně malý. Může se projevit pouze v řídkých situacích, kdy je spotřeba blízko fyzického stropu daného maximálním instalovaným výkonem otopových zařízení. Navíc je stabilita odhadu menší. Pro jiné úlohy než pro odhad typické spotřeby v TDD, např. při určování dne s maximální spotřebou a stanovování výše této spotřeby (tzv. „load faktor“) může ovšem být popis levého chvostu klíčovou záležitostí.

Hladké jádro Jak v modelu z oddílu 3.2, tak v modelu z oddílu 3.3, jakož i v dalších podobných formulacích známých z literatury vystupuje člen odpovídající „normálové spotřebě“. V různých modelech je poněkud odlišně parametrizován, ale obecně odpovídá spotřebě za „normálových podmínek“ (daných např. typickou teplotou a dalšími standardními podmínkami) nebo nějaké její monotónní transformaci. I přesto, že dlouhodobě typickou teplotu lze pohodlně získat třeba od ČHMÚ, není odhad C_{kt} nebo obdobného členu $\Psi(t)$ v modelu GAMMA jednoduchý ani bezproblémový. Nejde přitom jen o technickou stránku věci. Neparametrický odhad hladkého jádra lze provést vcelku snadno a na modelovaných datech se chová velmi dobře. Jde spíše o generalizaci tohoto jádra *za* pozorovaná data, tedy o pořízení odhadu, který na základě několika let dostupných měření spotřeb bude předpovídat typické chování obecně. Je jasné, že něco takového je možné jen do určité, velmi omezené, míry. Zlepšení predikčních schopností za určitou mez je velmi obtížné. Lze si sice představit komplikovaný model konstruovaný se znalostí různých expertních představ (včetně klimatologických) a odhadnutý např. Bayesovsky. Složitost takového přístupu by však zřejmě silně převyšovala to, co je vnímáno jako praktické.

Zdá se, že poněkud průchodnější možností pro zlepšování detailního chování modelu je paradoxně zjednodušení jeho struktury. Oproti plynárenskému zadání požadujícímu „průběh spotřeby za standardních podmínek“ jako jednu

z částí modelu, se můžeme pokusit strukturu uvolnit a odhadovat teplotní odpověď jako takovou, bez separace „typické“ složky, tj. hladkého jádra. Tento přístup ale také není úplně jednoduchý. Ukazuje se totiž, že separace teplotní odpovědi a periodicity vede k menší stabilitě odhadů při prudkých teplotních výkyvech. Sníženou stabilitu by bylo ve zjednodušeném modelu nutné kompenzovat, např. komplikovanější strukturou časově zpožděných proměnných. Na druhé straně je model bez hladkého jádra slibný zejména pro segmenty, v nichž je spotřeba dána výlučně otopem, tedy teplotou. Vykazuje totiž lepší chování zejména v přechodových obdobích s výraznou variabilitou teploty (jaro, podzim) a lépe postihuje např. dřívější či pozdější příchod chladného, resp. teplého období. V období „plné zimy“ či „plného léta“ se oba typy modelů chovají podobně, teplotní citlivost modelu s hladkým jádrem je zde dostačující. Modelování přechodových období je ovšem výrazně komplikovanější kvůli nemalé hysterze (odlišnému chování při sestupu a vzestupu teploty) i kvůli výrazné heterogenitě v rozhodování jednotlivých zákazníků, která je navíc v průběhu let poněkud proměnlivá.

3.4.2 Model se smíšenými efekty Modely popsané v sekcích 3.2 či 3.3 separují individuální (ale časově konstantní) a časově-proměnlivou (ale všem jedincům daného segmentu společnou) složku. Obecně tak pro model v denním rozlišení platí

$$Es_{ikt} = p_{ik} f_{kt}, \quad (18)$$

kde p_{ik} je individuální složka a f_{kt} společná složka závislá na dni v roce, kalendářových jevech a aktuální teplotě.

Již na první pohled je patrné, že takový model lze nazírat jako multiplikační obdobu lineárního modelu se smíšenými efekty (LME) [41]. V případě lognormálního chyby je shoda dokonalá. Přesto individuální efekty odhadujeme jinak, a to na základě dlouhodobého „průměru“, tedy několikaletého průměru spotřeby po očištění od vlivů teploty i kalendářových jevů. To je dáno v případě TDD strukturou modelu (8). Při použití značení obdobného jako v rovnici (10) je

$$\hat{p}_{ik} = S_{ik,\mathcal{T}} \left(\sum_{t \in \mathcal{T}} \bar{D}_{kt} \bar{C}_{kt\bullet} \right) / \left(\sum_{t' \in \mathcal{T}} \bar{D}_{kt'} \bar{C}_{kt'\bullet} \right) \quad (19)$$

pro historické spotřeby $S_{ik,\mathcal{T}}$ za dny z nějakého delšího (typicky tříletého) intervalu \mathcal{T} ležícího plně v minulosti. Ve vzorci vystupuje vyhodnocení v normálových ($\bar{\cdot}$) i ve skutečných podmínkách ($\check{\cdot}$). Obdobně lze postupovat i v případě modelu GAMMA.

Alternativně je možné uvažovat o vylepšeních založených na tom, že jednotlivá p_{ik} se odhadují jako skutečné náhodné efekty, např. jako pozorování z lognormálního rozdělení s nějakým populačním průměrem a rozptylem. Nevýhodou je však vyšší výpočetní náročnost. Experimentálně byl takový přístup použit v práci [6] a vedl k zajímavým výsledkům. Práce se zabývala

komplikovanějším modelem zaměřeným na větší spotřebitele, kteří jsou sledováni poněkud podrobněji, než ti, modelovaní v rámci modelů TDD či GAMMA. Očekávaný „shrinkage effect“ byl prakticky ještě výhodnější než ve standardním LME případě. „Shrinkage“ je leckdy větší pro vyšší hodnoty a menší pro hodnoty malé, což zvyšuje stabilitu výsledků.

V práci [6] i v dalších analýzách se také ukázalo, že by mohlo být výhodné ještě další rozšíření modelů GAMMA i TDD, jež jsou v současné podobě spíše statické, a to rozšíření směrem k dynamičnosti. O něm se zmíníme v dalším odstavci.

3.4.3 Dynamický model Výhody, které plynou jak pro interpretaci tak pro praktickou implementaci z totální separace mezi individuální a časově proměnlivou složkou danou multiplikativním tvarem (18), jsme již rozebírali. Na druhé straně je jasné, že interakce mezi jedincem a tvarem či výší spotřeby by nemusela být zcela zanedbatelná. Její modelování však není samozřejmou či jednoduchou záležitostí. Podrobná data, která takový (v analýze rozptylu obvyklý) saturovaný model s interakcemi dovolují, jsou k dispozici jen u malého výběru zákazníků. Praktické použití takového modelu na celém zákaznickém kmeni je naprosto nereálné. Je zde zapotřebí jistý kompromis mezi saturovaným modelem a modelem bez interakce (výše popsáný multiplikativní tvar modelu spotřeby) jakožto dvěma extrémy. Jinými slovy, jde o parsimonní model podstatné části interakce (jedinec)*(tvar odpovědi).

Ukazuje se, že určitá část interakce je koncentrována do pomalé (hladké) změny p_{ik} v čase. Zejména v některých segmentech může být změna (velmi často jde o pokles) p_{ik} spojena se změnami cen, s dlouhodobými opatřeními (např. lepší tepelnou izolací budov), apod. Proto byl navržen model, v němž se připouští, že vývoj spotřeb p_{ik} probíhá jako časová řada, pak tudíž pracujeme s p_{ikt} . Tuto řadu lze popsat například na základě jednoduchého state-space modelu $p_{ikt} = g(p_{ik,t-1}, \varepsilon_{ikt})$, který při vhodné volbě funkce $g(.,.)$ a chyby ε_{ikt} favorizuje jen velmi pozvolné změny. Takto rozšířený model je modelem dynamickým. Žádoucí je přitom dynamická obdoba „shrinkage“ individuální dynamiky p_{ikt} k dynamice průměrné, která je dána dlouhodobým vývojem celého segmentu zákazníků z hlediska spotřeby. Formulace modelu na výběrových, tedy průběhově měřených, datech s velkým časovým rozlišením je relativně schůdná. Problémem je ale formulace modelu použitelného na datech rutinních, vysoce časově agregovaných.

Pro ilustraci zde uvedeme návrh dynamického zobecnění modelu TDD či GAMMA, jež je určeno pro data rutinně odečítaná ve zhruba ročních intervalech tak, že odečty jsou pořizovány v tzv. odečtových cyklech. Zákazníci spadající do stejné odečtové skupiny jsou odečítáni vždy v průběhu stejného kalendářního měsíce. Počet odečtových období označíme n . Pro jednoduchost značení zde budeme uvažovat model pro jediný segment. Přechod ke stratifikovanému modelu pro všechny potřebné segmenty je pak přímočarý.

Komplikací je sezónní průběh odečtů vznikající v důsledku specifických vlastností společně odečítaných skupin zákazníků („reading cycle effect“). Povšimněme si, že jde o sezónní proměnlivost dosti umělého charakteru. Je dána systematickými odlišnostmi mezi skupinami zákazníků odečítaných v jednotlivých měsících, jejichž pojitkem je často pouze to, že jsou lokalizováni blízko sebe. Tato sezónní složka je fundamentálně odlišná od sezónní složky modelované už ve standardních modelech GAMMA či TDD, jež souvisí s chováním teplotně očištěné spotřeby v průběhu roku.

Data (logaritmy spotřeb posunutých o konstantu 0,5 jako v odstavci 3.3) zde označíme jako Z_{it} s tím, že čas t je nyní indexován v měsících nikoli dnech. Data zákazníků s indexy i_1, \dots, i_{d_t} získaná v měsíci t seskupíme do vektoru $\mathbf{Z}_t = (Z_{i_1,t}, Z_{i_2,t}, \dots, Z_{i_{d_t},t})'$. Dimenze d_t vektoru je časově proměnlivá, aby bylo možné zohlednit jak fluktuaci zákazníků, tak chybějící data.

Rovnice pozorování specifikujeme jako

$$\mathbf{Z}_t = \mathbf{f}_t + \mathbf{H}_t \boldsymbol{\eta}_t + \boldsymbol{\varepsilon}_t, \quad (20)$$

kde \mathbf{f}_t je systematická část modelu, daná jako obdoba systematické části pravé strany rovnice (8). $\mathbf{H}_t \boldsymbol{\eta}_t$ odpovídá dynamicky se vyvíjejícímu $\log(p_{ikt})$, observační matice \mathbf{H}_t je popsána níže. Pozorovací chyby jsou nekorelované, $\text{var}(\boldsymbol{\varepsilon}_t) = \sigma^2 \mathbf{I}_{d_t}$. Stavový vektor $\boldsymbol{\eta}_t$ se vyvíjí dle stavové rovnice

$$\boldsymbol{\eta}_t = \mathbf{T}_t \boldsymbol{\eta}_{t-1} + \boldsymbol{\omega}_t, \quad (21)$$

kde $\text{var}(\boldsymbol{\omega}_t) = \boldsymbol{\Sigma}_t$. Komponenty stavového vektoru jsou $\boldsymbol{\eta}'_t = (\mu_t, \beta_t, \boldsymbol{\theta}_t, \phi_t)$, matice strukturálních disturbancí pak $\boldsymbol{\Sigma}_t = \text{diag}\{\boldsymbol{\Sigma}_{\mu\beta}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\phi;t}\}$ a přechodová matice $\mathbf{T}_t = \text{diag}\{\mathbf{T}_{\mu\beta}, \mathbf{T}_{\boldsymbol{\theta}}, \mathbf{T}_{\phi;t}\}$. Její blok $\mathbf{T}_{\mu\beta}$ specifikuje dlouhodobý trend. Osvědčila se volba

$$\mathbf{T}_{\mu\beta} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

jež odpovídá lokálně lineárnímu modelu. Efekt odečtového cyklu specifikujeme (z opatrnosti) také dynamicky ve tvaru

$$\mathbf{T}_{\boldsymbol{\theta}} = \begin{pmatrix} -1 & -1 & \dots & -1 & -1 \\ 1 & 0 & \dots & 0 & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}_{n \times n}.$$

Individuální odchylky od trendu jsou pak specifikovány v následujícím (nejpodstatnějším) bloku

$$\mathbf{T}_{\phi;t} = \begin{pmatrix} \mathbf{0}_{(d_t-1) \times w_t} & \mathbf{I}_{d_t-1} \\ \text{diag}\{\mathbf{I}_{d_t-1}, \mathbf{I}_{d_t-2}, \dots, \mathbf{I}_{d_t-(n-1)-1}\} & \mathbf{0}_{w_t \times (d_t-1)} \end{pmatrix},$$

kde $w_t = \sum_{j=1}^{n-1} (d_t - j - 1)$.

Kovarianční matice stavových disturbancí má obdobnou blokovou strukturu jako \mathbf{T}_t , tedy $\mathbf{\Sigma}_t = \text{diag}\{\mathbf{\Sigma}_{\mu\beta}, \mathbf{\Sigma}_\theta, \mathbf{\Sigma}_{\phi;t}\}$. Konkrétně pak pro jednotlivé bloky specifikujeme $\mathbf{\Sigma}_{\mu\beta} = \text{diag}\{\sigma_\mu^2, \sigma_\beta^2\}$ pro obecný lokálně lineární trend (při volbě $\sigma_\mu^2 = 0$ dostáváme obdobu „smooth trend“ modelu populárního v ekonometrii), dále $\mathbf{\Sigma}_\theta = \text{diag}\{1, 0, \dots, 0\}_{n \times n}$ a konečně $\mathbf{\Sigma}_{\phi;t} = \text{diag}\{\sigma_\phi^2(1 - \rho)\mathbf{I} + \sigma_\phi^2\rho \cdot \mathbf{1} \cdot \mathbf{1}', \mathbf{0}, \dots, \mathbf{0}\}$. Pro individuální odchylky od trendu zde máme „vnitrotřídní“ (intra-class) korelaci $\rho = -1/(d_t - 1)$, indukovanou restrikcí $\sum_{i=1}^{d_t} \phi_{it} = 0$. Ta je potřeba k zachování interpretace trendové složky a její separaci.

Observační matice pro lokálně lineární model (či smooth trend) je

$$\mathbf{H}_t = \left(\left(\mathbf{1} \ \mathbf{0} \right), \left(\mathbf{1} \ \mathbf{0} \ \dots \ \mathbf{0} \right), \left(\left(\begin{array}{c} \mathbf{I}_{d_t-1} \\ -\mathbf{1}'_{d_t-1} \end{array} \right) \ \mathbf{0} \ \dots \ \mathbf{0} \right) \right).$$

Alternativně bychom mohli pracovat s náhodnou procházkou, a to vyřazením komponenty β_t v přechodové, observační i kovarianční matici.

Protože podstatným rysem modelu (20, 21) jsou nestacionární složky zachycující systematické změny v čase, je zásadní věcí specifikace počátečních podmínek. Naštěstí však zde na rozdíl od typické analýzy časových řad máme k dispozici individuální historické informace, z nichž lze vycházet. Počáteční rozptyly volíme nadsazeně (overdispersed).

Podívejme se na interpretace tohoto dynamického rozšíření TDD modelu. Rozptyl disturbancí průměrného trendu σ_μ^2 kontroluje „dynamiku na populační úrovni“, σ_ϕ^2 z bloku $\mathbf{\Sigma}_{\phi,t}$ kontroluje variabilitu mezi individuálními přírůstky. Máme pak $\text{var}([Z_{it} - Z_{i,t-n}] - [Z_{jt} - Z_{j,t-n}]) = 2\sigma_\phi^2 + 4\sigma^2$ pro $i \neq j$. Trochu to připomíná to modelování podél časové a prostorové osy v časoprostorových geostatistických modelech. Analogicky k takovým modelům tu máme časový a „individuální“ směr, u obou je na rozdíl od typických prostorových modelů indexování diskrétní. Ne zcela typická je restrikce na úrovni realizace, $\sum_{i=1}^{d_t} \phi_{it} = 0$.

Model (20, 21) je specifikován, jakmile jsou specifikovány strukturální parametry. Ty ovšem zahrnují rozptyly a kovariance disturbancí. K odhadu používáme metodu maximální věrohodnosti. Numericky maximalizujeme logaritmus věrohodnostní funkce po reparametrizaci komponent rozptylu na jejich logaritmy a používáme „prediction error decomposition“, kterou lze získat snadno z Kalmánova filtru. Poté, co je strukturální část modelu odhadnuta, je aplikace modelu na nová data jen aplikací Kalmánova filtru. Praktické použití tohoto přístupu je však limitováno velkými dimenzemi příslušných systémových matic.

3.4.4 Chybná klasifikace do segmentů Klasifikace do segmentů není v datech z rutinních databází nikterak perfektní. V důsledku toho je skupina zákazníků s vykazovaným segmentem k ve skutečnosti směsí několika segmentů. Uvažujme pro jednoduchost dva segmenty a označme jako π_j pravděpodobnost, že zákazník pocházející ze segmentu j je v rutinní databázi

skutečně správně klasifikován do segmentu j , $j = 1, 2$. Existují studie, ze kterých lze odhady pravděpodobností π_j získat, např. [11].

Máme-li tyto odhady k dispozici a je-li zmíněná chybná klasifikace časově neproměnlivá, lze ji ošetřit podle předpisu

$$\begin{pmatrix} Z_{1t} \\ Z_{2t} \end{pmatrix} = \mathbf{A} \begin{pmatrix} Q_{1t} \\ Q_{2t} \end{pmatrix}, \text{ kde } \mathbf{A} = \begin{pmatrix} \pi_1 & 1 - \pi_1 \\ 1 - \pi_2 & \pi_2 \end{pmatrix}. \quad (22)$$

Přitom Z_{1t}, Z_{2t} představují dva „pozorované“ segmenty namíchané z jejich ideálních, „čistých“ protějšků Q_{1t}, Q_{2t} .

Pokud jsou váhy komponent směsi π_j známé (a není-li $\pi_1 = 1 - \pi_2$), můžeme neznámé Q_t celkem snadno „zrekonstruovat“ ze Z_t pomocí vztahu

$$\begin{pmatrix} Q_{1t} \\ Q_{2t} \end{pmatrix} = \frac{1}{\pi_1\pi_2 - (1 - \pi_1)(1 - \pi_2)} \begin{pmatrix} \pi_2 & -(1 - \pi_1) \\ -(1 - \pi_2) & \pi_1 \end{pmatrix} \begin{pmatrix} Z_{1t} \\ Z_{2t} \end{pmatrix}. \quad (23)$$

Předchozí ideu lze zakomponovat do dynamického modelu (20, 21) ze sekce 3.4.3. Vcelku jednoduše to lze provést, pokud oba segmenty proložíme simultánně, nikoli stratifikovaně, jak jsme naznačovali v sekci 3.4.3, a pozměníme pozorovací rovnici na $Z_t = \mathbf{A}\mathbf{H}_t\boldsymbol{\eta}_t + \boldsymbol{\varepsilon}_t$.

3.4.5 Prostorová struktura Popsaný model TDD pracuje s celostátními teplotami. Ve skutečnosti však existují mezi různými oblastmi v závislosti na zeměpisné šířce i nadmořské výšce nemalé rozdíly. Nabízí se proto možnost pracovat s regionálními teplotami, a to buď s běžně užívanými průměry pro kraj či spíše pro ještě menší území. Zásadní otázkou je, jak velký stupeň prostorového vyhlazení teplotního pole zvolit, resp. jak velké plochy území charakterizovat stejným teplotním údajem. V případě návrhu regionálních TDD bude potřeba adekvátně přizpůsobit výběr a rozsah vzorku průběhově měřených zákazníků. Regionální teploty pak musí být použity při konstrukci modelu. Někteří uživatelé modelu však mají tendenci vkládat nekorektně regionální teploty do modelu postaveného na celostátních teplotách.

Řešitelé se otázkou regionalizace zabývali poměrně podrobně. K analýzám byla využita data získaná z regionálních meteorologických měření prováděných samotnými plynárenskými subjekty v lokalitách s velkou hustotou obyvatelstva a průmyslových podniků. Tato měření pak byla využita pro modelování spotřeb zákazníků s velkým podílem otopové složky. Přínos využití regionálních teplot na nižší úrovni, než je velikost jednoho kraje, je však zatím sporný vzhledem k nákladům spojeným s instalací, provozováním a zpracováním dat z těchto dodatečných meteorologických měření.

3.5 Testování modelů v praxi

Model popsáný v oddíle 3.3 byl implementován do systému SAO provozovaného RWE Transgas. V současné době uživatel testuje přesnost mo-

delu TDD. Je při tom porovnáván výstup modelových hodnot spotřeby segmentu malých a středních zákazníků v denním rozlišení s naměřenými hodnotami spotřeby z dané oblasti (celá ČR, jednotlivé regiony, jednotlivé uzavřené oblasti), od kterých je odečtena naměřená hodnota velkých odběratelů, tzv. vlastní spotřeba a předpokládaná hodnota denních ztrát. Tento rozdíl se pak nazývá bilanční odchylka. Jak už bylo zmíněno výše, modelová hodnota je porovnávána s odhadovanou hodnotou, která sama je zatížena chybou. Kromě tohoto do modelu vstupují i hodnoty o předpokládané roční spotřebě daných zákazníků. Tato spotřeba je v podstatě počítána jako aritmetický průměr spotřeby za poslední 3 roky (podle [49]). Takový výpočet nezohledňuje aktuální trend ani další externí vlivy (např. aktuální cenu plynu), které mohou významně ovlivnit spotřebu plynu zákazníka v období od jeho posledního odečtu. Získat představu o chybě modelu není tedy vůbec jednoduché. Pro základní orientaci se pracuje s dvěma typy chyb, a to se střední absolutní odchylkou, střední relativní odchylkou.

Uživatel pochopitelně požaduje, aby bilanční odchylky byly co nejmenší, a za přijatelnou odchylku považuje tu, jejíž relativní hodnota je v rozmezí -10% až $+10\%$. K tomuto cíli je ovšem daleká cesta, neboť například relativní chyba při stejné „nejistotě“ modelu zákonitě kolísá (vzhledem k sezónnímu charakteru spotřeby zemního plynu musí tato odchylka být výrazně nižší v zimním období s vyšší spotřebou a vyšší v letním období s nižší spotřebou).

V době finalizace tohoto příspěvku můžeme konstatovat, že průměrná hodnota relativní odchylky modelu na úrovni celé ČR se pohybuje okolo 9% (testováno za období cca jednoho roku).

Problematika vyhodnocování přesnosti modelu TDD je předmětem probíhajících diskusí s uživatelem. Proces vedoucí k dalšímu vylepšení výkonnosti modelu stále probíhá, a to jak na úrovni vývoje nových verzí statistického modelu (viz oddíl 3.4), tak i na úrovni detekování a odstraňování chyb, které se stále, byť v čím dál menší míře, ve vstupních datech vyskytují.

4 Závěr

Jak již vyplynulo z výše uvedeného, problematika statistického modelování spotřeby zemního plynu od návrhu modelu, přes jeho další vývoj, implementaci, testování a rutinní využívání je velice složitý proces. Řešitelé považují za jeden z hlavních přínosů jejich práce fakt, že statistické modely ukázaly důležitost kvality dat měřených a získávaných v plynárenství. Používání statistického modelu vedlo v první fázi k detekci chyb, které evidentně nebyly způsobeny modelem, ale problémy ve vstupních datech. To vedlo k zahájení procesů kontrolujících a výrazně vylepšujících kvalitu dat. V plynárenských kruzích panoval názor, že v datech nejsou chyby anebo jsou to chyby nepodstatné, či chyby, které přináší život a nelze s nimi nic dělat. Vyvrácení tohoto mýtu zůstane významným přínosem projektu bez ohledu na jeho další vývoj.

Ve vyspělých zemích, jako jsou USA, Francie či Japonsko se vývoj patrně bude ubírat implementací moderních komunikačních a informačních techno-

logií v procesu automatického měření každého zákazníka. Pokud bude každý zákazník průběžně měřen, statistického modelování ve výše uvedeném smyslu nebude potřeba a problém se zredukuje pouze na úlohu predikce spotřeby. Nutno však podotknout, že proces osazení každého zákazníka měřicím přístrojem je velmi drahý a zdlouhavý. Například ve francouzské společnosti EDF je takový proces naplánován na období 10 let (2006–2016) a cena se vyšplhá na stovky milionů až několik miliard eur. Dá se očekávat, že země s méně výkonnou ekonomikou půjdou cestou měření zákazníků s velkým a středním odběrem, zatímco velký počet zákazníků s nižším odběrem bude modelován statisticky. Řešitelé budou rádi, pokud tomu budou moci v rámci svých možností přispívat i nadále.

Literatura

- [1] Bailey J. (2000). *Load profiling for retail choice examining a complex and crucial component of settlement*. Electricity Journal **13**, 10, 69–74.
- [2] Bartels R., Fiebig D. G., Nahm D. (1996). *Regional end-use gas demand in Australia*. Economic Record **72**, 319–331.
- [3] Bauwens L., Fiebig D. G., Steel M. F. J. (1994). *Estimating end-use demand: A Bayesian approach*. Journal of Business and Economic Statistics **12**, 221–231.
- [4] Bečvář J., Blahová L., Pihera M., Pelikán E., Brabec M. (2007). *Experiment mimořádných měsíčních odečtů u zákazníků kategorie maloodběr a domácnost*. Plyn **87**, 12, 272–274.
- [5] Berrisford H. G. (1965). *The relation between gas demand and temperature: A study in statistical demand forecasting*. Operational Research Quarterly **16**, 2, 229–246.
- [6] Brabec M., Konár O., Malý M., Pelikán E. (2008). *A nonlinear mixed effects model for prediction of natural gas consumption by individual customers*. International Journal of Forecasting **24**, 4, 659–678.
- [7] Brabec M., Konár O., Malý M., Pelikán E., Vondráček J. (2009). *A statistical model for natural gas standardized load profiles*. Journal of the Royal Statistical Society C (Applied Statistics) **58**, 1, 123–139.
- [8] Bunn D. W., Farmer E. D. (eds.) (1985). *Comparative models for electrical load forecasting*. Wiley, Chichester.
- [9] Cariou V. (2006). *Extension of multivariate regression trees to interval data. Application to electricity load profiling*. Computational Statistics **21**, 325–341.
- [10] Cottet R., Smith M. (2003). *Bayesian modeling and forecasting of intraday electricity load*. Journal of the American Statistical Association **98**, 839–849.
- [11] Čermáková J., Matějovic J., Naxerová O., Bečvář J., Brabec M., Brabec T., Konár O., Malý M., Pelikán E., Šimůnek M., Vondráček J. (2005). *Matematické modelování spotřeby zemního plynu zákazníků bez průběhového měření*. Plyn **85**, 2, 34–37.

- [12] Čermáková J. (2008). *Reengineering energetického distribučního procesu v oblasti informačních toků*. Disertační práce, FS ZČU Plzeň.
- [13] Chambers J. M., Hastie T. J. (eds.) (1991). *Statistical models in s*. Chapman & Hall, New York.
- [14] Feinberg E. A., Genethliou D. (2006). *Load forecasting*. In: Chow J. H., Wu F. F., Momoh J. (eds.) Applied mathematics for restructured electric power systems. Optimization, control, and computational intelligence. Springer, New York, 269–285.
- [15] Fildes R., Randall A., Stubbs P. (1997). *One day ahead demand forecasting in the utility industries: Two case studies*. Journal of the Operational Research Society **48** (1), 15–24.
- [16] Gabbi G., Zanotti G. (2005). *Climate variables and weather derivatives: Gas demand, temperature and seasonality effects in the Italian case*. http://www.fma.org/SLC/Papers/FMA_weather.pdf, accessed on 30 November 2008.
- [17] Geiger B., Hellwig M. (2003). *Gaslastprofile für nicht leistungsgemessene Kunden*. BWK – Das Energie Fachmagazin **55**, 72–78.
- [18] Gil S., Deferrari J. (2004). *Generalized model of prediction of natural gas consumption*. Transactions of the ASME **126**, 90–97.
- [19] Harvey A., Koopman S. J. (1993). *Forecasting hourly electricity demand using time-varying splines*. Journal of the American Statistical Association **88**, 1228–1236.
- [20] Hastie T., Tibshirani R., Friedman J. (2001). *The elements of statistical learning*. Springer, New York.
- [21] Ivezić D. (2006). *Short-term natural gas consumption forecast*. FME Transactions **34**, 165–169.
- [22] Konár O. (2006). *Matematické modelování spotřeby zemního plynu domácností a malooběratelů*. Sborník ROBUST'2006, JČMF, Praha, 159–166.
- [23] Johnson R. A., Wichern D. W. (1988). *Applied multivariate statistical analysis*. Prentice Hall, Englewood Cliffs.
- [24] Johnston J. (1984). *Econometric methods*. McGraw–Hill, New York.
- [25] Liedermann P. (2006). *Typové diagramy dodávky elektřiny – náhradní metoda pro rozúčtování odběrů bez průběhového měření*. Energetika **56**, 2, 402–405.
- [26] Liu J. M., Chen R., Liu L.-M., Harris J. L. (2006). *A semi-parametric time series approach in modeling hourly electricity loads*. Journal of Forecasting **25**, 537–559.
- [27] Lyness F. K. (1984). *Gas demand forecasting*. Statistician **33**, 9–21.
- [28] Mirkin B. (2005). *Clustering for data mining*. Chapman & Hall, Boca Raton.

- [29] Musílek P., Pelikán E., Brabec T., Šimůnek M. (2006). *Recurrent neural network based gating for natural gas load prediction system*. IEEE International Joint Conference on Neural Networks, Vancouver, Canada, July 16-21, 2006, Omnipress, Madison, 7127–7132.
- [30] Peirson J., Henley A. (1994). *Electricity load and temperature. Issues in dynamic specification..* Energy Economics **16**, 4, 235–244.
- [31] Pelikán E., Eben K., Šimůnek M., Kolman M., Hais J. (2000). *Predikční systém Elvíra pro predikci spotřeby zemního plynu*. Plyn **80** (11), 253–255.
- [32] Pelikán E., Eben K., Šimůnek M. (2001). *Predikce spotřeby elektrické energie a zemního plynu*. AUTOMA, odborný časopis pro automatizační techniku **7** (4), 15–18.
- [33] Pelikán E., Šimůnek M. (2005). *Gas consumption forecasting: Risk management and genetic algorithms*. Neural Network World Journal **15**, 425–436.
- [34] Pelikán E., Šimůnek M., Brabec T. (2004). *Load forecasting using the system Elvíra*. Proceedings of the 3rd SIMONE workshop (CD), Český Krumlov, May 11-14, 2004.
- [35] Potočnik P., Thaler M., Govekar E., Grabec I., Poredoš A. (2007). *Forecasting risks of natural gas consumption in Slovenia*. Energy Policy **35**, 4271–4282.
- [36] Potočnik P., Govekar E., Grabec I. (2007). *Short-term natural gas consumption forecasting*. Proceedings of the 16th IASTED International Conference on Applied Simulation and Modelling, Palma de Mallorca, August 29-31, 2007 (A Publication of IASTED, 581). ACTA Press, Anaheim, 353–357.
- [37] R Development Core Team (2008). *The R project for statistical computing*. <http://www.r-project.org/>, accessed on 30 April 2008.
- [38] Ramsay J. O., Silverman B. W. (2002). *Applied functional data analysis: methods and case studies*. Springer, New York.
- [39] Venables W. N., Ripley B. D. (1994). *Modern applied statistics with S-Plus*. Springer, New York.
- [40] Rodríguez-Poo J. (2000). *Constrained nonparametric regression analysis of load curves*. Empirical Economics **25**, 229–246.
- [41] Ruppert D., Wand M. P., Carroll R. J. (2003). *Semiparametric regression*. Cambridge University Press, Cambridge.
- [42] Sánchez-Úbeda E. F., Berzosa A. (2007). *Modeling and forecasting industrial end-use natural gas consumption*. Energy Economics **29**, 710–742.
- [43] Skinner N. H. (1984). *Load research and its application to electricity demand forecasting*. Statistician **33**, 65–73.
- [44] Suykens J., Lemmerling P., Favoreel W., De Moor B., Crepel M., Briol P. (1996). *Modelling the Belgian gas consumption using neural networks*. Neural Processing Letters **4**, 157–166.

- [45] Taylor J. W., McSharry P. E. (2007). *Short-term load forecasting methods: An evaluation based on European data*. IEEE Transactions on Power Systems **22**, 4, 2213–2219.
- [46] Vajk I., Hetthéssy J. (2005). *Load forecasting using nonlinear modelling*. Control Engineering Practice **13**, 7, 895–902.
- [47] Vondráček J., Pelikán E., Konár O., Čermáková J., Eben K., Malý M., Brabec M. (2008). *A statistical model for the estimation of natural gas consumption*. Applied Energy **85**, 5, 362–370.
- [48] Wang G. C. S. (2004). *Forecasting practices in electric and gas utility companies*. Journal of Business Forecasting Methods & Systems **23**, 11–15.
- [49] Vyhláška č. 524/2006 Sb. o pravidlech pro organizování trhu s plynem a tvorbě, přiřazení a užití typových diagramů dodávek plynu (příloha 9) se změnami: 184/2007 Sb., 321/2007 Sb., 354/2008 Sb.

Poděkování: Tato práce byla podporována grantem GA AV ČR č. 1ET400300513 a výzkumným záměrem ÚI AV ČR č. AV0Z10300504.

Adresa: E. Pelikán, M. Brabec, M. Malý, O. Konár, J. Vondráček, Ústav informatiky AV ČR, Pod vodárenskou věží 2, 182 07 Praha 8

E-mail: pelikan@cs.cas.cz

REGRESSION IN SOBOLEV SPACES USING TOTAL LEAST SQUARES

Michal Pešta

Keywords: Total least squares, nonparametric regression, Sobolev spaces.

Abstract: *Total least squares* (TLS) serves as a data modeling technique which can be used for many types of statistical analyses, e.g. a regression. In the regression setup, both dependent and independent variables are considered to be measured with errors. Thereby, the TLS approach in statistics is sometimes called an *errors-in-variables* (EIV) modeling and, moreover, this type of regression is usually known as an *orthogonal regression*.

We propose a class of *nonparametric* estimates for the EIV models over the sets of sufficiently *smooth* functions. The estimation takes place over the balls of functions which are elements of a suitable *Sobolev space*—special type of Hilbert spaces that facilitate calculation of the (total) least squares projection. The Hilbertness allows us to take projections and hence to decompose spaces into mutually orthogonal complements. Then we transform the problem of searching for the best fitting function in an infinite dimensional space into a finite dimensional optimization problem.

The regression set-up proposed by [6] will be extended and combined with the total least squares approach introduced by [1].

Abstrakt: *Úplne najmenšie štvorce* (TLS) slúžia ako modelovacia technika, ktorá môže byť použitá pre rôzne druhy štatistických analýz (napr. regresia). V regresii sú obe (závislá i nezávislá) premenné považované za merané s chybami. Preto je TLS prístup v štatistike niekedy nazývaný *chyby-v-premenných* (EIV) modelovanie a navyše je tento typ regresie obvykle známy ako *ortogonálna regresia*.

Uvažujeme triedu *neparametrických* odhadov pre EIV modely cez množiny dostatočne *hladkých* funkcií. Odhady prebiehajú cez gule funkcií, ktoré sú prvkami vhodného *Sobolevovho priestoru*. Sobolevove priestory sú špeciálny typ Hilbertových priestorov, ktoré umožňujú projekciu vzhľadom k úplne najmenším štvorcům. Hilbertovskosť nám napomáha robiť projekciu a teda rozložiť priestor do navzájom kolmých doplnkov. Potom prevedieme problém hľadania najlepšie aproximujúcej funkcie v priestore nekonečnej dimenzie na konečno-dimenzionálny optimalizačný problém.

Regresné nastavenie navrhnuté v [6] bude rozšírené a skombinované s prístupom úplne najmenších štvorcov predstavených v [1].

1 Introduction

Observing several characteristics, which may be thought of as variables, straightforwardly postulates a natural question: “What is the relationship

between these measured characteristics?” Some of the characteristics might be explained by a (functional) dependence on the other characteristics. Therefore, we consider the first mentioned variables as *dependent* or *response* and the rest as *independent* or *explanatory*.

Our proposed model of dependence contains errors in the response variables and in the explanatory variables as well. One may think of this as *measurements errors* or *disturbances* on *both* sides of the functional relationship, i.e. errors-in-variables model.

1.1 Idea and motivation

Let us consider the simplest one-dimensional situation when one observes *input data* $[\mathbf{x}, \mathbf{y}] \equiv [(x_1, \dots, x_n)^\top, (y_1, \dots, y_n)^\top]$. Moreover, these *observations* are considered to be measured with additive random *errors* $[\boldsymbol{\delta}, \boldsymbol{\varepsilon}]$. *Unobservable true values* $[\mathbf{x} + \boldsymbol{\delta}, \mathbf{y} + \boldsymbol{\varepsilon}]$ satisfy an *unknown* functional relationship, i.e. *regression*

$$y_i + \varepsilon_i = f(x_i + \delta_i), \quad i = 1, \dots, n.$$

Unknown function f is thought to be *smooth*. The smoothness needs to be properly defined somehow. Suppose that our “smoothness” is the only assumption and, thereby, we want a modeling technique to be applicable on *various* types (large number) of data. Finally, we are searching for a suitable estimate \hat{f} , where the misfit needs to be “*as small as possible*”.

2 Sobolev spaces and total least squares

A wide applicability of the method for finding a suitable estimate in our setup yields a *nonparametric* approach as an adequate technique. Smoothness of the estimator function \hat{f} needs to be ensured, but, e.g., kernels, splines, or wavelets can be *too restrictive*. Therefore, we fit a function from a *general class* of smooth functions—*Sobolev spaces* equipped with a corresponding *Sobolev norm*

$$\left(\mathcal{H}^m, \|\cdot\|_{Sob,m} \right) := \left\{ g \in \mathbb{L}^2 : \|g\|_{Sob,m} := \left(\sum_{i=0}^m \int |g^{(i)}(t)|^2 dt \right)^{1/2} < +\infty \right\}.$$

Previous definition indicates that unknown function $f \in \mathcal{H}^m$ needs to have derivatives up to the order m and, hence, one may speak about the *order* of corresponding Sobolev space. In many physical or econometric relationships, the order $m = 2$ seems to be quite satisfactory.

2.1 Graphical illustration

The observed data should be “as close as possible” to the true unobservable values, or in other words, the errors $[\boldsymbol{\delta}, \boldsymbol{\varepsilon}]$ should be “as small as possible”.

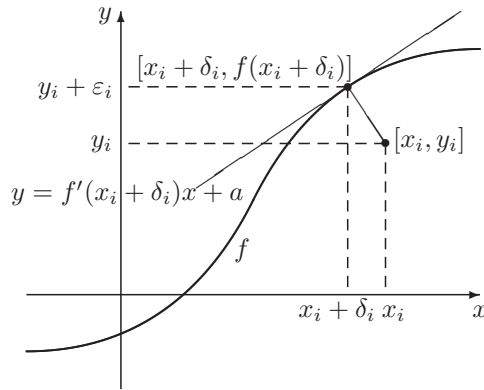


Figure 1: Orthogonal regression in the EIV setup with a smooth function.

This can be reached by measuring the misfit in the “shortest” way, i.e. taking the perpendicular distance into account as demonstrated in Figure 1.

Since one assumes that $m \geq 1$, a tangent can be constructed for function f in its each point and, hence, the orthogonal distance from input values can be measured. This orthogonal distance clearly corresponds to the Euclidean norm of the errors $[\delta, \varepsilon]$, i.e. total least squares. TLS method is just another name for orthogonal regression in statistics.

On the other hand, the smoothness (or “wildness”) of unknown function f is measured by its Sobolev norm. Hence, one should realize that the *better* the fit the *wilder* the function and vice versa. This can be written in an informal way

$$\frac{\text{small}}{\text{large}} \left\| \begin{bmatrix} \delta \\ \varepsilon \end{bmatrix} \right\|_2 = \frac{\text{large}}{\text{small}} \|f\|_{Sob,m}.$$

2.2 Estimate

Searching for an estimate \hat{f} is simply nothing else than finding a reasonable *compromise* between misfit (Euclidean norm of the error vector) and smoothness (Sobolev norm of the estimated function). This compromise can be easily incorporated using so-called *smoothing parameter* $\chi > 0$:

$$\min_{f \in \mathcal{H}^m, \delta \in \mathbb{R}^n, \varepsilon \in \mathbb{R}^n} \left\{ \left\| \begin{bmatrix} \delta \\ \varepsilon \end{bmatrix} \right\|_2^2 + \chi \|f\|_{Sob,m}^2 \right\}, \quad \text{s.t. } \mathbf{y} + \varepsilon = \mathbf{f}(\mathbf{x} + \delta). \quad (1)$$

But the optimizing problem (1) is very complicated to solved directly.

Let us consider for simplicity interval $[0, 1]$ as a bounded domain, where our x -data lie. Sobolev space on this interval $\mathcal{H}^m[0, 1]$ is a Banach space and one can define *Sobolev inner product* for each $g, h \in \mathcal{H}^m[0, 1]$:

$$\langle g, h \rangle_{Sob,m} := \sum_{i=0}^m \int_0^1 g^{(i)}(t)h^{(i)}(t)dt. \quad (2)$$

Hence, $\mathcal{H}^m[0, 1]$ is also a Hilbert space. Using Riesz representation theorem and Arzelà-Ascoli theorem, [6] proved the following: for all $f \in \mathcal{H}^m[0, 1]$ and for any $a \in [0, 1]$, there exists $\psi_a \in \mathcal{H}^m[0, 1]$ such that

$$f(a) = \langle \psi_a, f \rangle_{Sob, m} = \sum_{i=0}^m \int_0^1 \psi_a^{(i)}(t) f^{(i)}(t) dt, \quad (3)$$

and ψ_a is called a *representor* at the point a . Hence, one may easily derive so-called *representor matrix* $\Psi_{n \times n}(\mathbf{t})$ whose columns (and rows) equal the representors *evaluated* at t_1, \dots, t_n

$$\Psi_{ij}(\mathbf{t}) = \langle \psi_{t_i}, \psi_{t_j} \rangle_{Sob, m} = \psi_{t_i}(t_j) = \psi_{t_j}(t_i), \quad \forall i, j. \quad (4)$$

The representor matrix is *symmetric* and *positive definite* as proved in [4].

A form of the representors was derived by [5]:

$$\begin{aligned} \psi_a(t) = \sum_{k=1}^{2m} \exp[\Re(e^{i\theta_k})t] \left\{ I_{[t \leq t_j]} \gamma_k(t_j) \cos[\Im(e^{i\theta_k})t] \right. \\ \left. + I_{[t > t_j]} \gamma_{2m+k}(t_j) \sin[\Im(e^{i\theta_k})t] \right\}. \end{aligned}$$

Here, the coefficients γ_{ks} and θ_{ks} are determined as a solution of the ordinary differential equation with some boundary conditions.

Let $\mathcal{M} := \text{span}\{\psi_{x_i + \delta_i} : i = 1, \dots, n\}$ and, afterwards, its orthogonal complement $\mathcal{M}^\perp = \{h \in \mathcal{H}^m[0, 1] : \langle \psi_{x_i + \delta_i}, h \rangle_{Sob, m} = 0, i = 1, \dots, n\}$. The Sobolev space can be written as a direct sum of its orthogonal subspaces, i.e. $\mathcal{H}^m[0, 1] = \mathcal{M} \oplus \mathcal{M}^\perp$ since $\mathcal{H}^m[0, 1]$ is a Hilbert space. Function $h \in \mathcal{M}^\perp$ takes on the value zero at $x_1 + \delta_1, \dots, x_n + \delta_n$ due to the property (3). Each $f \in \mathcal{H}^m[0, 1]$ can be written in form

$$f = \sum_{i=1}^n c_i \psi_{x_i + \delta_i} + h, \quad h \in \mathcal{M}^\perp. \quad (5)$$

Then, one can rewrite the objective function from (1) incorporating the corresponding restriction simply by its substitution, applying the relation (5) and the representation (3), utilizing linearity of the Sobolev inner product (2), and using the definition of representor matrix (4) with its property of being symmetric

$$\begin{aligned}
\left\| \begin{bmatrix} \boldsymbol{\delta} \\ \boldsymbol{\varepsilon} \end{bmatrix} \right\|_2^2 + \chi \|f\|_{Sob,m}^2 &= \|\boldsymbol{\delta}\|_2^2 + \left\| \mathbf{y} - \left\langle \boldsymbol{\psi}_{\mathbf{x}+\boldsymbol{\delta}}, \sum_{i=1}^n c_i \psi_{x_i+\delta_i} + h \right\rangle_{Sob,m} \right\|_2^2 \\
&+ \chi \left\langle \sum_{i=1}^n c_i \psi_{x_i+\delta_i} + h, \sum_{i=1}^n c_i \psi_{x_i+\delta_i} + h \right\rangle_{Sob,m}^2 \\
&= \|\boldsymbol{\delta}\|_2^2 + \|\mathbf{y} - \boldsymbol{\Psi}(\mathbf{x} + \boldsymbol{\delta})\mathbf{c}\|_2^2 \\
&+ \chi \mathbf{c}^\top \boldsymbol{\Psi}(\mathbf{x} + \boldsymbol{\delta})\mathbf{c} + \chi \|h\|_{Sob,m}^2
\end{aligned}$$

where for an arbitrary $g \in \mathcal{H}^m$ the following convention is used $\langle \boldsymbol{\psi}_{\mathbf{t}}, g \rangle_{Sob,m} = (\langle \boldsymbol{\psi}_{t_1}, g \rangle_{Sob,m}, \dots, \langle \boldsymbol{\psi}_{t_n}, g \rangle_{Sob,m})^\top$. Note further that $\sum_{i=1}^n c_i \psi_{x_i+\delta_i}$ minimizes (1) and, moreover, is the unique solution of that optimizing problem, because $\psi_{x_i+\delta_i}$ are the base elements of \mathcal{M} . Therefore, the *infinite dimensional* minimizing (1) is transformed into the *finite dimension*

$$\min_{\mathbf{c} \in \mathbb{R}^n, \boldsymbol{\delta} \in \mathbb{R}^n} \left\{ \|\boldsymbol{\delta}\|_2^2 + \|\mathbf{y} - \boldsymbol{\Psi}(\mathbf{x} + \boldsymbol{\delta})\mathbf{c}\|_2^2 + \chi \mathbf{c}^\top \boldsymbol{\Psi}(\mathbf{x} + \boldsymbol{\delta})\mathbf{c} \right\}. \quad (6)$$

A solution $\{\hat{\mathbf{c}}, \hat{\boldsymbol{\delta}}\}$ of the finite optimizing problem (6) always exists and is unique, which can be proved similarly as in [5]. A *derivative of representor matrix* $\boldsymbol{\Psi}_{n \times n}^{(1)}(\mathbf{t})$ needs to be defined as a matrix whose columns are equal to the first derivatives of the representors evaluated at t_1, \dots, t_n ; i.e.

$$\Psi_{i,j}^{(1)} = \psi'_{t_j}(t_i), \quad i, j = 1, \dots, n.$$

Now, by setting all the partial derivatives of the objective function in (6) with respect to all elements of \mathbf{c} and $\boldsymbol{\delta}$ equal zero, and taking into account the existence of the inverse of representor matrix (due to its positive definiteness), one can end up with a system of equations

$$\begin{aligned}
\left[\boldsymbol{\Psi}(\mathbf{x} + \hat{\boldsymbol{\delta}}) + \chi \mathbf{I} \right] \hat{\mathbf{c}} &= \mathbf{y}, \\
\left[\mathbf{y} - \boldsymbol{\Psi}(\mathbf{x} + \hat{\boldsymbol{\delta}})\hat{\mathbf{c}} - \frac{\chi}{2}\hat{\mathbf{c}} \right] \boldsymbol{\Psi}^{(1)}(\mathbf{x} + \hat{\boldsymbol{\delta}}) &= \hat{\boldsymbol{\delta}},
\end{aligned}$$

which can be solved iteratively.

Once we find $\hat{\mathbf{c}}$ and $\hat{\boldsymbol{\delta}}$, a *unique* estimate \hat{f} can be obtained by

$$\hat{f} = \sum_{i=1}^n \hat{c}_i \psi_{x_i+\hat{\delta}_i}.$$

3 Examples

Our technique will be demonstrated on two totally different real data sets. If we do not have any idea about the nature of our data, one cannot simply

use a special technique. Here comes our method. Surely, our technique can behave worse on one concrete data set than an “appropriate” method for that kind of data. On the other hand, we do not lose as much as in a situation when an inappropriate method is chosen due to the lack of information about the data, e.g., a Pareto type model for estimating a probability density which appears to be bimodal.

The first data are the result of a National Institute of Standards and Technology (NIST) study involving the *thermal expansion of copper*. The response variable is the coefficient of thermal expansion and the predictor variable is temperature in kelvin. The data contain 236 observations and were firstly described by [2]. The precision of the thermometer used is surely not zero and that is why some disturbances in measured temperature should be taken into account in our model. Our fit can be seen in Figure 2.

The second data are monthly averaged atmospheric pressure differences between Easter Island (Pacific) and Darwin (Australia) and can be found in [3]. This difference drives the trade winds in the southern hemisphere. Cycles in the pressure differences correspond to the *El Niño* and the *Southern Oscillation*. These data contain 168 observations and errors should be taken in the account in the explanatory variable (time) as well. The reason for this is very simple—one cannot know whether the data were collected weekly on the same day or daily at the same hour, and also simultaneously on both locations. The fitted curve for our technique is again shown in Figure 2.

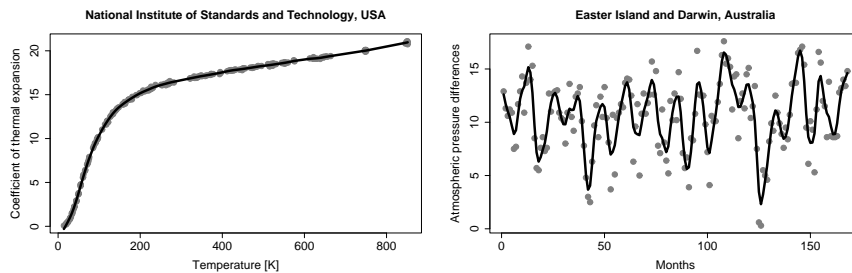


Figure 2: Thermal Expansion of Copper (left) and El Niño – Southern Oscillation (right).

4 Discussion and conclusions

In this paper, regression in Sobolev spaces using TLS is developed. Sobolev spaces provide the only general restriction—smoothness—on the unknown estimated function. Total least squares helps to incorporate (measurement) errors in the explanatory variable and in the predictor as well. Joining these two mathematical constructions together yields a method which easily provi-

des an estimate with demanded properties as shown in Section 2. Thus, one may conclude that it is a *very general* and *widely applicable* nonparametric smoothing technique as it was demonstrated in Section 3.

Moreover, our method *works without a prior knowledge* of functional relation or error distribution. This makes our technique quite robust with respect to the nature of data.

4.1 Remarks

It has to be remarked that regression in Sobolev spaces using TLS can be easily extended into a *multivariate* case, meaning more dimensions for dependent and independent variable as well.

A smoothing parameter χ incorporated in (1), which controls the *trade-off* between the infidelity to the data versus the roughness of the estimated solution, can be chosen according, e.g., a *cross-validation* criterion

$$\mathcal{CV}(\chi) = \frac{1}{n} \sum_{i=1}^n \left[y_i - \hat{f}_{-i}(x_i + \delta_i) \right]^2$$

where \hat{f}_{-i} is obtained by solving

$$\min_{f, \delta_{-i}, \varepsilon_{-i}} \left\{ \left\| \begin{bmatrix} \delta \\ \varepsilon \end{bmatrix}_{-i} \right\|_2^2 + \chi \|f\|_{Sob,m}^2 \right\} \quad \text{s.t. } \mathbf{y}_{-i} + \varepsilon_{-i} = \mathbf{f}(\mathbf{x}_{-i} + \delta_{-i})$$

where the subscript $-i$ denotes omitted the i -th element of the corresponding vector. A starting value of χ can be chosen “ad hoc” by trying on values from the logarithmic lattice. But the cross-validation need not to be the only one which can provide a reasonable choice of the previously mentioned parameter, i.e. generalized cross-validation or various information criteria.

Regression in Sobolev spaces allows to add so-called *isotonic* restrictions (for more details see [5]) to the estimated function, e.g., monotonicity or convexity. Therefore, one can perform a cumulative density function or a probability density function estimation. Another application of isotonicity can be found in testing unimodality of a general regression curve.

Unfortunately, we have to admit that our method has a disadvantage. There is a problem of the invariant estimate with respect to a change of the variable magnitudes (scale). Let us consider the simplest one-dimensional setup discussed above. When the values of the explanatory variable are divided by a factor of two and the values of the response variable are kept unchanged, our technique provides a different estimated curve than the originally fitted curve shrunk twice in the horizontal way. This problem can be solved incorporating *scaled* total least squares (STLS) with a scaling parameter $\gamma > 0$. Hence, optimizing problem of finding an estimate \hat{f} is shown as follows

$$\min_{f \in \mathcal{H}^m, \delta \in \mathbb{R}^n, \varepsilon \in \mathbb{R}^n} \left\{ \|\delta\|_2^2 + \gamma \|\varepsilon\|_2^2 + \chi \|f\|_{Sob,m}^2 \right\}, \quad \text{s.t. } \mathbf{y} + \varepsilon = \mathbf{f}(\mathbf{x} + \delta). \quad (7)$$

On the other hand, an additional theoretical problem of the choice of scaling parameter appears when compared with previous simpler optimizing problem (1). Altogether, STLS approach (7) can be viewed as a skewed orthogonal regression, when one takes into account not the perpendicular distance to the unknown function, but the skewed one with a certain angle represented by parameter γ . This scaling parameter can also serve as an *emphasizing parameter*, because it distributes emphasis on the errors corresponding to independent and dependent variables.

4.2 Further research

For a further research, one may be interested in the extension of the TLS approach in the parametric nonlinear regression. Our method might be extended for functional data, because Hilbert-Schmidt operator nicely “shakes hands” with Sobolev spaces.

Our approach of regression in Sobolev spaces using TLS needs to be studied from the statistical point of view as well. Consistency and asymptotic normality of the estimate should be explored. Moreover, bootstrap techniques could be applied for confidence intervals and hypothesis testing.

References

- [1] Golub G.H., C.F. Van Loan (1980). *An analysis of the total least squares problem*. SIAM Journal on Numerical Analysis, **17**, 6, 883–893.
- [2] Hahn T.A. (1970). *Thermal expansion of copper from 20 to 800K—standard reference material 736*. Journal of Applied Physics, **41**, 5096–5101.
- [3] Kahaner D.C. Moler, Nash S. (1989). *Numerical methods and software*. Englewood Cliffs, NJ: Prentice Hall, pp. 441–445.
- [4] Pešta M. (2006). *Isotonic regression in Sobolev spaces*. Master Thesis, Charles University in Prague, Faculty of Mathematics and Physics.
- [5] Pešta M. (2006). *Ideas and remarks on isotonic regression in Sobolev spaces*. Robust 2006 Conference Proceedings, edited by J. Antoch and G. Dohnal, 253–260.
- [6] Yatchew A., L. Bos (1997). *Nonparametric least squares estimation and testing of economic models*. Journal of Quantitative Economics, **13**, 81–131.

Acknowledgement: The present work was supported by grants GA ČR 201/05/H007 and GAUK B-MAT 49809.

Address: M. Pešta, MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8 – Karlín

E-mail: pesta@karlin.mff.cuni.cz

MOMENT ESTIMATES FOR INHOMOGENEOUS PLANAR CLUSTER PROCESSES

Michaela Prokešová

Keywords: Planar cluster process, inhomogeneous Neyman-Scott process, minimum contrast estimation, K-function, composite likelihood, Palm likelihood estimation.

Abstract: In the paper we discuss parameter estimation procedures for planar point process models, which exhibit small scale clustering and large scale inhomogeneity. The interest in such models was stimulated recently by the need to analyze and availability of complex ecological datasets coming e.g. from large scale tropical rain forest studies. Currently the most popular class of such models are the so called SOIR inhomogeneous Neyman-Scott processes. In the paper we give an overview of both the classical and recently developed estimation procedures for these processes and compare their performance by a simulation study.

Abstrakt: Článek se zabývá odhadem parametrů rovinných bodových procesů, které se vyznačují shlukováním v malém měřítku a nehomogenitou ve velkém měřítku. Nedávný zájem o takové modely byl vyvolán dostupností a potřebou analyzovat složité datové soubory pocházející například z rozsáhlých mapování tropického dešťného pralesa. V současné době jsou nejpopulárnější třídou takových modelů takzvané SOIR nehomogenní Neyman-Scottovy procesy. V článku podáme přehled klasických i nově vyvinutých odhadovacích metod pro tyto procesy a jejich chování porovnáme pomocí simulační studie.

1 Introduction and background

Many of the emerging spatial point pattern data are inhomogeneous in nature (e.g. [4]). Conventional point process models often have difficulty in modeling such data since they were developed under stationarity (e.g. homogeneity). To solve the problem, a large number of new models have been recently proposed, the majority of them being inhomogeneous Gibbs point process models. These are able to model repulsive spatial interaction, but are limited in modeling attractive point patterns. Therefore [10] proposed a new class of inhomogeneous Neyman-Scott process that allow attractive interaction between points. This class of models is analytically simple yet practically sensible for modeling attractive point patterns especially those arising from ecological studies (like tropical rain forest surveys). Statistical properties of estimators of the first-order structure (i.e. the inhomogeneity parameters) were also studied in the paper. The second-order parameters (i.e. interaction parameters) were estimated by the minimum contrast estimation. Nevertheless

this is not the only option and several other procedures appearing in the literature recently in different contexts can be used. In this paper we will give an overview of these methods and will compare them by means of a simulation study.

But let us first review some basic notions we will need in the sequel (for more detailed exposition of these notions see e.g. [3] or [6]). We will consider a two-dimensional (planar) point process X . Let A denote an arbitrary Borel set in \mathbf{R}^2 , then $|A|$ denotes the area of A , $|X \cap A|$ denotes the number of points X has in A . For any given point $u \in \mathbf{R}^2$ let du be the infinitesimal region that contains the point u . Following [4] we can define the (first-order) intensity function λ of X as

$$\lambda(u) = \lim_{|du| \rightarrow 0} \left(\frac{E[X \cap du]}{|du|} \right),$$

i.e. the mean number of points from X occurring in du ; and the second-order intensity function $\lambda^{(2)}(u, v)$ as

$$\lambda^{(2)}(u, v) = \lim_{|du|, |dv| \rightarrow 0} \left(\frac{E[(X \cap du)(X \cap dv)]}{|du||dv|} \right).$$

When X is simple (it does not have multiple points) then the intuitive meaning of $\lambda^{(2)}(\cdot, \cdot)$ is that $\lambda^{(2)}(u, v)|du||dv|$ is the approximate probability that du and dv each contains a point from X , where $u \neq v$. If X is stationary, then $\lambda(u) = \lambda = \text{const}$ and $\lambda^{(2)}(u, v) = \lambda^{(2)}(0, u - v)$ which we will denote just by $\lambda^{(2)}(u - v)$.

There are two important summary characteristics for stationary point processes. The so-called pair correlation function (pcf) (sometimes called simply the g -function) and the K -function which are defined as follows

$$g(u, v) = g(u - v) = \frac{\lambda^{(2)}(u, v)}{\lambda^2}, \quad K(r) = \int_{\|u\| < r} \lambda^{(2)}(u) du / \lambda^2, \quad r > 0.$$

The interpretation of the K -function is the mean number of points of X observed in a ball of radius r centered in point of X normalized by $\frac{1}{\lambda}$. Thus for a homogeneous Poisson point process $K(r) = 2\pi r^2$.

2 Inhomogeneous Neyman-Scott processes

Consider a planar homogeneous Poisson process whose first-order intensity is equal to μ . Each event of the process is considered as a parent which will in turn generate a Poisson number of offspring with an expected value equal to ν . Conditional on the location of a parent, the offsprings are dispersed independently following some common probability density function (pdf) f . For any location u , let $z(u)$ be a vector of covariates that are recorded in that location. An offspring at u is retained with a probability $\exp(z(u)\beta^T)/M$,

where $M = \max(\exp(z(u)\beta^T))$ and β is vector of unknown (inhomogeneity) parameters. The resulting offspring process is then called an inhomogeneous Neymann-Scott point process.

We see that the covariates (which often reflect conditions of the local environment) control the survival rate of the offspring in a given location. Or from another point of view: the initial homogeneous cluster point process is thinned by some thinning probabilities which depend on the covariates z ([10]). This method of introducing inhomogeneity into the point process results in point process belonging into the class of second-order intensity reweighted (SOIR) point processes first introduced in [1]. We omit here the exact definition, but remind the important property of these processes – all the higher order intensity functions $\lambda^{(k)}, k \geq 2$ can be written as

$$\lambda^{(k)}(u_1, \dots, u_k) = \left(\prod_{i=1}^k \lambda(u_i) \right) \lambda^*(u_1, \dots, u_k),$$

where λ^* is invariant with respect to simultaneous translation of all its arguments u_1, \dots, u_k i.e. it is equal to the k -th order intensity function of some homogeneous point process. Thus in the SOIR processes the interactions from the homogeneous model are not destroyed and all the higher order intensity functions are just reweighted by the same functions like the first order intensity function. This implies that we can also define inhomogeneous versions of the g and K functions [1] by

$$g_i(u, v) = g_i(u - v) = \frac{\lambda^{(2)}(u, v)}{\lambda(u)\lambda(v)}, \quad K_i(r) = \int_{\|u\| < r} g_i(u) du, \quad r > 0, \quad (1)$$

which makes it possible to adapt the estimation procedures from the homogeneous case also to these inhomogeneous Neyman-Scott processes.

The first- and second-order intensity functions of the Neyman-Scott processes are given as follows

$$\lambda(u) = \mu \nu \exp(z(u)\beta^T) / M, \quad (2)$$

$$\lambda^{(2)}(u, v) = \lambda(u)\lambda(v)g_i(u - v), \quad (3)$$

where $g_i(\cdot)$ depends only on the parameter μ and the (possibly vector) parameter σ parametrizing the dispersion pdf f . We will denote $\theta = (\mu, \sigma)$ and call θ the inhomogeneity parameter (or the second-order parameter, since it affects only the second order structure of the process). For example, if the pdf for an offspring location relative to its parent is a bivariate radially symmetric normal distribution (e.g. [4]), then

$$g_i(u - v) = g_\theta(u - v) = 1 + \exp(-\|u - v\|^2 / (4\sigma^2)) / (4\pi\sigma^2\mu), \quad (4)$$

where g_θ denotes the dependence of the pcf on the parameter θ and $\|\cdot\|$ denotes the Euclidean norm.

3 Parameter estimation procedures

Let us now review the available estimation procedures for inhomogeneous Neyman-Scott processes. All of these procedures are 2-step procedures, where in the first step the first-order (or inhomogeneity) parameters are estimated and then conditionally on them the second-order (or interaction) parameters are estimated in the second step.

Let us denote $\beta_0 = \log(\mu\nu/M)$. Then (2) changes into $\lambda(u) = \lambda_\beta(u) = \exp(\beta_0 + z(u)\beta^T)$ and is parametrized by $\beta = (\beta_0, \beta_1, \dots, \beta_k)$ and (3) becomes $\lambda_{\beta,\theta}^{(2)}(u, v) = \lambda_\beta(u)\lambda_\beta(v)g_\theta(u - v)$, where the values of $g_\theta(u - v)$ are independent of the parameter β .

Following [10] we estimate the inhomogeneity parameter β by an estimation equation approach. Let W denote the observation window in which the realization of X is observed. The estimator $\hat{\beta}$ is defined as the solution of

$$u(\beta) = \sum_{x \in X \cap W} (\beta_0 + z(x)\beta^T) - \int_W \exp(\beta_0 + z(u)\beta^T) du = 0. \quad (5)$$

Note that $u(\beta)$ is actually the derivative of the log-likelihood of an inhomogeneous Poisson process \tilde{X} observed on W with intensity function $\lambda_\beta(u)$, thus the estimate of the first-order parameter β is obtained by ignoring the interactions in the point process X .

In the sequel let us denote $\hat{\lambda}(u) = \lambda_{\hat{\beta}}(u)$ the intensity function $\lambda_\beta(u)$ where the estimate $\hat{\beta}$ is plugged in - i.e. this is the estimate of the intensity function $\lambda(u)$ obtained by the first step estimation.

3.1 Minimum contrast estimation

This is the method for the estimation of the interaction parameter θ which was used in the original paper [10]. The minimum contrast estimation is a classical procedure in spatial statistics - see e.g. [4].

The interaction parameter θ is estimated in the second step as the argument minimizing the contrast function

$$\int_0^R \left((\hat{K}(r))^c - (K_i(r; \theta))^c \right)^2 dr. \quad (6)$$

Here $K_i(r; \theta)$ is the inhomogeneous K -function (1) and $\hat{K}(r)$ is the estimate of this statistic from the observed data

$$\hat{K}(r) = \sum_{x, y \in X \cap W}^{\neq} \frac{I(\|x - y\| \leq r)}{\hat{\lambda}(x)\hat{\lambda}(y)|(W - x) \cap (W - y)|}, \quad (7)$$

where \neq over the sum denotes that we consider only pairs of distinct points $x \neq y$ from X , $I(\cdot)$ is the indicator function and $\hat{\lambda} = \lambda_{\hat{\beta}}$ is the parametric

estimate of the intensity function from the first step. The two parameters R and c are tuning parameters and it is suggested in the literature ([4]) to choose $c = 0.25$ for attractive point patterns, but there are no recommendations for the choice of R . The performance of the estimate can depend strongly on the chosen value of R as we will see in the simulation study.

3.2 Composite likelihood estimation

In [5] the composite likelihood method was introduced for fitting homogeneous point process models. The composite likelihood is defined by

$$CL = \sum_{\substack{x \neq y \in X \cap W, \\ \|x-y\| < R}} \left[\log \lambda^{(2)}(y-x) - \log \left(\int_W \int_W \lambda^{(2)}(u-v) I(\|u-v\| < R) dudv \right) \right], \tag{8}$$

and the estimate of the interaction parameter is the argument of maxima of the composite likelihood function. The parameter R is a user specified constant. We can define a two-step estimation procedure for estimation of inhomogeneous point patterns by first estimating the inhomogeneity parameter β from the estimation equation (5), thus obtaining the estimate of the intensity function $\hat{\lambda} = \lambda_\beta$. Then in the second step we plug-in $\hat{\lambda}$ into the equation for composite likelihood (8) and obtain

$$CL(\theta) = \sum_{x \neq y \in X \cap W, \|x-y\| < R} \left[\log(\hat{\lambda}(x)\hat{\lambda}(y)g_\theta(y-x)) - \log \left(\int_W \int_W \hat{\lambda}(u)\hat{\lambda}(v)g_\theta(u-v)I(\|u-v\| < R)dudv \right) \right],$$

where $g_\theta(\cdot, \cdot)$ is the inhomogeneous g-function. By maximizing $CL(\theta)$ with respect to θ we obtain the composite likelihood estimate of the inhomogeneity parameter θ .

3.3 Palm likelihood estimation

For homogeneous point processes this method was introduced in [9]. It was generalized to inhomogeneous point processes in [8]. It is again a 2-step estimation procedure where in the first step the inhomogeneity parameter β is estimated from the estimation equation (5). The so-called Palm likelihood L_P is then given by

$$\log L_P = \sum_{x \neq y \in X \cap W, \|x-y\| < R} \log \left(\hat{\lambda}(x)\hat{\lambda}(y)g_\theta(y-x) \right) - \int_{B(0,R)} \int_{W \cap (W-u)} \hat{\lambda}(x)\hat{\lambda}(x+u)g_\theta(u)dxdu,$$

where $B(0, R)$ denotes a ball centered in the origin with radius R and R is a tuning constant. The MPL estimate of the interaction parameter θ is the value which maximizes $\log L_P = \log L_P(\theta)$.

4 Simulation study

To compare the performance of the three estimation procedures introduced in Section 3 when they are used for the estimation of the second-order parameters of middle sized data we applied them to an exponentially inhomogeneous Neyman-Scott process on the unit square W . Thus the thinning function was of the form $\exp(x_1\beta_1)$ for $x = (x_1, x_2) \in W$ and the intensity function $\lambda_\beta(\cdot)$ was $\lambda_\beta(x) = \exp(\beta_0 + \beta_1 x_1)$, $x \in W$, where $\beta = (\beta_0, \beta_1)$ is the inhomogeneity parameter. We have chosen the pdf f to be a bivariate radially symmetric normal distribution like at the end of Section 2 implying the pcf of X to be given by (4) with the interaction parameter $\theta = (\mu, \sigma)$. Let us denote $\Lambda = \int_W \lambda_\beta(x) dx$, the total intensity of X on W .

We have fixed $\beta_1 = 1$ and to investigate how the behaviour of the estimators depends on the values of the interaction parameters, we have simulated patterns with $\mu = 12.5, 25$ for $\Lambda = 100$ and $\mu = 25, 50$ for $\Lambda = 200$ and for each of them we used three different cluster spread parameters $\sigma = 0.01, 0.02, 0.04$. 500 replications from each model were simulated and the three different estimation procedures were applied to the same set of simulated patterns. The relative mean squared errors (i.e. divided by the square of the true parameter value) for the estimates of σ and μ are given in the tables below. Because of limited space we omit the results for $\Lambda = 200, \mu = 25$ since it showed very similar behaviour to the case $\Lambda = 200, \mu = 50$.

In the first two tables values for the CL and MPL estimators $\hat{\mu}$ and $\hat{\sigma}$ (first or second half of the table, respectively) are given. To investigate the influence of the tuning parameter R we have used three different values $R = 0.2, 0.3, 0.4$. This is indicated in the second column of the tables.

CL		$\lambda = 100, \mu = 12.5$			$\lambda = 100, \mu = 25$			$\lambda = 200, \mu = 50$		
	R	0.01	0.02	0.04	0.01	0.02	0.04	0.01	0.02	0.04
$\hat{\mu}$	0.2	.218	.303	.476	.145	.185	.380	.071	.136	.474
	0.3	.164	.235	.440	.112	.162	.421	.061	.127	.724
	0.4	.146	.225	.396	.105	.163	.425	.059	.122	.719
$\hat{\sigma}$	0.2	.041	.029	.029	.028	.027	.038	.017	.020	.033
	0.3	.039	.030	.042	.029	.026	.042	.013	.020	.042
	0.4	.038	.030	.046	.027	.026	.047	.014	.021	.050

MPL		$\lambda = 100, \mu = 12.5$			$\lambda = 100, \mu = 25$			$\lambda = 200, \mu = 50$		
	R	0.01	0.02	0.04	0.01	0.02	0.04	0.01	0.02	0.04
$\hat{\mu}$	0.2	.123	.193	.491	.082	.142	.524	.046	.109	.605
	0.3	.123	.193	.490	.081	.142	.522	.046	.109	.590
	0.4	.124	.193	.490	.079	.140	.520	.046	.109	.590
$\hat{\sigma}$	0.2	.020	.023	.041	.027	.022	.043	.012	.018	.047
	0.3	.019	.023	.047	.026	.022	.044	.012	.018	.053
	0.4	.019	.023	.048	.026	.023	.044	.012	.015	.059

In the third table values of the relative MSE for the minimum contrast estimator using the K-function are given. Here also, we used different values

of the tuning parameter R , now given as a multiple of the true value of the interaction parameter σ (as is done in the literature [4]). The MC estimates were computed using the procedure `kppm` from the R package `Spatstat`, for estimation of inhomogeneous Thomas processes. This procedure allows the user to choose the tuning parameter R , but it can find a sensible default by itself (when it is not specified by the user). This automated estimate is denoted by S in the table.

MC		$\lambda = 100, \mu = 12.5$			$\lambda = 100, \mu = 25$			$\lambda = 200, \mu = 50$		
	R	0.01	0.02	0.04	0.01	0.02	0.04	0.01	0.02	0.04
$\hat{\mu}$	3σ	.115	.155	.250	.128	.128	.523	.049	.085	.392
	4σ	.096	.132	.310	.076	.097	.889	.034	.081	.482
	5σ	.094	.137	.388	.071	.106	1.052	.034	.095	.592
	6σ	.095	.146	.439	.073	.121	1.248	.036	.111	.623
	S	.179	.228	.491	.147	.220	.649	.095	.204	.569
$\hat{\sigma}$	3σ	.495	.308	.127	1.570	1.293	2.121	.044	.077	.964
	4σ	.041	.025	.042	.052	1.662	.533	.016	.020	.058
	5σ	.031	.018	.041	.028	1.029	.092	.012	.017	.054
	6σ	.020	.022	.053	.024	.130	.078	.012	.019	.075
	S	.056	.631	.215	2.050	1.403	4.371	36.836	16.040	14.054

5 Results

From observing the tables we see, that generally the rule: tighter clusters = better estimates, holds. The only exception is the minimum contrast estimation for smaller intensities (i.e. $\Lambda = 100$). We also see, that the number of clusters is more important than the total intensity of the point process. The larger number of the observed clusters, the better the estimates. But of course, this rule ceases to hold, when the number of clusters is so big, that they melt together and are hard to identify (seems this is the case with $\Lambda = 200, \mu = 50$).

When comparing the procedures we see the qualitatively similar behaviour of the CL and MPL estimates, the MPL estimate being slightly more effective. The MC estimator gives better estimates of the parameter μ but worse estimates of the cluster spread parameter σ .

From the table for the MC estimator we can also see that the quality of the estimate varies quite a lot for the different values of the tuning parameter R . Especially obvious is the problem with the possible underestimation of the interaction when choosing too a small value of R or on the other side the danger of “smoothing out” too much interaction when the chosen R is too big. Another interesting observation is that the automated estimator provided by the function `kppm` performs quite well concerning the estimation of μ but when it comes to the estimation of the spread parameter σ it fails completely. Thus when we want to use the MC estimator some preprocessing

and estimation of the approximate range of interaction (e.g. from the empirical pair-correlation function) is advisable to get a usable value of R to be used in the estimation.

Concerning the dependence of the CL and MPL estimators on the tuning constant R , the MPL estimate is less dependent of the choice of R than the CL estimator, both of them being much less dependent than the MC estimator.

References

- [1] Baddeley A.J., Møller J., Waagepetersen R. (2000). *Non- and semi-parametric estimation of interaction in inhomogeneous point patterns*. Statist. Neerlandica **54**, 329–350.
- [2] Baddeley A.J., Turner R. (2005). *spatstat: an R package for analyzing spatial point patterns*. Journal of Statistical Software **12**, 1–42.
- [3] Daley D.J., Vere-Jones D. (1988). *An introduction to the theory of point processes*. Springer Verlag, New York.
- [4] Diggle P.J. (2003). *Statistical analysis of spatial point patterns*. Oxford University Press, New York.
- [5] Guan Y. (2006). *A composite likelihood approach in fitting spatial point process models*. J. Am. Stat. Assoc. **101**, 1502–1512.
- [6] Møller J., Waagepetersen R.P. (2003). *Statistical inference and simulation for spatial point processes*, CRC Press, Boca Raton.
- [7] Møller J., Waagepetersen R.P. (2007). *Modern statistics for spatial point processes*. Scandinavian Journal of Statistics **34**, 643–684.
- [8] Prokešová M., Vedel Jensen E.B. (2008). *Palm likelihood estimation for inhomogeneous cluster processes*. In preparation.
- [9] Tanaka U., Ogata Y., Stoyan D. (2007). *Parameter estimation and model selection for Neymann-Scott point processes*. Biometrical Journal **49**, 1–15.
- [10] Waagepetersen R.P. (2007). *An estimating function approach to inference for inhomogeneous Neyman-Scott processes*. Biometrics **63**, 252–258.
- [11] Waagepetersen R.P., Guan Y. (2008). *Two-step estimation for inhomogeneous spatial point processes*. J. R. Stat. Soc. B, submitted.

Acknowledgement: This work was supported by the grant 201/08/P100 from GAČR and by the research project MSM 0021620839 financed by MŠMT ČR.

Address: M. Prokešová, MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8 – Karlín

E-mail: prokesov@karlin.mff.cuni.cz

NĚKTERÉ MÍRY NEURČITOSTI MĚŘENÍ ZA NEJISTOTY

Zdeněk Půlpán, Michal Čihák, Jaroslava Bomberová

Klíčová slova: Fuzzy množiny, entropie, entropie pro fuzzy množiny.

Abstrakt: Měření za nejistoty má svá omezení, která jsou dána krom jiného i absencí prototypu jednotky měřené vlastnosti. Měření pak nemá tu univerzálnost, kterou má například měření délky nebo hmotnosti. V případě absence „jednotkové vlastnosti“ je nutné zavést „prototypy“ vlastností, se kterými se pak porovnávají vlastnosti měřeného vzorku. Příspěvek nabízí několik možností odhadu neurčitosti a informace takových „měření“.

Abstract: Fuzzy measure of uncertainty under indefinite and vague condition. It is very important to estimate the measure of uncertainty of data in the cases of vague conditions. We recommend some methods with help of fuzzy theory.

1 Úvod

Každá procedura měření je zatížena neurčitostí. Neurčitostí je provázána i interpretace souboru dat získaných měřením. Problém odhadu neurčitosti měření nemá jednoznačné řešení. Adekvátnost určitého formálního popisu je ale možné odhadovat „mírou podobnosti“ jeho vybraných vlastností a vlastností modelovaných. Ukážeme na několika příkladech, jak lze neurčitost odhadovat v souvislosti s měřením za nejistoty a v podmínkách vágnosti.

Studované objekty se snažíme za účelem měření popsat z několika relativně dobře rozlišitelných hledisek (vlastností). Každý objekt tak může být charakterizován uspořádanou k -ticí, jejíž každá složka představuje numericky nebo nenumericky vyjádřené hodnocení, opírající se o definici měření v rámci příslušného hlediska. Je-li například jedno z hledisek hmotnost, využíváme toho, že můžeme mít k dispozici jednotku hmotnosti a těleso libovolné hmotnosti s jednotkou porovnáme; měříme-li však teplotu, nemůžeme předchodí postup aplikovat, nemáme totiž prototyp 1°C . Podobně neexistuje prototyp jednotkové schopnosti, vytrvalosti, inteligence, užítku, znalosti, atd. V případě absence jednotky měření a tomu odpovídající procedury, vytváříme soubor S etalonů, kterými porovnáváme měřený objekt. Je-li soubor etalonů uspořádaným souborem, můžeme pracovat s ordinální stupnicí měřené vlastnosti (hlediska). Tak bylo potřeba vytvořit řadu referenčních bodů teplotní stupnice, které přesně definují jisté teploty na základě fyzikálních vlastností některých látek (vody, rtuti, ...) mezi kterými se další teplotní hodnoty získávají interpolací. Pak 1°C odpovídá $\frac{1}{B-A}$ rozdílů určité úrovně některé měřitelné vlastnosti fyzikálního jevu, měnícího se mezi stavy danými definicí např. pro $B^\circ\text{C}$ a $A^\circ\text{C}$, přičemž $B > A$.

Jakou informaci o sledovaném objektu z měření k vlastností můžeme získat? Na uvedenou otázku aspoň částečně odpovíme, když posoudíme mechanismy hodnocení ve všech hlediscích. Soustředíme se jen na odhady těch vlastností, jejichž úroveň jsme schopni posoudit na základě porovnání s předem připraveným systémem etalonů. Východiskem pro odhad informace je neurčitost. Budeme se proto zabývat odhady neurčitosti na konečných množinách, informaci pak definujeme jako rozdíl dvou neurčitostí.

Pro diskrétní rozložení pravděpodobnosti se již v 60. letech 20. století ustálily základní požadavky na entropii a informaci ve formě Chinčín-Faddějevových axiomů [9], [10]. Přibližně v téže době se Ingarden a Urbanik pokusili zavést pojem informace bez pravděpodobnosti [11] a pro potřeby humanitních věd se formulovaly principy sémantické informace [2]. Základní výsledky teorie informace z 80. let 20. století jsou shrnuty v rozsáhlé monografii I. Vajdy [12]. Rozvoj fuzzy matematiky umožnil rozšířit pojem informace i pro fuzzy množiny [13], [14] a syntetizovat je do specifických požadavků [1]. V současné době fuzzy matematika nabízí velké množství variant modelů zpracování dat vágní povahy. Některé z nich zde nabízíme.

2 Neurčitost obsažená ve fuzzy množině

Předpokládejme, že máme připravenou definici úplného souboru škálových hodnot S_h pro měření, vztahujícímu se k určité vlastnosti h , například souborem n etalonů

$$S_h = \{s_1, s_2, \dots, s_n\}. \quad (1)$$

Úkolem měření je přiřazení některé škálové hodnoty $s \in S_h$ vybranému objektu A . Vycházíme z představy, že takové přiřazování je obecně nejednoznačné a můžeme ho proto charakterizovat fuzzy množinou

$$A(h) = \{s_i | \mu_A(s_i); i = 1, 2, \dots, n\}, \quad 0 \leq \mu_A(s_i) \leq 1. \quad (2)$$

Číslem $\mu_A(s_i)$ vyjadřujeme míru rozhodnutelnosti (nebo nerozhodnutelnosti) volby škálové hodnoty s_i na základě aplikace standardizované¹ procedury měření. V situaci měření je někdy rozumné předpokládat, že existuje jediná hodnota s_{i_0} taková, že $\mu_A(s_{i_0}) > \mu_A(s_i)$ pro všechna $i \neq i_0$.

Označme systém všech fuzzy množin nad S_h znakem $F(S_h)$. Pro každou fuzzy množinu $X = \{s_i | \mu_i; 0 \leq \mu_i \leq 1, i = 1, 2, \dots, n\}$ označíme znakem $|X|$ součet $|X| = \sum_{i=1}^n \mu_i$. Entropií $H_1(A(h))$ příslušnou fuzzy množině $A(h) \in F(S_h)$ nazveme hodnotu vypočtenou ze vztahu

$$H_1(A(h)) = \frac{|A(h) \cap_T A^c(h)|}{|A(h) \cup_S A^c(h)|}, \quad (3)$$

¹Procedura měření je standardizována, když je soubor škálových hodnot S_h stanoven s přiměřenou přesností například normou.

kde $A^c(h) = \{s_i | 1 - \mu_A(s_i); i = 1, 2, \dots, n\}$. Operace \cap_T , resp. \cup_S představuje některou z t -norm, resp. k ní duální t -konormu, často například užíváme klasické Zadehovy operace \cap a \cup

$$A(h) \cap A^c(h) = \{s_i | \min(\mu_A(s_i), 1 - \mu_A(s_i)); i = 1, 2, \dots, n\}, \quad (4)$$

$$A(h) \cup A^c(h) = \{s_i | \max(\mu_A(s_i), 1 - \mu_A(s_i)); i = 1, 2, \dots, n\}. \quad (5)$$

Pro entropii definovanou vztahem (3) platí $0 \leq H_1(A(h)) \leq 1$, bez ohledu na to, jakou t -normu a k ní duální t -konormu použijeme.

Nejvyšší hodnoty nabývá fuzzy entropie (3) pro fuzzy množiny, jejichž $\mu(s_i) = 0,5$ pro všechna i . Naopak nejmenší hodnoty dosahuje (3) když pro každé i je $\mu(s_i)$ buď rovno 1 nebo 0. Chceme-li pak (3) užívat k odhadu neurčitosti měření podle hlediska h prostřednictvím $A(h)$, musí funkce příslušnosti $\mu(s_i)$ hodnotou 0,5 vyjadřovat případnou nejvyšší nejistotu v určení škálové hodnoty s_i jako jeden z výsledků procesu měření (a hodnoty 0 a 1 pak jako nejvyšší možnou jistotu).

Označme nyní $a_i = \mu_A(s_i)$, $i = 1, 2, \dots, n$, a transformujeme hodnoty a_i funkcí $g: [0; 1] \rightarrow [0; 1]$ tvaru

$$g(x) = \begin{cases} \frac{x}{1-x} & \text{pro } 0 \leq x \leq 1/2, \\ \frac{1-x}{x} & \text{pro } 1/2 < x \leq 1. \end{cases}$$

Novou mírou neurčitosti pak bude g -normalizovaná entropie $H_2(A(h))$ pro $A(h) \in F(S_h)$

$$H_2(A(h)) = \frac{1}{n} \sum_{i=1}^n g(a_i). \quad (6)$$

Zřejmě je také $0 \leq H_2(A(h)) \leq 1$. Největší a nejmenší hodnoty dosahuje (6) za stejných podmínek jako (3).

Jsou-li $A, B \in F(S_h)$ a $A \subset B$ (tj. $\mu_A(s_i) \leq \mu_B(s_i)$, $i = 1, 2, \dots, n$), pak

$$H_2(A \cup B) = \frac{1}{n} \sum_{i=1}^n g(\max[\mu_A(s_i), \mu_B(s_i)]) = \frac{1}{n} \sum_{i=1}^n g(\mu_B(s_i)) = H_2(B), \quad (7)$$

Je-li $A \cap B = 0$, $A, B \in F(S_h)$, pak

$$H_2(A \cup B) = H_2(A) + H_2(B). \quad (8)$$

Toto jsou velmi důležité vlastnosti, které nemá entropie (3).

De Luca a Termini v [1] byli při volbě míry neurčitosti inspirováni Shannonovou entropií; zde je však třeba se vyhnout statistické interpretaci tohoto pojmu. Pro každou fuzzy množinu $A(h) \in F(S_h)$ definovali nejprve nezáporné číslo $H_3(A(h))$ vztahem

$$H_3(A(h)) = -K \sum_{i=1}^n \mu_A(s_i) \ln \mu_A(s_i), \quad K > 0. \quad (9)$$

Užijeme-li již zmíněných Zadehových interpretací pro operace \cap a \cup mezi fuzzy množinami, pak pro H_3 z (9) a $A, B \in F(S_h)$ platí [2], [6]

$$H_3(A \cup B) + H_3(A \cap B) = H_3(A) + H_3(B). \quad (10)$$

Zavedeme-li direktní součin $A \otimes B$ fuzzy množin $A, B \in F(S_h)$ jako fuzzy množinu C na $F(S_h \times S_h)$, přičemž $\mu_C(x, y) = \mu_A(x)\mu_B(y)$, $x, y \in S_h$, pak platí pro H_3 z (9) následující aditivita [1], [6]

$$H_3(A \otimes B) = |A| \cdot H_3(B) + |B| \cdot H_3(A). \quad (11)$$

Jsou-li funkce příslušnosti fuzzy množin A a B normovány tak, že $|A| = |B| = 1$, přejde předchozí vztah v jednodušší

$$H_3(A \otimes B) = H_3(A) + H_3(B), \quad (12)$$

který nám připomíná obdobný vztah pro Shannonovu entropii dvou sdružených a nezávislých náhodných veličin.

De Luca a Termini navrhli míru neurčitosti na základě (9) pro $A(h) \in F(S_h)$ vztahem

$$H_4(A(h)) = H_3(A(h)) + H_3(A^c(h)). \quad (13)$$

Užitím funkce $h(x) = -x \log_2 x - (1-x) \log_2(1-x)$ pro $x \in (0; 1)$, $h(0) = h(1) = 0$, můžeme (13) přepsat do tvaru $H_4(A(h)) = K_1 \sum_{i=1}^n h(a_i)$, kde $K_1 > 0$, $a_i = \mu_A(s_i)$, $i = 1, 2, \dots, n$. Volíme-li $K_1 = 1/n$, dostáváme pro H_4 vyjádření

$$H_4(A(h)) = \frac{1}{n} \sum_{i=1}^n h(\mu_A(s_i)). \quad (14)$$

Proto také

$$0 \leq H_4(A(h)) \leq 1. \quad (15)$$

Uvážíme-li monotónnost funkce $h(x)$ v $[0; 1/2]$ a v $(1/2; 1]$ a to, že maxima nabývá v bodě $1/2$, vidíme, že pak (14) nabývá maxima když $\mu_A(s_i) = 0,5$ pro všechna i , minima když $\mu_A(s_i)$ je buď 0 nebo 1 . A to jsou vlastnosti, které u entropie požadujeme.

Pro stanovení míry neurčitosti pro $A(h) \in S_h$ můžeme také užít výrazu (inspirovaném Kullbackovou mírou v pravděpodobnostních prostorech), vyjadřujícím vztah fuzzy množiny $A(h)$ k pevné referenční množině $C = \{s_i | 0,5; s_i \in S_h\} \in F(S_h)$

$$H_5(A(h)) = 1 - \frac{1}{n} \left[\sum_{0,5 \leq a_i \leq 1} a_i \log_2(2a_i) + \sum_{0 \leq a_i < 0,5} (1-a_i) \log_2(2(1-a_i)) \right], \quad (16)$$

kde $a_i = \mu_A(s_i)$, $i = 1, 2, \dots, n$. Snadno ověříme, že pro každou fuzzy množinu $A(h) \in F(S_h)$ je také

$$0 \leq H_5(A(h)) \leq 1. \quad (17)$$

Největší hodnotu pro H_5 z (16) dostaneme, když $A(h) = C$ a nejmenší pak pro tu fuzzy množinu, která má a_i buď rovno 0 nebo 1 pro všechna i .

Další možnost odhadu neurčitosti, vztahující se k fuzzy množině $A(h) \in F(S_h)$ nabízí vztah

$$H_6(A(h)) = 1 - \frac{1}{n} \sum_{i=1}^n [2 \max(\mu_A(s_i), 1 - \mu_A(s_i)) - 1]. \quad (18)$$

I pro tuto neurčitost platí, že

$$0 \leq H_6(A(h)) \leq 1, \quad (19)$$

přičemž $H_6(C) = 1$ a minimální hodnoty nabývá opět, když $\mu_A(s_i)$ je rovno buď 0 nebo 1 pro všechna i .

Shannonova neurčitost (20), obsažená ve fuzzy množině $A(h) \in F(S_h)$ je analogií k entropii, zavedené C. Shannonem pro rozdělení diskretních náhodných veličin; k tomu je však nutné hodnoty funkce příslušnosti $\mu_A(s_i)$ transformovat na hodnoty $p_i \in [0; 1]$ tak, aby $\sum_{i=1}^n p_i = 1$. Potom definujeme

$$H_7(A(h)) = - \sum_{i=1}^n p_i \log_2 p_i, \quad (20)$$

kde

$$p_i = \mu_A(s_i) \left[\sum_{i=1}^n \mu_A(s_i) \right]^{-1} \quad (21)$$

a dodefinováváme $0 \cdot \log_2 0 = 0$.

Protože

$$0 \leq H_7(A(h)) \leq \log_2 n, \quad (22)$$

je možné uvedenou neurčitost normovat tak, aby platilo $H_7(A(h)) \in [0; 1]$.

Uveďme nyní některé příklady aplikace předložené teorie.

Příklad 1: Pro fuzzy množinu $A(h) = \{s_1|0,7; s_2|0,4\}$ dostáváme z (3), (6), (14), (16), (18) a (20) postupně $H_1 \approx 0,54$; $H_2 \approx 0,54$; $H_4 \approx 0,93$; $H_5 \approx 0,75$; $H_6 \approx 0,70$; $H_7 \approx 0,95$. Z uvedených hodnot můžeme soudit, že výrazy (14) a (20) dávají hodnoty velmi citlivé na větší intuitivně chápanou neurčitost. \triangle

Příklad 2: Sledujme neurčitost v situaci, kdy se má po vyšetření pacienta rozhodnout, kterému ze specialistů bude pacient předán k další léčbě; po předání pacienta do rukou specialisty je třeba chorobu přesněji diagnostikovat a navrhnout léčbu. Předpokládejme, že diagnostický postup se skládá nejprve z výběru vhodného hlediska, tj. z určení prvku $h \in \mathcal{H} = \{h_1, h_2, \dots, h_k\}$, kde \mathcal{H} je soubor všech relevantních hledisek. Pak jde o stanovení diagnózy s z množiny možných diagnóz S_h . Celková neurčitost rozhodnutí o jisté diagnóze se skládá ze dvou částí: a) z neurčitosti volby prvku $h \in \mathcal{H}$, označíme

ji H^a ; b) z neurčitosti stanovení diagnózy $s \in S_h$, označíme ji $H^{b|a}$. Neurčitost výsledného rozhodování H můžeme odhadovat na základě charakterizace obou postupů fuzzy množinami jako jejich součet

$$H = H^a + H^{b|a}, \quad (23)$$

přičemž druhá neurčitost je podmíněna volbou určitého hlediska $h \in \mathcal{H}$.

Použijeme-li k odhadu neurčitostí některého ze vztahů (3), (6), (14), (18) nebo (20), bude celková neurčitost $H \in [0; 2]$ (ovšem za předpokladu stejné váhy obou rozhodovacích kroků).

Poznámka: Každý ze vztahů pro entropii je definicí metody i jednotky. Proto ve vztahu (23) musíme v obou výrazech pro H^a i $H^{b|a}$ používat stejného definičního vztahu pro entropii. Informaci $I(h)$ získanou z diagnózy užitím hlediska h pak odhadneme z rozdílu

$$I(h) = 2 - H; \quad I(h) \in [0; 2]. \quad (24)$$

Příklad 3: V dotaznících zpracovávajících názory respondentů posuzujeme rozložení četností odpovědí na diskretní ordinální škále s jistým středem x_0 . Škála je konstruována tak, že nejmenší, resp. největší, škálová hodnota představuje vždy maximální naplnění jednoho ze dvou pólů zkoumaného jevu. Informace I o tendenci zkoumaného jevu (k maximu naplnění některého z pólů) je obsažena v četnosti volby škálových hodnot větších než x_0 vzhledem k četnosti volby hodnot menších než x_0 a může být odhadována z rozdílu

$$I \approx 1 - h(\hat{p}); \quad I \in [0; 1], \quad (25)$$

kde $\hat{p} = n^+ / (n^+ + n^-)$ a n^+ , resp. n^- je četnost volby škálové hodnoty větší nebo rovné x_0 , resp. menší než x_0 . Čím je ukazatel I bližší 1, tím jednoznačnější je rozhodování „průměrného“ respondenta. Stupnice informace vzhledem k \hat{p} není rovnoměrná, vyjadřuje dobře intuitivně chápaný vztah mezi \hat{p} a informací I o tendenci odpovědi. \triangle

Jak jsme již viděli, fuzzy entropii $H(A(h))$, $A(h) \in F(S_h)$, můžeme definovat různým způsobem. Obecně se však požaduje, aby $H(A(h))$ splňovala co nejvíce z následujících podmínek:

- a) $H(A(h)) = 0$, právě když fuzzy množina $A(h)$ reprezentuje „ostrou“ množinu, tj. když $\mu(s_i) = 0$ nebo 1 pro $i = 1, 2, \dots, n$;
- b) $H(A(h))$ nabývá maxima, právě když fuzzy množina $A(h)$ má „nejmenší ostrost“, tj. když $\mu(s_i) = 0,5$ pro $i = 1, 2, \dots, n$;
- c) hodnoty $H(A(h))$ rostou s růstem „neostroti“ fuzzy množiny $A(h)$, tj. $H(A'(h)) \geq H(A(h))$ když pro $\mu_{A'}$ platí: je-li $\mu_{A'}(s_i) \leq 0,5$, pak $\mu_A(s_i) \leq \mu_{A'}(s_i)$, je-li $\mu_{A'}(s_i) > 0,5$, pak $\mu_A(s_i) > \mu_{A'}(s_i)$, $s_i \in S_h$;
- d) pro fuzzy množinu $A^c(h) \in F(S_h)$, jejíž funkce příslušnosti splňuje podmínku $\mu_{A^c}(s_i) = 1 - \mu_A(s_i)$, $s_i \in S_h$, platí

$$H(A^c(h)) = H(A(h)); \quad (26)$$

e) jsou-li $A, B \in F(S_h)$, pak

$$H(A \cup B) = H(A) + H(B) - H(A \cap B), \quad (27)$$

kde $\mu_{A \cup B}(s) = \max_{s \in S_h}(\mu_A(s), \mu_B(s))$, $\mu_{A \cap B}(s) = \min_{s \in S_h}(\mu_A(s), \mu_B(s))$.

Jestliže entropie $H(A(h))$, $A(h) \in F(S_h)$ splňuje všechny podmínky a) až e), nazývá se „dobrou mírou“ neurčitosti (podobně pak z ní odvozenou informaci nazýváme „dobrou mírou“ informace).

V práci [7] je uveden příklad parametrizované entropie pro fuzzy množinu $A(h) \in F(S_h)$ ve tvaru

$$H_8^\beta(A(h)) = K A_\beta \sum_{i=1}^n \left[\mu_A^\beta(s_i) + (1 - \mu_A(s_i))^\beta - 1 \right], \quad (28)$$

kde $\beta > 0$, $\beta \neq 1$, $A_\beta = (2^{1-\beta} - 1)^{-1}$ a $K > 0$ je normalizační konstanta (β je vhodný parametr, který specifikuje vlastnosti entropie). Entropie (28) má pro každé $A(h) \in F(S_h)$ a libovolný přípustný parametr β vlastnosti a) až e). Volbou normalizační konstanty K určujeme jednotku entropie. Je výhodné spojovat volbu konstanty K s určitou interpretací stupnice neurčitosti. Ukážeme to na následujícím příkladě.

Příklad 4: Je-li $n = 3$ a fuzzy množina $A(h) = \{s_1|1; s_2|0,3; s_3|0\} \in F(S_h)$, je

$$H_8^{0,5}(A(h)) = K(\sqrt{2} + 1) \left[0 + (\sqrt{0,3} + \sqrt{0,7} - 1) + 0 \right] \approx 0,928K.$$

Má-li být pro fuzzy množinu $B(h) = \{s_i|0,5; i = 1, 2, 3\} \in F(S_h)$ hodnota největší neurčitosti rovna 1, musí platit

$$H_8^{0,5}(B(h)) = K(\sqrt{2} + 1) \cdot 3 \left[(\sqrt{0,5} + \sqrt{0,5} - 1) \right] = 3K = 1,$$

takže volíme $K = 1/3$. Pro neurčitost v uvedených jednotkách pak máme $H_8^{0,5}(B(h)) \approx 0,31$. \triangle

Informaci můžeme určit odečtením příslušné entropie od její maximální možné hodnoty. Tak například pro entropii definovanou vztahem (28) se příslušná informace určí ze vztahu

$$I_8^\beta(A(h)) = H_{8,\max}^\beta - H_8^\beta(A(h)) = nK - H_8^\beta(A(h)), \quad (29)$$

kde $H_{8,\max}^\beta$ je maximální hodnota entropie definované vztahem (28).

3 Odhad znalosti z neurčitosti pedagogického měření

Zjišťování znalostí je příkladem měření zatíženého velkou neurčitostí. Měřícím prostředkem je zde většinou objektivizovaný validní dotazník (test).

Ten je rozdělen na jednotlivé sady spolu významově souvisejících položek. Na položky respondent reaguje odpovědí, která může být skórovatelná buď dichotomicky (0–1, resp. ANO–NE, apod.) nebo pomocí vícestupňové škály (0–1–2, resp. špatně–částečně–správně nebo 0–1–2–3–4 s podrobnější interpretací). V našem případě měření znalostí (pro jednoduchost) předpokládáme, že neurčitosti související s dotazníkem a jeho skórováním jsou vzhledem k respondentům i v čase konstantní. Pak míru znalostí respondenta posuzujeme ze znalostní funkce Z , kterou určíme z příslušné míry neurčitosti nejdříve pro j -tou dotazníkovou položku. Předpokládejme, že skórování bylo realizováno tak, že každé z možných škálových hodnot byla expertně přiřazena míra neurčitosti adekvátního vyjádření úrovně odpovědi na příslušnou položku. Jsou-li $0, 1, 2, \dots, k$, $k \geq 1$, škálové hodnoty, pak výsledek hodnocení experta je možné zaznamenat jako konvexní fuzzy množinu $A_j = \{i | \mu_i; i = 1, 2, \dots, k\}$, $j = 1, 2, \dots, n$, kde n je počet dotazníkových položek v dané sadě a μ_i jsou míry souhlasu hodnotitele s i -tou škálovou hodnotou, $\mu_i \in [0; 1]$ ($\mu_i = 0$ znamená úplný nesouhlas, $\mu_i = 1$ plný souhlas, hodnoty $\mu_i \in (0; 1)$ míru váhání mezi oběma krajními mezemi). Z entropií H pro fuzzy množiny A_j (viz vztahy (3), (6), (9), (14), (16), (18), (20), (28)) stanovíme míru znalosti Z_j podle [6] takto:

$$Z_j = \begin{cases} 1 - \frac{H(A_j)}{2}, & \text{je-li } \sum_{i \geq x_0} \mu_i \geq \frac{1}{2} |A_j|, \\ \frac{H(A_j)}{2}, & \text{je-li } \sum_{i \geq x_0} \mu_i < \frac{1}{2} |A_j|, \end{cases} \quad (30)$$

kde $x_0 = k/2$. Celkovou znalost Z respondenta pak určíme jako vážený součet

$$Z = \sum_{j=1}^n \nu_j Z_j, \quad 0 \leq \nu_j \leq 1, \quad \sum_{j=1}^n \nu_j = 1, \quad (31)$$

přičemž ν_j jsou nějak objektivizované² váhy, nebo $\nu_j = 1/n$ v případě, že nemáme prostředek pro objektivizaci vah.

Příklad 6: Použijeme stupnice 0–1–2 pro hodnocení dotazníkových odpovědí. Nechť pro jednu z položek je hodnocení zaznamenáno fuzzy množinou $A = \{0|0,2; 1|0,7; 2|0,9\}$. Pro $H = H_1$ dostaneme z (3) neurčitost $H_1 = 0,25$. Míru znalosti pak počítáme pomocí první varianty (30), protože $\sum_{i \geq 1} \mu_i = 1,6 \geq 0,9 = \frac{1}{2} |A|$ a tedy $Z^1 = 1 - 0,25/2 = 0,875$.

Nechť ale například $A = \{0|1; 1|0,2; 2|0\}$. Z (3) dostaneme $H_1 \approx 0,07$. Míru znalosti pak počítáme pomocí druhé varianty (30), protože $\sum_{i \geq 1} \mu_i = 0,2 < 0,6 = \frac{1}{2} |A|$ a tedy $Z^1 \approx 0,07/2 \approx 0,04$.

Je-li $A = \{0|0; 1|0,2; 2|1\}$, pak je opět $H_1 \approx 0,07$, ale míru znalosti počítáme pomocí první varianty (30), protože $\sum_{i \geq 1} \mu_i = 1,2 < 0,6 = \frac{1}{2} |A|$ a tedy $Z^1 \approx 1 - 0,07/2 \approx 0,97$.

²O objektivizaci vah viz například [14]

Vztah (31) není vztahem ideálním; závisí na konstrukci dotazníku i jeho skórování. Protože za entropii H je možné určit kterýkoliv z dříve diskutovaných vztahů, není volba konkrétní podoby vztahu (30) jediná; je to důsledek toho, že univerzální jednotka znalosti neexistuje. Dodejme ještě, že údaje o skutečných znalostech ukazatel vyčíslený podle (30) podává pouze relativně, nejen vzhledem k volbě neurčitosti, ale i v závislosti na dotazníku a podmínkách jeho zadávání (např. stanovení časových limitů, možnosti užívat jistých pomůcek, atd.) i skórování (záleží na přesnosti interpretace škálových hodnot nebo hodnot funkce příslušnosti dané fuzzy množiny).

Navržená metodika umožňuje zpracovat psychologické nebo medicínské dotazníky [6], [14], nebo jen verbálně formulovatelné výsledky šetření [13]. První z autorů tuto metodiku také s úspěchem používá k hodnocení výsledků své výuky. Na rozdíl od klasické metodiky (například 0–1 hodnocení položek) se zdá být uvedená metodika objektivnější (ztrácí se méně informace – podrobněji viz [8]).

Literatura

- [1] De Luca A., Termini S. (1972). *A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory*. Information and Control **20**.
- [2] Hintikka J. (1970). *On semantic information*. In: Hintikka, J., Suppes, P. (eds.) Information and Inference, D. Reidel Publ. Comp., Dordrecht, 1970.
- [3] Klir H. J. (1987). *Where do we stand on measures of uncertainty, ambiguity, fuzziness, and the like?* Fuzzy Sets and Systems **24**.
- [4] Klir G. J., Fogler T. A. (1988). *Fuzzy sets, Uncertainty and Information*, Prentice Hall.
- [5] Půlpán Z. (1990). *Fuzzy – pragmatische information*. GrKg/Human-kybernetik, **31**, 3.
- [6] Půlpán Z. (2001). *K problematice hledání podstatného v humanitních vědách*. Academia, Praha.
- [7] Arora H., Petry F., Beaubouef T. (1997). *New information measures for fuzzy sets*. IFSA'97 Prague, **4**, 75–78.
- [8] Půlpán Z. (2006). *Ztráty informace v důsledku restrikce měřící škály*, UPOL, Olomouc.
- [9] Rényi A. (1965). *On the foundation of information theory*. Review of the International Statistical Institute, **33**, 1, 1965.
- [10] Faddeev D. K. (1957). *Zum Begriff der Entropie in der Wahrscheinlichkeitschemas*. Arbeiten zur Informationstheorie, I, Deutscher Verlag der Wissenschaften, Berlin.
- [11] Ingarden R. S., Urbanik K. (1962). *Information without probability*. Colloquium Mathematicum, **9**, 131–150.
- [12] Vajda I. (1982). *Teória informácie a statistického rozhodovania*. Alfa, Bratislava.

- [13] Půlpán Z. (2000). *Informace pro fuzzy množiny v diagnostice nemocí*, Acta Medica, Supplementum, Hradec Králové **43**, 1, 7–12.
- [14] Půlpán Z. (1997). *K problematice vágnosti v humanitních vědách*, Academia, Praha.

Poděkování: Práce byla podpořena projektem Specifického výzkumu: Teorie zpracování didaktického a psychologického experimentu PdF UHK č. 2118.

Adresa: Z. Půlpán, M. Čihák, J. Bomberová, Pedagogická fakulta UHK, katedra matematiky, Rokitanského 62, 500 03 Hradec Králové

E-mail: zdenek.pulpan@uhk.cz, michal.cihak@uhk.cz

DETERMINANTY PZI VE ZPRACOVATELSKÉM PRŮMYSLU ČR V LETECH 2000-2006

Eva Ryšavá

Klíčová slova: PZI, přímé zahraniční investice, zpracovatelský průmysl.

Abstrakt: Je důležité porozumět chování zahraničních investorů a tomu, co hraje roli při jejich rozhodování, kam by měli umístit své investice. Cílem příspěvku je najít ekonometrický model determinant přímých zahraničních investic (PZI) ve zpracovatelském průmyslu v ČR v letech 2000–2006. Model zahrnuje několik základních ekonomických veličin (práce, kapitál, R&D, zisk a další). Vedle jednoduchých technik odhadu (OLS, fixní efekty) je použit odhad zobecněnou momentovou metodou (GMM).

Abstract: It is important to understand the decision of foreign investors where to place their investments. The aim of this paper is to find an econometric model describing determinants of foreign direct investment (FDI) in manufacturing industry of the CR between 2000–2006. Our model includes several basic economic variables (labor, physical capital, R&D, profit and others). Together with simple techniques of estimation (OLS, fixed effects) it is used generalized method of moments (GMM).

1 Úvod

Přímé zahraniční investice (PZI) znamenají pro hostitelskou ekonomiku jak výhody (rozšiřování trhu, přístup k novým technologiím, nové zkušenosti, příliv nového kapitálu a manažerských dovedností, zvyšování zaměstnanosti a konkurence aj.), tak i některé nevýhody (riziko převzetí domácích společností s cílem vytlačit domácí produkci, transfer domácích úspor do zahraničí za nevýhodných podmínek aj.).

V roce 1995 dosáhly toky PZI do České republiky částky 195,5 miliard Kč, v roce 2005 to bylo 1491,6 miliard. K dominantním příjemcům zahraničního kapitálu přitom stále patří zpracovatelský průmysl, i když v posledních deseti letech ze svého podílu ztratil. Data, popisující toky PZI do ČR, jsou shrnuta v tabulce 1. Cílem článku je popsat historický vývoj PZI v ČR v posledních letech a popsat literaturu zabývající se studiem determinant PI v ČR nebo ve střední Evropě. Hlavním cílem článku je odhadnout ekonometrický model determinant PZI v ČR v letech 2000–2006. Kapitola 2 popisuje zajímavé historické okamžiky, které ovlivnily toky PI v posledních letech. Kapitola 3 shrnuje dosavadní výzkum v oblasti. Kapitola 4 se zabývá popisem dat a metod. Kapitola 5 popisuje dosažené výsledky a kapitola 6 shrnuje článek.

rok	objem PZI v ČR (mld Kč)	meziroční přírůstek (mld Kč)	podíl zpracovatelského průmyslu
2000	818,3	186,8	38 %
2001	982,3	164	38 %
2002	1165,5	183	46 %
2003	1161,8	-3,7	42 %
2004	1280,6	119	40 %
2005	1491,6	211	38 %
2006	1666,8	175	36 %

Tabulka 1: PZI ve zprac. průmyslu ČR, 1999–2006. Zdroj: ČNB.

2 PZI v České republice v letech 2000–2006

V roce 1998 vznikl systém státních investičních pobídek a v roce 2000 byl přijat zákon o investičních podmínkách. V zákoně byla uvedena kritéria, za kterých jsou pobídky uděleny.

Dle výročních zpráv České Národní Banky (ČNB) došlo na konci 90. let k mimořádnému přílivu zahraničního kapitálu do českého zprac. průmyslu. Privatizovaly se podniky, expandovaly velké řetězce a dceřiné společnosti. V roce 2001 byla dokončena privatizace a došlo k poklesu investic do základního kapitálu. Narostly investice do kancelářských strojů a počítačů, potravin, papírenského průmyslu. V roce 2002 došlo k významné privatizaci společnosti Transgas. V roce 2003 byl příliv PZI nižší. Uskutečnil se zpětný odkud akcií Eurotel od zahraničního investora, dochází k prodeji zahraničního podílu v českém Telecomu investorům hlavně z USA a VB. Ke konci sledovaného období došlo k investicím do základního kapitálu, prodeji státních podílů v českém Telecomu a Unipetrolu, navyšují se stávající investice a vznikají nové podniky. V roce 2006 nedochází k realizaci žádné výraznější investice, příliv PZI byl tvořen zejména navyšováním kapitálu ve stávajících podnicích se zahraniční účastí. Nejvyšší PZI ve zpracovatelském průmyslu se alokují do automobilové výroby, koksování a chemického průmyslu.

Z geografického hlediska jsou ve sledovaném období nejvýznamnějšími investory Nizozemí, Německo a Rakousko. ČNB monitoruje zhruba 3000–4000 společností se zahraniční účastí, z nichž se zhruba jen 70 podniků podílí na celkových tocích PZI více než z poloviny.

3 Dosavadní empirický výzkum

Otázkou PZI se v minulosti zabývalo mnoho autorů empirických studií. Ve velkém množství z těchto analýz hrála důležitou roli i Česká republika. V případě determinant PZI se můžeme zabývat otázkou, proč zahraniční investoři

preferují určitou zemi před ostatními nebo proč plynou do určitých odvětvích průmyslu v dané zemi vyšší toky PZI [4]. U posledně jmenovaného má většina testovaných hypotéz kořeny v mikroekonomických teoriích alokace výroby a obchodu. Klasickým přístupem je aplikace teorií komparativních výhod, nicméně větší část testovaných hypotéz je odvozena z nových teoriích alokace, obchodu a průmyslové organizace.

Ekonometrický model determinant PZI může vysvětlovat PZI jako funkci mnoha faktorů. Jedním z nich může být velikost trhu. Takové výsledky prezentovali například Lankes a Venables [12], Pye [15]. Jiní autoři ukázali, že důležitým faktorem může být též očekávaný růst trhu [3].

Velikost zahraničního kapitálu mohou ovlivňovat pracovní náklady, což ukázali Pye [15], Bevan a Estrin [7] nebo Benáček a Víšek [5]. Kromě toho, Pye [15] specifikoval ve své práci i další důležité faktory: politická a ekonomická stabilita země. Stabilitu shledali důležitou i Lankes a Venables [12]. Rozhodování zahraničních investorů může záviset na důrazu na vědu a výzkum. Benáček a Víšek [5] ukázali, že investoři preferují investovat do odvětví zprac. průmyslu s vyššími výdaji na R&D.

Existují další faktory, které ovlivňují rozhodování investorů: vzdálenost od trhů západní Evropy [7, 3], přírodní zdroje [11], celková produktivita faktorů [5] nebo byrokratické překážky [14].

Existují dvě studie, které popisují determinanty v českém zpracovatelském průmyslu [5, 6]. V první studii autoři popisují determinanty PZI v českém zpracovatelském průmyslu v roce 1994, ve druhé studii zkoumají tutéž otázku v období mezi lety 1991 a 1997. Autoři došli k závěru, že nebylo možné najít ekonometrický model, který by popisoval determinanty pro všechny sektory zpracovatelského průmyslu. V české ekonomice existovaly segmenty, ve kterých se investoři chovali rozdílně. Za pomoci robustní techniky odhadu se autorům nakonec podařilo najít determinanty pro zúžený soubor pozorování.

4 Data a metodologie

K dispozici jsme měli data popisující 14 sektorů zpracovatelského průmyslu (členěné dle klasifikace OKEČ) mezi lety 2000-2006. Počet pozorování je 98 (=14x7). Data pocházejí z různých zdrojů. Hodnoty jednotlivých PZI poskytuje ČNB, ostatní veličiny poskytl Český statistický úřad (ČSÚ). Zde důležité zdůraznit, že data z ČSÚ prochází každoroční revizí, což znamená, že data dostupná v určitém roce se mohou odlišovat od stejných dat, která jsou k dispozici o pár let dříve nebo později.

Za vysvětlovanou veličinu jsme stejně jako Benáček a Víšek [6] zvolili intenzitu PZI (relace ročních toků přímých zahraničních investic k přidané hodnotě PZI/PH). Veličiny vstupující do regresního modelu byly vybrány na základě ekonomických teorií alokace v otevřené ekonomice. Umístění PZI je spojeno s komparativními výhodami jednotlivých odvětví.

Fyzický kapitál a práce. Použili jsme kombinaci kapitálové náročnosti přidané hodnoty (K/PH) a práce na jednotku čisté produkce (L/PH). Tuto kombinaci vyjádříme podílem K/L (vybavenost práce kapitálem). Autoři v podobných studiích předpokládají, že postkomunistické země disponují komparativní výhodou v poměrně dobře kvalifikované pracovní síle. Očekáváme kladné znaménko u proměnné L/PH: čím vyšší je pracovní náročnost přidané hodnoty, tím více je odvětví konkurenceschopné a tím více je atraktivní pro investora. Fyzický kapitál coby substitut pracovní náročnosti je drahým faktorem a lze očekávat záporné znaménko odhadnutého parametru. Očekáváme tedy záporné znaménko u parametru pro veličinu K/L.

Celková produktivita faktorů (TFP). Tuto proměnnou jsme použili jako proxy veličinu vyjadřující technickou efektivnost použitých faktorů: čím vyšší je TFP, tím nižší množství faktorů je potřebné k produkci dané jednotky [4]. Očekáváme kladné znaménko. Existuje více způsobů, jak odhadnout TFP. Pomocí úprav Cobby-Douglasovy produkční funkce dostaneme vztah

$$TFP_{it} = \frac{PH_{it}}{K_{it}^{\alpha} L_{it}^{1-\alpha}}$$

kde α značí podíl kapitálu na přidané hodnotě. Je často předpokládáno, že rozumná hodnota parametru α se pohybuje mezi 0,25 [13] a 0,35 [10]. V našich analýzách zvolíme velikost α jako 0,3 [8].

Vývoj cenového indexu v čase (PPI). Tento inflační index vyjadřuje změny cen produktů průmyslových výrobců oproti roku 2000. Čím vyšší je index, tím větší je potenciál pro růst odvětví a investice do tohoto odvětví. Očekáváme tedy pozitivní znaménko u této veličiny.

Věda a výzkum (R&D). Kvalita výrobků může být dalším důležitým faktorem, a to zejména v posledních letech, kdy význam R&D roste. Úroveň R&D může být měřena počtem zaměstnanců zaměstnaných v oblasti R&D nebo pomocí celkových výdajů na R&D v korunách. V této studii je použit druhý způsob. Vyšší výdaje jsou znakem vyšší kvality výrobků. Čekáme tedy kladné znaménko.

Zisk na pracovníka. Zahraniční investice mohou být přitahovány firmami s vyšším ziskem anebo přítomnost PZI v daném odvětví může být příčinou vyššího zisk. Očekáváme kladné znaménko.

Energetická náročnost. V článku jsme využili informaci o spotřebě benzínu a nafty, zemního plynu, elektřiny, uhlí. Uvážíme-li fakt, že během posledních deseti let všechny výše uvedené zdroje energie zdražily, nečekáme, že by byl tento faktor signifikantní.

Mzdy. Poslední (a spíše doplňkovou) veličinou v našem ekonometrickém modelu jsou hrubé mzdy zaměstnanců. Očekáváme kladné znaménko tohoto faktoru jako efekt (spillover) vyšších zisků v odvětvích s vyššími PZI.

Z možností odhadu jsme vybrali přístup Chenga a Kwana [9]: porovnat v modelu současnou a minulou hodnotu PZI s ostatními vysvětlujícími veličinami. Předpokládali jsme, že trvá určitou dobu než se PZI dostane do svého rovnovážného stavu a že zpožděné hodnoty mohou být vyjádřeny jako přízpusobovací proces:

$$Y_{it} - Y_{it-1} = \alpha(Y_{it}^* - Y_{it-1}) \quad (1)$$

$$Y_{it} = (1 - \alpha)Y_{it-1} + \alpha Y_{it}^* \quad (2)$$

kde Y_{it}^* je rovnovážný stav PZI a α je menší než 1 pro zachování stability. Rovnovážný stav PZI je ovlivněn jednotlivými faktory X_{it} , které představují vektor $k \in 1 \dots K$ vysvětlujících proměnných:

$$Y_{it}^* = \beta X_{it} + v_{it} \quad (3)$$

kde v_{it} je chybový člen zahrnující jak efekt odvětví a časový efekt. Po malých úpravách dostaneme:

$$\begin{aligned} Y_{it} &= \delta Y_{it-1} + \lambda X_{it} + \varepsilon_{it} \\ \varepsilon_{it} &= \mu_i + \eta_t + u_{it} \end{aligned} \quad (4)$$

kde $\delta = 1 - \alpha$ a $\lambda = \alpha\beta$ jsou koeficienty, které máme odhadnout (β je vektor rozměrů $1 \times K$); $\varepsilon_{it} = \alpha v_{it}$, μ_i je efekt odvětví a η_t je časový efekt.

Model (1) neřeší jeden závažný problém při odhadu pomocí nejmenších čtverců (OLS). Zpožděná proměnná Y_{t-1} a časově invariantní efekt odvětví μ_i jsou korelovány a OLS odhad modelu nebude konzistentní. Vyjádříme model pomocí prvních diferencí:

$$\Delta Y_{it} = \delta \Delta Y_{it-1} + \lambda \Delta X_{it} + \Delta \varepsilon_{it} \quad (5)$$

Vzhledem k tomu, že ΔY_{it-1} a $\Delta \varepsilon_{it}$ jsou stále korelovány, použili jsme odhad zobecněnou momentovou metodou (GMM). Arellano a Bond [1] představili odhad GMM (difference, DIFF-GMM) který řeší a odhaduje model jako systém rovnic, jednu pro každou časovou periodu. Rovnice se liší pouze v použitých instrumentech. Arellano a Bover [2] představili další verzi odhadu GMM (system, SYS-GMM) ve snaze zlepšit původní DIFF-GMM odhad pro malé počty pozorování. Vzhledem k tomu, že náš počet pozorování není příliš velký, rozhodli jsme se použít tento odhad. Správnost instrumentů je testována pomocí Sarganova testu a Arellanových-Bondových statistik [11]. Vzhledem k nízkému počtu pozorování jsme se rozhodli rozšířit analýzy také o odpovídající OLS odhady a odhady pomocí fixních efektů.

5 Výsledky

Tabulka 2 uvádí výsledky odhadů regresních modelů. Ve sloupcích (a) a (b) jsou uvedeny odhady pomocí OLS a fixních efektů. Výsledky těchto odhadů

mohou čelit určitému vychýlení, proto jsou ve čtvrtém sloupci uvedeny i odhady GMM ve sloupci (c). Všechny regresní modely zahrnují nula-jedničkové časové indikátory. Koeficienty determinace v modelech (a) a (b) jsou uspokojivé (dosahují 73% a 51%). Správnost GMM ověřují dva testů specifikace. Sarganův test zamítá nulovou hypotézu o správnosti použití instrumentů. Na druhou stranu Arellanův-Bondův test pro AR(2) proces (AB 2) napovídá, že se model nepotýká s problémem autokorelace druhého řádu. Budeme předpokládat, že je model správně specifikován.

Výsledky ve sloupci (c) ukazují že fyzický kapitál a práce hrají důležitou roli při rozhodování investorů kam umístit své investice, stejné výsledky udává i OLS odhad. Zdá se, že v posledních letech investoři směřují svůj kapitál spíše do pracovně náročných odvětví, zatímco fyzický kapitál je pro ně příliš drahým faktorem. To je v souladu s naším očekáváním. Proměnná vyjadřující zisk na pracovníka je signifikantní pouze pro model (a). Dle našich hypotéz, zisk v odvětvích s vysokými PZI by měl být vyšší než v těch odvětvích, které PZI nepřitahují.

	OLS (a)	FE (b)	GMM (c)
zpožděné PZI	0.62***(0.07)	0.12(0.09)	0.16(0.14)
K/L	-42.69*** (9.32)	8.86 (33.70)	-89.35*** (22.60)
Zisk	402.87** (189.23)	89.26(99.97)	206.62(255.72)
R & D	925.33** (434.49)	-227.50(807.89)	2800.29** (1102.46)
Energie	0.15*** (0.03)	0.17*** (0.02)	0.22*** (0.05)
Mzdy	-0.02(0.01)	-0.001(0.01)	0.11*** (0.04)
PPI	-4.05(3.62)	11.10** (4.37)	-12.99*** (4.74)
TFP	-150.39(397.34)	-516.84(320.04)	-941.23(685.92)
N	98	98	84
R^2	0.73	0.51	–
Sarganův test	–	–	0.002
AB 2	–	–	0.802

Tabulka 2: Notes: * signifikantní na 10%; ** signifikantní na 5 %; *** signifikantní na 1 %. V závorkách jsou uvedeny směrodatné odchylky. Hausmannův test zamítá nulovou hypotézu (náhodné efekty).

Zahraniční investoři kladou důraz na vědu a výzkum, parametr je signifikantní jak pro model (a), tak pro model (c). Vyšší výdaje na vědu a výzkum znamenají vyšší kvalitu výrobků a vyšší zájem ze strany zahraničního investora². Protichůdné výsledky podávají odhady parametrů u proměnné PPI.

²Samozřejmě je potřeba uvažovat možnost, že vyšší investice se mohou projevit jako efekt vyšších výdajů na vědu a výzkum.

Index je signifikantní pro model (b) a (c), avšak u druhého jmenovaného se záporným znaménkem. Dle našeho očekávání by toto znaménko mělo být kladné. Při bližším zkoumání dat lze zjistit, že cenový index má pro některá odvětví spíše klesající charakter - zejména tam, kde jsou toky PZI vysoké. V takových odvětvích mohou být ceny konkurencí tlačeny dolů.

Překvapující je signifikantní odhad parametru energetické náročnosti, a to s kladným znaménkem. To je v rozporu s naším očekáváním. Je možné, to může být způsobeno například nižšími cenami na počátku sledovaného období³ anebo pokračování trendu investování do energeticky náročných odvětví. Proměnná popisující efektivnost využití faktorů (TFP) není překvapivě signifikantní. Proměnná mzdy je signifikantní v modelu (c).

6 Závěr

Článek analyzuje některé aspekty chování zahraničních investorů v českém zpracovatelském průmyslu. Odhadli jsme panelová data 14ti odvětví zpracovatelského průmyslu mezi lety 2000–2006 pomocí tří technik odhadu (OLS, fixní efekty, GMM). Výsledky naznačují, že kvalifikovaná pracovní síla je jednou z komparativních výhod v odvětvích českého zprac. průmyslu. Investoři preferují odvětví, ve kterých je kladen důraz na vědu a výzkum, jsou více zisková a jsou více energeticky náročná. Problematickým faktorem se ukázal být vývoj cenového indexu v čase a také souhrnná efektivita faktorů. Lze se domnívat, že data nemají homogenní charakter a chování investorů může být heterogenní. Domněnku můžeme podpořit například použitím metody nejmenších useknutých čtverců (LTS), vyčleněním určitých sektorů z datového souboru a opětovnou analýzou pomocí uvedených technik. Můžeme se dohadovat, zda by takovým odvětvím byl např. kožedělný průmysl (nízké toky PZI) nebo automobilový průmysl (extrémně vysoké toky PZI). Výsledky by se pravděpodobně zlepšily. Toto zkoumání je předmětem další, obsáhlejší studie.

Literatura

- [1] Arellano M. and Bond S. (1991). *Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations*. Review of Economic Studies, 58, 277–297.
- [2] Arellano M. and Bover O. (1995). *Another look at the instrumental variable estimation of error-components models*. Journal of Econometrics, 68, 29–51.
- [3] Barrell R., Holland D. (1999). *Foreign Direct Investment in Central European Manufacturing: Consequences for enterprise restructuring and labor productivity*. Research of NIESR, London.

³Ceny energií vzrostly prudce zejména na konci tohoto období a v posledních dvou letech. Ceny elektřiny ještě v roce 2001 nebo 2003 patřily k nejnižším v EU.

- [4] Benáček V. (2000), *Přímé zahraniční investice v České republice: praxe, teorie a aplikace*. Politická ekonomie, 1, 7–24.
- [5] Benáček V. and Víšek J.Á. (1999). *The determinants of FDI flows into czech manufacturing industries: theoretical background for and empirical testing*. Research of Charles University, IES and ACE Brusel.
- [6] Benáček V. and Víšek J.Á. (1999a). *The determinants of FDI in the Czech economy: The Robust Analysis applied on Industrial data of 1991-1997*. Charles University, mimeo.
- [7] Bevan A., Estrin S. (2000). *The Determinants of Foreign Direct Investment in Transition Economies*. CEPR discussion paper No. 2638, London, Center for Economic Policy Research.
- [8] Caselli F. (2005). *Accounting for cross-country income differences*. Handbook of economic growth, Elsevier, 679–742.
- [9] Cheng L. and Kwan Y. (2000). *What are the determinants of the location of foreign direct investment? The Chinese experience*. Journal of International Economics 51, 379–400.
- [10] Collins S.M., Bosworth B.P., Rodrik D. (2000). *Economic growth in East Asia: Accumulation versus assimilation*. Brooking papers of economic activity, 1996, 2, 135–203.
- [11] Kinoshita Y. and Campos N.F. (2003). *Why Does FDI Go Where it Goes? New Evidence from the Transition Economies*. William Davidson Institutute Working Paper Number 573.
- [12] Lankes H.P., Venables A.J. (1997). *Foreign Direct Investment in Eastern Europe and the former Soviet Union: Result from a Survey of Investors*. In Zecchini S. (ed.): *Lessons from the Economic Transition*. Central and Eastern Europe in the 1990s. Kluwer Academic Publishers and OECD, Norwell, Massachusetts.
- [13] Prescott E.C. (1998). *Needed: A theory of total factor productivity*. International Economic Review, 39, 3, 525–551.
- [14] Pomery C. (1997). *The First CzechInvest Annual Survey on FDI in the Czech Republic*. CzechInvest, Praha.
- [15] Pye R.B.K. (1998). *Foreign Direct Investment in Central Europe: The Experience of Major Western Investors*. European Management Journal, Special Issue.

Poděkování: Tato práce byla podporována grantem GAUK 2157/2007.

Adresa: E. Ryšavá, IES FSV UK, Opletalova 26, 110 00 Praha 1

E-mail: rysavaeva@volny.cz

ZJIŠŤOVÁNÍ OPTIMÁLNÍHO POČTU SHLUKŮ VE STATISTICKÝCH PROGRAMOVÝCH SYSTÉMECH

Hana Řezanková, Dušan Húsek

Klíčová slova: Shluková analýza, počet shluků, statistické programy.

Abstrakt: Příspěvek se zabývá přístupy ke stanovení optimálního počtu skupin objektů, které jsou implementovány v rámci shlukovacích algoritmů v komerčních statistických programových systémech. Některé z těchto přístupů jsou ilustrovány na úloze nalezení skupin poslanců ruského parlamentu na základě výsledků hlasování v roce 2004. Jsou sledovány možnosti programových systémů S-PLUS, SAS, SPSS a SYSTAT.

Abstract: The paper deals with approaches to determination of the optimal number of groups of objects which are implemented in clustering algorithms in commercial statistical software packages. Some of these approaches are applied to the example of finding groups of deputies of the Russian parliament on the base of roll-call votes in 2004. The capabilities of software packages S-PLUS, SAS, SPSS, and SYSTAT are described.

1 Úvod

Shluková analýza je užitečným nástrojem pro zkoumání vztahů v souborech dat. Jejím výsledkem může být nalezení skupin (shluků) podobných objektů, proměnných, kategorií, případně i zjištění složitějších vazeb mezi objekty a proměnnými, viz [7]. Může být aplikována v různých oblastech vědy a výzkumu. Tento článek bude zaměřen na shlukování objektů.

Vícerozměrná statistická analýza zkoumá m -rozměrná *pozorování* objektů (statistických jednotek, datových jednotek). Počet těchto objektů budeme označovat písmenem n . Prvky vektoru pozorování jsou hodnoty *statistických znaků* neboli *proměnných*. Vstupní matici budeme v dalším textu označovat symbolem \mathbf{X} a pro její prvky budeme používat značení x_{il} , kde $i = 1, 2, \dots, n$ a $l = 1, 2, \dots, m$. Matice \mathbf{X} je tedy rozměru $n \times m$. Počet shluků budeme dále označovat písmenem k a jednotlivé shluky obvykle jako C_h , kde $h = 1, 2, \dots, k$. Cílem shlukové analýzy je tedy přiřadit každý zkoumaný objekt do některého z k shluků. Tyto shluky mohou být buď *disjunktní* (objekt je přiřazen právě do jednoho shluku), nebo *překrývající se*. Přiřazení objektů je buď *pevné* (objekt je přiřazen jednoznačně), nebo *fuzzy*, kdy se stanovuje míra příslušnosti u_{ih} , spočtená pro každý i -tý objekt a h -tý shluk C_h , přičemž $0 \leq u_{ih} \leq 1$ pro všechna $i = 1, \dots, n$ a všechna $h = 1, \dots, k$ a $\sum_{h=1}^k u_{ih} = 1$ pro všechna i .

Příspěvek se zaměřuje na metody ohodnocení výsledků shlukování, které jsou implementované v současných statistických programových systémech. Tyto metody mohou posloužit jednak k určení, kterou metodou byly vytvořeny kvalitnější shluky (výběru metody), jednak ke stanovení optimálního počtu shluků. Zjišťovány byly některé možnosti systémů S-PLUS, SAS, SPSS a SYSTAT.

2 Indexy, koeficienty a jiná kritéria

Všechny metody hodnocení se v podstatě zaměřují na zjišťování, jak kompaktní shluky byly vytvořeny a jak jsou jednotlivé shluky odlišné. Je sledován základní princip disjunktního shlukování, aby objekty z určitého shluku si byly velmi podobné a aby pro každé objekty ze dvou různých shluků platilo, že jsou hodně odlišné. Při použití statistické terminologie jde o to, aby variabilita uvnitř shluku byla co nejmenší a naopak variabilita mezi shluky co největší. Samozřejmě platí, že čím je počet shluků větší, tím jsou shluky více homogenní. Z tohoto důvodu se při usuzování na vhodný počet shluků uplatňují různá speciální kritéria.

Na kvalitu shluků lze usuzovat pomocí poměru determinace, který se konstruuje analogicky jako v analýze rozptylu. Nazývá se *RSQ (R-square) index*. Jde o podíl meziskupinové variability SS_B (*Sum of Squares Between groups*) na celkové variabilitě SS_T (*Total Sum of Squares*). Vzhledem k tomu, že platí vztah $SS_T = SS_B + SS_W$, kde SS_W je součet čtverců, který charakterizuje vnitroskupinovou variabilitu (*Sum of Squares Within a group*), může být RSQ index vyjádřen pomocí vnitroskupinové variability, tj.

$$I_{RSQ} = \frac{SS_T - SS_W}{SS_T} = \frac{\sum_{i=1}^n \sum_{l=1}^m (x_{il} - \bar{x}_l)^2 - \sum_{h=1}^k \sum_{\mathbf{x}_i \in C_h} \sum_{l=1}^m (x_{il} - \bar{x}_{hl})^2}{\sum_{i=1}^n \sum_{l=1}^m (x_{il} - \bar{x}_l)^2}. \quad (1)$$

R-kvadrát index se používá například při porovnání různých postupů při hierarchické shlukové analýze. Pro stanovení optimálního počtu shluků se využívá *SPRSQ (semipartial R-squared) index*, který reprezentuje pokles hodnoty R-kvadrát indexu způsobený spojením dvou shluků, tj.

$$I_{SPRSQ}(k) = I_{RSQ}(k+1) - I_{RSQ}(k). \quad (2)$$

Čím je hodnota SPRSQ indexu nižší, tím menší je změna jak meziskupinové, tak vnitroskupinové variability. Malá změna indikuje, že spojením dvou shluků se vnitroskupinová variabilita zvýšila málo a že je tudíž menší počet shluků lepším výsledkem.

Dalším indexem inspirovaným analýzou rozptylu je *pseudo F (PSF) index*, založený na analogii s F statistikou. Podle jeho autorů je označován též jako *CHF (Calinski-Habarasz F) index*. Počítá se podle vzorce

$$I_{CHF} = \frac{\frac{SS_B}{k-1}}{\frac{SS_W}{n-k}} = \frac{(n-k) \cdot SS_B}{(k-1) \cdot SS_W}, \quad (3)$$

v němž použité symboly mají stejný význam jako u R-kvadrát indexu. Vysoké hodnoty CHF indexu indikují dobře oddělené shluky.

Ze součtů čtverců dále vychází *index PTS*, resp. *PST2* (*pseudo T-square*). Tato *pseudo T-kvadrát statistika* hodnotí spojení dvou shluků do jednoho. Spojení h -tého a h' -tého shluku je hodnoceno pomocí vztahu

$$I_{PTS} = \frac{B_{hh'}}{\frac{W_h + W_{h'}}{n_h + n_{h'} - 2}}, \quad (4)$$

kde $B_{hh'}$ je mezishlukový součet čtverců a W_h a $W_{h'}$ jsou vnitroshlukové součty čtverců. PTS index se obvykle počítá pro jednotlivé úrovně hierarchického shlukování. Kvantifikuje odlišnosti mezi dvěma shluky, které jsou v daném kroku spojovány. Je-li při grafickém znázornění pro k shluků na křivce zřejmý „skok“, pak vhodný počet shluků je $(k + 1)$.

Pouze na vnitroskupinové variabilitě je založen *RMSSTD* (*Root Mean Square STandard Deviation*) *index*, jenž měří homogenitu výsledných shluků:

$$I_{RMSSTD} = \sqrt{\frac{SS_W}{m(n-k)}}, \quad (5)$$

kde SS_W je součet čtverců, který charakterizuje vnitroskupinovou variabilitu (viz RSQ index). Nižší hodnoty indexu tedy indikují lepší rozdělení objektů do shluků.

K dalším indexům pro stanovení optimálního počtu shluků patří *Daviesův a Bouldinův (DB) index*, který kvantifikuje průměrnou podobnost mezi shlukem a jeho nejpodobnějším protějškem. Vzdálenost shluků $D_{hh'}$ je reprezentována vzdáleností centroidů. Uvažuje se přitom Minkowského vzdálenost. Dále je definována míra disperze $s_{D,h}$ shluku C_h jako kvadratický průměr vzdáleností jednotlivých objektů od svého centroidu. Označme

$$R_h = \max_{h', h' \neq h} \left\{ \frac{s_{D,h} + s_{D,h'}}{D_{hh'}} \right\}.$$

Výsledný DB index je definován jako průměr z hodnot R_h , tj.

$$I_{DB} = \frac{\sum_{h=1}^k R_h}{k}. \quad (6)$$

Nejlepší je takové rozdělení objektů do shluků, při němž je hodnota indexu minimální.

Jiným kritériem je *Dunnův index* (též *separační index*). Pro něj se definuje jednak vzdálenost mezi shluky $D_{hh'}$ jako minimální vzdálenost dvou objektů

z různých shluků, jednak poloměr shluku $diam_h$ jako maximální vzdálenost dvou objektů z daného shluku. Výsledný Dunnův index je dán vzorcem

$$I_D = \min_{1 \leq h \leq k} \left\{ \min_{1 \leq h' \leq k} \frac{D_{hh'}}{\max_{1 \leq l \leq k} diam_l} \right\}. \quad (7)$$

Vysoké hodnoty tohoto indexu indikují kompaktní a dobře oddělené shluky.

Pro hodnocení shluků vytvořených pomocí metod rozkladu (tj. nehierarchických) se používá *obrysový koeficient* (*silhouette coefficient*), který vychází z hodnot počítaných pro obrysový graf. Pro $\mathbf{x}_i \in C_g$ spočteme hodnoty ψ_i na základě průměrných vzdáleností sledovaného objektu s ostatními objekty v jednotlivých shlucích, a to

$$\psi_i = \frac{\eta_i - \mu_i}{\max\{\eta_i, \mu_i\}},$$

kde η_i je průměrná vzdálenost i -tého objektu s objekty z téhož shluku a μ_i je minimum z průměrných vzdáleností počítaných pro ostatní shluky. Z hodnot ψ_i se počítá průměrná hodnota za celý rozklad, tj.

$$\psi_P = \sum_{i=1}^n \frac{\psi_i}{n}. \quad (8)$$

Čím vyšší je tato průměrná hodnota, tím jsou shluky kompaktnější. Obrysový koeficient lze použít jednak pro porovnání metod, jednak pro stanovení optimálního počtu shluků.

Dále můžeme pro stanovení optimálního počtu shluků využít *informační kritéria*. Rozlišují se dvě, přičemž jedno je založeno na bayesovském přístupu. To bývá označováno zkratkou *BIC* (*Bayesian Information Criterion*) a nazývá se *Schwarzovo bayesovské kritérium*. Druhé kritérium je Akaikeovo, označuje se *AIC* (*Akaike Information Criterion*). Podrobněji viz [7]. V praxi se postupuje tak, že jsou nejdříve spočteny hodnoty *BIC*, resp. *AIC*, pro každý počet shluků v rámci specifikovaného intervalu. Na jejich základě se stanoví počáteční odhad pro počet shluků, který odpovídá minimu ze zjištěných hodnot. Poté je počáteční odhad upraven zjišťováním největšího vzrůstu ve vzdálenosti mezi dvěma nejbližšími shluky v každé etapě hierarchického shlukování.

Další speciální postupy jsou uplatňovány v rámci *metod založených na hustotě*. Provádí se posloupnost testů, v nichž nulová hypotéza označuje, že počet shluků je hodnota k nebo menší, a alternativní hypotéza označuje, že počet shluků je větší než k . Výsledkem aplikace metody založené na hustotě je sice stanovení počtu shluků, to ovšem závisí na parametrech shlukování, které zadává uživatel.

Je-li aplikována *fuzzy shluková analýza*, pak hodnocení vychází z koeficientů příslušnosti u_{ih} . Základním koeficientem uváděným v literatuře a implementovaným v současném softwaru je *koeficient rozkladu*, který vyjadřuje

míru překrytí mezi shluky. Zatímco v knihách [1] a [2] je spojován se jménem Bezdek (1981), v systému S-PLUS je podle knihy [5] nazýván *Dunnův koeficient rozkladu*. Vzhledem k tomu, že jeden Dunnův index byl již definován, budeme v souladu s [2] používat značení *PC (Partition Coefficient) index*, tj. I_{PC} . Tento index se počítá podle vzorce

$$I_{PC} = \frac{1}{n} \sum_{i=1}^n \sum_{h=1}^k u_{ih}^2 \quad (9)$$

a nabývá hodnot z intervalu $\langle 1/k; 1 \rangle$. Můžeme rozlišit dvě extrémní situace, a to *úplně fuzzy shlukování*, kdy $I_{PC} = 1/k$, a *pevné (disjunktní) shlukování*, kdy $I_{PC} = 1$. Uvažujeme-li počet shluků od dvou do $(n - 1)$, pak optimální počet lze získat pomocí maxima z indexů počítaných pro všechny počty shluků z tohoto intervalu. Protože však se změnou počtu shluků se mění také interval, v němž se může hodnota koeficientu nacházet, je lépe porovnávat hodnoty *normalizovaného koeficientu rozkladu*, které jsou z intervalu od 0 do 1.

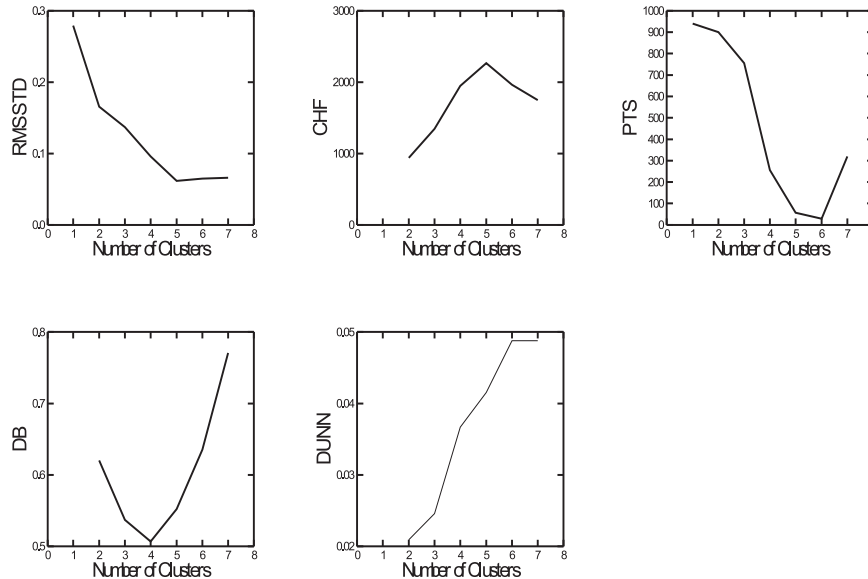
Pokud jde o jednotlivé programové produkty, pak se částečně překrývají možnosti systémů SAS a SYSTAT. Součástí výstupů hierarchické shlukové analýzy v systému SAS je jak číselný výstup hodnot indexů, tak grafický (pouze vybraných), SYSTAT poskytuje grafické srovnání hodnot. SAS zahrnuje indexy RSQ, SPRSQ, PSF (číselně též pro metodu k -průměrů) a PST2, viz též [9], SYSTAT indexy CHF, PTS, RMSSTD, DB a Dunnův. SAS navíc poskytuje tzv. kubické shlukovací kritérium (CCC) a v rámci procedury MODECLUS umožňuje provádět posloupnost testů, viz výše. Systém SPSS umožňuje využít informační kritéria (BIC a AIC), a to v rámci dvoukrokové shlukové analýzy, která je založena na principu algoritmu BIRCH, viz [10]. Ostatní výše uvedené postupy jsou implementovány v systému S-PLUS. Hodnota obrysového koeficientu je součástí obrysového grafu, který je nabízen u algoritmů k -medoidů (PAM a CLARA) a u fuzzy shlukové analýzy (FANNY). Součástí výstupu tohoto algoritmu jsou hodnoty Dunnova koeficientu rozkladu.

3 Aplikace na data z hlasování v parlamentu

Možnosti programových systémů jsou ilustrovány na shlukování poslanců ruského parlamentu v roce 2004 na základě jednotlivých hlasování, viz [3]. Datový soubor obsahuje 430 vektorů odpovídajících hlasování jednotlivých poslanců (20 poslanců s malou účastí bylo vynecháno). Těchto hlasování bylo 3150. Odpovídající proměnné jsou binární, přičemž hodnota 1 indikuje kladný hlas a hodnota 0 jiný výsledek. Cílem shlukové analýzy bylo nalézt skupiny podobně hlasujících poslanců. Lze předpokládat, že tyto skupiny by mohly souviset s příslušností poslanců ke svým politickým stranám. Jednotlivé politické strany byly zastoupeny v počtech 285, 51, 37, 41 a 16 poslanců bylo nezávislých.

Shlukování jsme prováděli jednak na základě původních dat (PD), jednak na základě dat částečně analyzovaných. V tomto případě byla na objekty aplikována faktorová analýza a vstupem do shlukování byly první dva, resp. tři, vektory faktorových zátěží (FZ). Jde o postup navržený pro shlukování binárních proměnných, viz [4].

Na obrázku 1 jsou ohodnoceny výsledky získané hierarchickou shlukovou analýzou (na základě prvních dvou vektorů faktorových zátěží) pro počty shluků od dvou do sedmi. Podle RMSSTD a CHF indexu je vhodné rozdělení objektů do pěti shluků, PTS indikuje jako optimální dva shluky (maximum je v hodnotě 1), zatímco DB index čtyři shluky a Dunnův index šest a sedm shluků (obdobné grafy pro CHF a PTS lze získat v systému SAS).



Obrázek 1: Ohodnocení rozdělení objektů do různých počtu shluků (FZ, euklidovská vzdálenost, metoda průměrné vazby – SYSTAT).

V tabulce 1 jsou uvedeny hodnoty obrysového koeficientu pro metody rozkladu (PAM, CLARA a FANNY ze systému S-PLUS) získané na základě původních dat (PD) a na základě dvou a tří faktorů (2f a 3f). Lze porovnávat pouze výsledky analýz se stejným počtem proměnných. (Výsledky pro fuzzy shlukovou analýzu aplikovanou na původní data nejsou uvedeny, protože objekty byly přiřazeny ke shlukům se shodnými hodnotami měř příslušnosti.) Jako optimální byly vyhodnoceny dva a čtyři shluky, přičemž lépe ohodnoceno bylo rozdělení objektů do skupin pomocí metod k -medoidů.

Pokud jde o hodnocení výsledků dvoukrokové shlukové analýzy, pak při analýze původních dat a využitím věrohodnostní míry nepodobnosti byly pomocí obou kritérií vyhodnoceny jako optimální dva shluky. Při analýze

vektorů faktorových zátěží se výsledky lišily v závislosti na počtu použitých vektorů, na použité míře a na použitém kritériu.

S využitím normalizovaného Dunnova koeficientu rozkladu ve fuzzy shlukové analýze byly jako optimální vyhodnoceny tři shluky (bez normalizace byla největší hodnota pro dva shluky).

Počet shluků	PD PAM (CLARA)	FZ(3f) PAM (CLARA)	FZ(2f) PAM (CLARA)	FZ(3f) FANNY	FZ(2f) FANNY
2	0,39	0,57 (0,57)	0,63 (0,64)	0,58	0,63
3	0,18	0,53 (0,52)	0,61 (0,62)	0,53	0,61
4	0,21	0,67 (0,67)	0,65 (0,67)	0,41	0,51

Tabulka 1: Hodnoty obrysového koeficientu pro různé počty shluků.

Při rozdělení objektů do dvou shluků na základě faktorových zátěží byla velikost jednoho shluku blízká (resp. stejná) velikosti zastoupení největší politické strany (285), a to jak v případě metod PAM a FANNY, tak v případě metody k -průměrů a metod hierarchického shlukování (při použití euklidovské vzdálenosti a metod úplného spojení a průměrné vazby). Aplikací dvoukrokové shlukové analýzy byl shluk o velikosti 287 získán dokonce na základě původních dat (při použití věrohodnostní míry nepodobnosti). Podrobněji viz [8].

4 Závěr

V současných programových systémech již bývají implementovány metody, které umožňují hodnotit kvalitu shluků, vybírat vhodné postupy shlukování a určovat vhodné počty shluků. Konečný závěr však musí obvykle učinit uživatel sám na základě zobrazených hodnot či grafů. Jsou však i výjimky, například programový systém SPSS v rámci dvoukrokové shlukové analýzy umožňuje zobrazit výstup pro optimální počet shluků (ze zadaného intervalu). Uživatel ovšem musí zvolit jedno ze dvou kritérií.

Jak je vidět z výsledků experimentu týkajícího se shlukování poslanců, optimální počty shluků se mohou lišit v závislosti na použité metodě. Některé získané počty (2, 4, 5) mohou mít reálnou interpretaci ve smyslu politických seskupení. Při zjišťování velikostí shluků však pouze při rozdělení do dvou shluků jeden z nich velikostně přibližně odpovídá zastoupení největší politické strany v parlamentu. Při určení počtu shluků je kromě zmíněných koeficientů důležité též věcné a interpretační hledisko. Konečné rozhodnutí musí tedy učinit analytik. Jiný příklad na určování počtu shluků při shlukování okresů na základě ekonomických ukazatelů je popsán v [6].

Literatura

- [1] Abonyi J., Feil B. (2007). *Cluster analysis for data mining and system identification*. Birkhäuser Verlag AG, Berlin.
- [2] Gan G., Ma C., Wu J. (2007). *Data clustering: theory, algorithms, and applications*. ASA-SIAM, Philadelphia.
- [3] Húsek D., Frolov A.A., Polyakov P., Řezanková H. (2006). *Neural network analysis of Russian parliament voting patterns*. In: Issa, G. (ed.) Computer Science and Information Technology – CSIT 2006. Applied Science Private University, Amman, 328–334.
- [4] Húsek D., Řezanková H., Frolov A.A. (2006). *Overlapping clustering of binary variables*. In: Knowledge Extraction and Modelling [CD-ROM]. TILAPIA Edizioni, Italy.
- [5] Kaufman L., Rousseeuw P. (2005). *Finding groups in data: An introduction to cluster analysis*. Wiley, Hoboken.
- [6] Löster T., Popenková M. (2008). *Determination of optimum number of clusters*. In: Mezinárodní statisticko-ekonomické dny [CD-ROM]. VŠE, Praha.
- [7] Řezanková H., Húsek D., Snášel V. (2007). *Shluková analýza dat*. Professional Publishing, Praha.
- [8] Řezanková H., Húsek D. (2008). *Methods for the determination of the number of clusters in statistical software packages*. In: Mezinárodní statisticko-ekonomické dny [CD-ROM]. VŠE, Praha.
- [9] Stankovičová I., Vojtková M. (2007). *Viacrozmerné štatistické metódy s aplikáciami*. Iura Edition, Bratislava.
- [10] Žambochová M. (2008). *Algoritmus BIRCH a jeho varianty pro shlukování velkých souborů dat*. In: Mezinárodní statisticko-ekonomické dny [CD-ROM]. VŠE, Praha.

Poděkování: Tato práce byla vznikla v rámci projektu 1ET100300414 a výzkumného záměru AV0Z10300504 GA AV ČR.

Adresa: H. Řezanková, Vysoká škola ekonomická v Praze, nám. W. Churchilla 4, 130 67 Praha 3

D. Húsek, Ústav informatiky AV ČR, Pod Vodárenskou věží 2, 182 07 Praha 8

E-mail: rezanka@vse.cz, dusan@cs.cas.cz

MULTI-SAMPLE SPATIAL MEDIAN TEST NOT REQUIRING DISTRIBUTIONS OF THE SAME TYPE

Ján Somorčík, František Rublík

Keywords: Location parameters, multi-sample comparison, spatial median.

Abstract: We present a test for the multivariate multi-sample location problem. It is based on spatial medians. Therefore, it is more robust than the classical tools. Moreover, it can be used also in situations when the underlying distributions of the samples are not of the same type. We prove some asymptotic properties and present results of a Monte Carlo study.

Abstrakt: Predstavujeme test rovnosti parametrov polohy viacerých mnohorozmenných rozdelení, založený na priestorových mediánoch. Je robustnejší než klasické metódy. Navyše sa dá použiť aj v situáciách, keď rozdelenia pravdepodobnosti v súboroch nie sú rovnakého typu. V článku dokazujeme niekoľko asymptotických vlastností a uvádzame výsledky Monte Carlo simulácií.

1 Introduction

We consider the d -dimensional q -sample location problem. It means that we have q independent random samples from d -variate distributions with location parameters $\theta_1, \dots, \theta_q$ and we wish to test the hypothesis

$$H_0 : \theta_1 = \dots = \theta_q.$$

We suppose that the d -variate distributions possess densities w.r.t. Lebesgue measure. Typically, it is assumed that the densities are of the form $f(\cdot - \theta_a)$ ($a = 1, \dots, q$). This means that the underlying distributions can differ only in location parameters. A lot of tests have been developed to test the above hypothesis, a good overview can be found in [6]. We have been motivated by the Lawley-Hotelling test based on the comparison of sample means (see e.g. [6]). Its test statistic is

$$T^2 := \sum_{a=1}^q n_a (\bar{X}^{(a)} - \bar{X})^T S^{-1} (\bar{X}^{(a)} - \bar{X}),$$

where n_a is the sample size of the a -th sample, $\bar{X}^{(a)}$ is the sample mean of the a -th sample. \bar{X} and S denote the sample mean and the sample covariance matrix based on the pooled sample of all $n := n_1 + \dots + n_q$ data points. Existence of finite second order moments of the underlying distributions is required. Then the asymptotic distribution of T^2 under H_0 is $\chi_{(q-1)d}^2$.

It is well-known that the performance of the Lawley-Hotelling test is rather poor (it has low power) when the underlying distributions are heavy-tailed (see e.g. [5]). The reason is that the sample covariance matrix and the sample means are very sensitive to outliers. A more robust estimate of location is, for example, the spatial median. Therefore, in [5] we have replaced the sample means by sample spatial medians to obtain a more robust test.

First, a few words about spatial median. The sample spatial median $\hat{\mu}$ of the data points X_1, \dots, X_n is defined as

$$\hat{\mu} := \arg \min_{M \in \mathbf{R}^d} \sum_{i=1}^n \|X_i - M\|, \quad (1)$$

where $\|\cdot\|$ denotes the usual Euclidean norm. Uniqueness and existence of $\hat{\mu}$ is ensured unless the data points lie on a single line (see [3], one of the shortest contributions ever published in *Annals of Statistics*). There is no explicit formula to compute the spatial median. Hence, an iterative algorithm is needed. The most popular one seems to be the Weiszfeld's algorithm. It was developed already in 1937 and refined in [7] to ensure its convergence for an arbitrary starting point.

The sample spatial median $\hat{\mu}$ can be seen as an estimate of

$$\mu := \arg \min_{M \in \mathbf{R}^d} \mathbf{E}(\|X - M\| - \|X\|)$$

which is the spatial median of the underlying probability distribution. Now, we introduce a weak assumption from [2] about this distribution.

Assumption 1. *Let the density of the underlying distribution be bounded on every bounded subset of \mathbf{R}^d .*

Under Assumption 1 the sample spatial median is asymptotically normal:

$$\sqrt{n}(\hat{\mu} - \mu) \rightarrow N_d(0, V) \quad \text{in distribution.} \quad (2)$$

See [5] for the definition of the asymptotic covariance matrix V and [2] for the proof.

(1) defines the sample spatial median as the point from which the sum of distances to the data points is minimal. The robustness of spatial median against outliers is not obvious from this definition. To see it, compute the gradient of the function $\sum_{i=1}^n \|X_i - M\|$ with respect to M which must be a zero vector for $M := \hat{\mu}$. It follows that the sample spatial median is such a point that the unit-length vectors pointing from $\hat{\mu}$ to the data points are balanced, i.e. their sum is a zero vector. Now, the robustness of the spatial median can be seen from Figure 1: irrespective of how far X_4 has moved to the “north-east” the spatial median of X_1, \dots, X_4 does not change, whereas the sample mean “follows” X_4 .

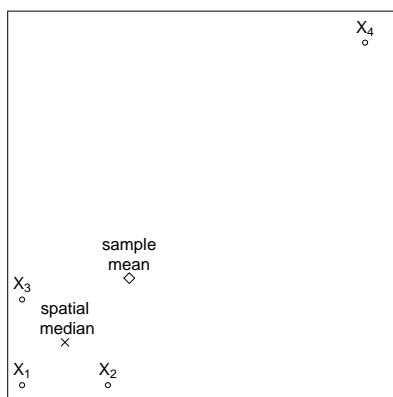


Figure 1: Robustness of spatial median.

2 Test statistics based on spatial medians

In [5] we introduced two test statistics whose form is inspired by the Lawley-Hotelling T^2 :

$$M_1 := \sum_{a=1}^q n_a (\hat{\mu}_a - \bar{\mu})^T \hat{V}^{-1} (\hat{\mu}_a - \bar{\mu}),$$

$$M_2 := \sum_{a=1}^q n_a (\hat{\mu}_a - \hat{\mu})^T \hat{V}^{-1} (\hat{\mu}_a - \hat{\mu}),$$

where the $\hat{\mu}_a$'s are the sample spatial medians, $\bar{\mu} := (1/n) \sum_{a=1}^q n_a \hat{\mu}_a$, $\hat{\mu}$ is the sample spatial median of the pooled sample and \hat{V} (see [5]) is an estimate of the asymptotic covariance matrix V of the sample spatial medians. We have shown that the asymptotic distribution of M_1 and M_2 under H_0 is $\chi_{(q-1)d}^2$. See [5] for more asymptotic properties and comparison with other test statistics.

However, there is an important deficiency concerning M_1 and M_2 : they require that the underlying densities are of the form $f(\cdot - \theta_a)$, i.e. of the same type, differing at most in location parameters. But it is natural to test H_0 also in more general situations. Think, for example, about the situation that we have random samples from d -variate spherically symmetric distributions with (possibly) different centers of symmetry $\theta_1, \dots, \theta_q$ but also with different scatter matrices describing the variability of the spherical distributions.

This motivated us to adjust M_1 or M_2 to get rid of the assumption that the distributions of the samples must be of the same type. Our solution is as follows:

$$M_3 := \sum_{a=1}^q n_a (\hat{\mu}_a - \tilde{\mu})^T \hat{V}_a^{-1} (\hat{\mu}_a - \tilde{\mu}),$$

where \hat{V}_a ($a = 1, \dots, q$) are the estimates of the asymptotic covariance matrices V_a of the sample spatial medians $\hat{\mu}_a$ and

$$\tilde{\mu} := \hat{S}^{-1} \sum_{a=1}^q n_a \hat{V}_a^{-1} \hat{\mu}_a$$

where $\hat{S} := \sum_{a=1}^q n_a \hat{V}_a^{-1}$. Hence, $\tilde{\mu}$ is a weighted average of the sample spatial medians $\hat{\mu}_a$. The impact of a particular $\hat{\mu}_a$ increases with increasing sample size n_a and decreases with “increasing” \hat{V}_a because \hat{V}_a measures the variability of the estimate $\hat{\mu}_a$. Our idea is not new, it was used e.g. in [4] in a different multi-sample testing problem.

Strictly speaking, the test statistics M_1 , M_2 and M_3 test the equality of the spatial medians μ_1, \dots, μ_q of the underlying distributions. If these distributions are of the same type the equality of the μ_a 's implies also the equality of the location parameters θ_a 's, no matter how the term “location parameter” is defined. But if the underlying distributions are not of the same type (it is the case when M_1 and M_2 can not be used but M_3 can) the above implication is not necessarily true. However, in many practical situations each of the underlying distributions possesses certain kind of symmetry and the location parameters θ_a 's are defined as the centers of these symmetries. Typically, also the spatial medians μ_a 's of the distributions coincide with the centers of the symmetries. Therefore, the equality of the location parameters θ_a 's is the consequence of the equality of the spatial medians μ_a 's.

To establish the asymptotic distribution of M_3 we will need an assumption about the asymptotic “proportions” of the samples:

Assumption 2. Let $\exists p_a := \lim(n_a/n) > 0$ for $a = 1, \dots, q$.

The following theorem provides a tool for testing the hypothesis H_0 . Its proof and also the proofs of the following theorems can be found in the Appendix.

Theorem 1. Let the underlying densities of the samples satisfy Assumption 1. Then under Assumption 2 the asymptotic distribution of M_3 under H_0 is $\chi_{(q-1)d}^2$.

In [5] it was shown that in case of underlying distributions of the same type the test statistics M_1 and M_2 are asymptotically equal under H_0 , i.e. $M_1 = M_2 + o_P(1)$. Now, a natural question arises about the relationship between M_3 and M_1 (or M_2). The following theorem gives the answer.

Theorem 2. *Let the distributions of the samples be of the same type (i.e. their densities are of the form $f(\cdot - \theta_a)$). Let f satisfy Assumption 1. Then under Assumption 2: $M_3 = M_1 + o_P(1)$.*

Now, we are going to study the asymptotic performance of M_3 when H_0 is not true. Consider the sequence of Pitman alternatives, i.e. the spatial medians do not share the same value μ but the spatial median of the distribution of the a -th sample is

$$\mu + \frac{h_a}{\sqrt{n}},$$

where the h_a 's are some constant vectors satisfying

$$\sum_{a=1}^q p_a V_a^{-1} h_a = 0, \quad (3)$$

which means that asymptotically the “shifts” of the distributions are balanced.

Theorem 3. *Under Pitman alternatives and Assumptions 1 and 2 the asymptotic distribution of M_3 is noncentral chi-squared $\chi_{(q-1)d}^2(\delta)$ where the noncentrality parameter is $\delta := \sum_{a=1}^q p_a h_a^T V_a^{-1} h_a$.*

Condition (3) is just technical and enables us to compare M_3 with other tests by the ratio of the noncentrality parameters. Note that in case of underlying distributions of the same type the noncentrality parameters of M_1 , M_2 and M_3 are the same (cf. [5]).

3 Monte Carlo study

We have performed a simulation study to illustrate the finite sample performance of the test statistics M_1 , M_2 and especially the performance of M_3 . Also four other multivariate multi-sample test statistics were included in the study: Lawley-Hotelling T^2 , L_N based on component-wise ranks (see [6]) and W_{ϕ_1} , W_{ϕ_2} based on spatial signs (see [5]). Only M_3 and W_{ϕ_1} do not require underlying distributions of the same type, however, W_{ϕ_1} needs spherical symmetry.

We have been generating $q = 3$ samples of $n_1 = n_2 = n_3$ data points from 3-variate distributions (i.e. $d = 3$). The first sample has been generated from $N_3(\theta_1, I_3)$, the second and third from spherically symmetric Cauchy distributions (see [6]) with centers of symmetry θ_2 and θ_3 . If not stated otherwise, the location parameters $\theta_1, \theta_2, \theta_3$ were set to $(0, 0, 0)^T$. We have simulated three different cases, each of them 5000-times. The 5% critical value of χ_6^2 was used to reject H_0 . The results are in Table 1.

The simulated probabilities of Type I error of M_1 , M_2 and M_3 are slightly higher than the nominal level 5%. In case of M_3 it is just because of too

	M_1	M_2	M_3	T^2	L_N	W_{ϕ_1}	W_{ϕ_2}
H_0 true	.060	.064	.063	.029	.049	.049	0.057
$\theta_1 = (0.3, 0.3, 0)^T$.497	.500	.580	.038	.421	.567	.369
$\theta_2 = (0.3, 0.3, 0)^T$.480	.489	.485	.042	.315	.456	.283

Table 1: Simulated probabilities of Type I error and powers.

small sample sizes. For M_1 and M_2 it is because of underlying distributions of different type (compare with simulated probabilities of Type I error of M_1 and M_2 proposed in [5] in case of $n_1 = n_2 = n_3 = 100$ and underlying distributions of the same type).

If the first sample is shifted (i.e. the one with lower variability) the simulated powers of M_1 and M_2 lag behind the simulated powers of M_3 and W_{ϕ_1} . However, in case of shift in the second sample (here the variability is higher) the simulated powers of M_1 and M_2 are similar to that of M_3 . The presence of the heavy-tailed Cauchy distribution makes the performance of the Lawley-Hotelling T^2 really poor. Also note that the powers of L_N and W_{ϕ_2} are significantly smaller than the powers of M_3 and W_{ϕ_1} .

4 Conclusions

The test statistic M_3 based on spatial medians turns out to be a quite robust tool for testing the multivariate multi-sample location problem when the underlying distributions are not of the same type. Some non-parametric multivariate test statistics are computationally intensive and, therefore, not easy to use for larger data sets. Thanks to Weiszfeld's spatial median algorithm M_3 can be obtained quite quickly. Also its χ^2 -approximation seems to work already for relatively small sample sizes.

As the simulations suggest, the violation of the assumption of underlying distributions of the same type does not necessarily mean a poor performance of tests based on that assumption. Nevertheless, M_3 ensures that the favourable properties of the spatial median test statistics M_1, M_2 remain valid also in more general situations.

Appendix

Proof of Theorem 1:

M_3 can be written in the matrix form:

$$M_3 = \begin{pmatrix} \sqrt{n_1}(\hat{\mu}_1 - \tilde{\mu}) \\ \vdots \\ \sqrt{n_q}(\hat{\mu}_q - \tilde{\mu}) \end{pmatrix}^T \underbrace{\begin{pmatrix} \hat{V}_1^{-1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \hat{V}_q^{-1} \end{pmatrix}}_{=: \hat{W}^{-1}} \underbrace{\begin{pmatrix} \sqrt{n_1}(\hat{\mu}_1 - \tilde{\mu}) \\ \vdots \\ \sqrt{n_q}(\hat{\mu}_q - \tilde{\mu}) \end{pmatrix}}_{=: Z}.$$

In [1] it was shown that $\hat{V}_a = V_a + o_P(1)$. It means that $\hat{W}^{-1} = W^{-1} + o_P(1)$, where

$$W := \begin{pmatrix} V_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & V_q \end{pmatrix}.$$

Further, Z can be rewritten into the form

$$Z = \underbrace{\left[I_{qd} - \begin{pmatrix} \sqrt{\frac{n_1}{n}} I_d \\ \vdots \\ \sqrt{\frac{n_q}{n}} I_d \end{pmatrix} \left(\sqrt{\frac{n_1}{n}} n \hat{S}^{-1}, \dots, \sqrt{\frac{n_q}{n}} n \hat{S}^{-1} \right) \hat{W}^{-1} \right]}_{=: \hat{B}} \cdot X,$$

where $X := (\sqrt{n_1}(\hat{\mu}_1 - \mu)^T, \dots, \sqrt{n_q}(\hat{\mu}_q - \mu)^T)^T$, μ is the common value of μ_1, \dots, μ_q . Assumption 2 and the fact that $\hat{V}_a = V_a + o_P(1)$ imply that $\hat{B} = B + o_P(1)$ where

$$B := I_{qd} - (\sqrt{p} \otimes I_d)(\sqrt{p}^T \otimes R^{-1})W^{-1},$$

$\sqrt{p} := (\sqrt{p_1}, \dots, \sqrt{p_q})^T$, $R := \sum_{a=1}^q p_a V_a^{-1}$ and \otimes denotes the Kronecker product. Summarizing the above asymptotic results about \hat{W} and \hat{B} we obtain that

$$M_3 = X^T (B + o_P(1))^T (W^{-1} + o_P(1)) (B + o_P(1)) X = X^T B^T W^{-1} B X + o_P(1),$$

where the second equality follows from the fact $X = O_P(1)$ (ensured by (2)). It is easy to verify that

$$(\sqrt{p}^T \otimes I_d) W^{-1} (\sqrt{p} \otimes R^{-1}) = I_d, \quad (4)$$

which implies that

$$M_3 = X^T \underbrace{[W^{-1} - W^{-1}(\sqrt{p} \otimes R^{-1})(\sqrt{p}^T \otimes I_d)W^{-1}]}_{=: A} X + o_P(1).$$

From (2) we have that asymptotically $X \sim N_{qd}(0, W)$. Hence, to show that the asymptotic distribution of the quadratic form $X^T A X$ is $\chi_{(q-1)d}^2$ it is sufficient to prove that $W A W A W = W A W$ and $\text{trace}(A W) = (q-1)d$. These two equalities can be verified by a straightforward computation making use of (4).

Proof of Theorem 2:

Let V denotes the common value of V_1, \dots, V_q . Then $W = I_q \otimes V$, $R = V^{-1}$ and one easily obtains that $A = (I_q - \sqrt{p} \sqrt{p}^T) \otimes V^{-1}$. But for this form of

the matrix A it was shown (see the proofs in [5]) that $M_1 = X^T AX + o_P(1)$ and the proof is complete.

Proof of Theorem 3:

As in the proof of Theorem 1 one obtains that $M_3 = X^T AX + o_P(1)$. But here the asymptotic distribution of X is $N_{qd}(h^*, W)$, where

$$h^* := (\sqrt{p_1}h_1^T, \dots, \sqrt{p_q}h_q^T)^T.$$

It was already shown that $WAWAW = WAW$ and $\text{trace}(AW) = (q-1)d$. To complete the proof we have to verify that $WAh^* \in \mathcal{M}(WAW)$ and $h^{*T}Ah^* = h^{*T}AWAh^*$, where $\mathcal{M}(\cdot)$ denotes the linear subspace spanned by the columns of the matrix. The first condition is satisfied because W is regular and therefore we can write $WAh^* = WAWW^{-1}h^* \in \mathcal{M}(WAW)$. The validity of the second condition follows from the equality $AWA = A$ which can be verified by (4). The noncentrality parameter is given by $\delta = h^{*T}AWAWAh^*$. From (4) it is easy to see that $AWAWA = A$. Further, we apply (3) and obtain that $\delta = \sum_{a=1}^q p_a h_a^T V_a^{-1} h_a$.

References

- [1] Bai Z.D., Chen X.R., Miao B.Q., Rao C.R. (1990). *Asymptotic theory of least distances estimate in multivariate linear models*. *Statistics* **21**, 503–519.
- [2] Chaudhuri P. (1992). *Multivariate location estimation using extension of R-estimates through U-statistics type approach*. *Ann. Statist.* **20**, 897–916.
- [3] Milasevic P., Ducharme G.R. (1987). *Uniqueness of the spatial median*. *Ann. Statist.* **15**, 1332–1333.
- [4] Rublík F. (2001). *Tests of some hypotheses on characteristic roots of covariance matrices not requiring normality assumptions*. *Kybernetika* **37**, 61–78.
- [5] Somorčík J. (2006). *Tests using spatial median*. *Austrian Journal of Statistics* **35**, 331–338.
- [6] Um Y., Randles R.H. (1998). *Nonparametric tests for the multivariate multi-sample location problem*. *Statistica Sinica* **8**, 801–812.
- [7] Vardi Y., Zhang C.H. (2000). *The multivariate L_1 -median and associated data depth*. *Proceedings of the National Academy of Science USA* **97**, 1423–1426.

Acknowledgement: The work has been supported by the VEGA grant No. 1/0077/09 of the Science Grant Agency of the Slovak Republic.

Address: J. Somorčík, FMFI UK, KAMŠ, Mlynská dolina, 842 48 Bratislava;
F. Rublík, ÚM SAV, Dúbravská cesta 9, 841 04 Bratislava

E-mail: somorcik@fmph.uniba.sk, umerrubl@savba.sk

STOCHASTICKÝ MODEL VÝVOJE EPIDEMIE S VAKCINACÍ

Jakub Staněk

Klíčová slova: Model vývoje epidemie, simulace, stochastická diferenciální rovnice, vakcinace.

Abstrakt: Tento příspěvek se zabývá stochastickým modelem vývoje epidemie, který je přirozenou stochastickou verzí Kermack-McKendrickova modelu s vakcinací. Tento model je popsán stochastickou diferenciální rovnicí (R3). V příspěvku jsou porovnány střední hodnoty řešení (R3) s řešením deterministického modelu, je diskutováno chování modelu v závislosti na velikosti populace a na závěr je prezentováno simulační řešení optimalizace lineární vakcinace.

Abstract: This contribution concerns a stochastic model for the spread of an infection. This model is a natural version of the Kermack-McKendrick model with vaccination and it is given by the stochastic differential equation (R3). Expected value of a solution to (R3) and a solution to the deterministic model are compared, behavior of a solution to (R3) depending on the size of population is discussed and finally, a solution of the optimizing of linear vaccination obtained by simulations is presented.

1 Úvod

S rozvojem cestování po světě v současné době roste nebezpečí přenosu různých chorob, a tudíž je nutné neustále zlepšovat léčebné postupy či volit nové strategie prevence. Nedílnou součástí vývoje léčebných postupů je zlepšování našich znalostí o chování dané choroby, k čemuž nám mohou pomoci matematické modely chování vývoje epidemie.

Ačkoliv matematických modelů vývoje epidemie je mnoho, většina modelů nám dává informaci pouze o chování epidemie při daných vstupních podmínkách (počet nemocných a zdravých jedinců v určitý časový okamžik, intenzita přenosu choroby nebo léčení atd.), ale neumožňuje nám zasahovat do vývoje epidemie vakcinací. Model, který je zde prezentován, byl navržen tak, aby nám umožňoval sledovat, jak vakcinace ovlivňuje průběh epidemie, umožňuje nám porovnávat různé strategie vakcinace a hledat optimální strategii.

V první části článku je uveden Kermack-McKendrickův model (viz [2]) a jeho verze s vakcinací, ze které stochastický model vychází. V druhé části je pak představen stochastický model, provedeno srovnání numerických studií tohoto modelu s deterministickým modelem s vakcinací a provedena numerická optimalizace lineární vakcinace vzhledem ke ztrátové funkci. Všechny tři uvedené modely popisují chování vysoce infekční nemoci s krátkou dobou

uzdravování (typickým onemocněním tohoto typu je chřipka). V těchto modelech díky poměrně krátkému trvání epidemie předpokládáme konstantní velikost populace, která je rozdělena na tři podpopulace, a to na skupiny jedinců zdravých, nemocných a imunních.

2 Deterministické modely

Předpokládejme konstantní velikost populace n , která je rozdělena do tří skupin, jejichž velikost se v čase mění. Tyto skupiny jsou označeny jako x_t popisující počet jedinců, kteří jsou zdraví, ale mohou být nakaženi (dále jen „zdraví“), y_t značící počet nakažených jedinců, kteří jsou přenašeči nemoci (dále jen „nemocní“), a z_t , což je počet jedinců, kteří nemoc již prodělali, nemohou být znovu nakaženi a nejsou infekční (dále jen „imunní“). Kermack-McKendrickův model, který předpokládá homogenní populaci (ve smyslu výše imunity jedinců a množství kontaktů jednotlivých osob s ostatními osobami), je popsán diferenciální rovnicí

$$\begin{aligned} dx_t &= -\beta y_t x_t dt, & x_0 &> 0, \\ dy_t &= \beta y_t x_t dt - \gamma y_t dt, & y_0 &> 0, \\ dz_t &= \gamma y_t dt, & z_0 &= 0, \end{aligned} \quad (\text{R1})$$

kde $\beta > 0$ je intenzita přenosu nemoci, $\gamma > 0$ je intenzita uzdravování (γ^{-1} je tedy průměrná doba trvání nemoci) a $x_t y_t$ je počet všech dvojic v čase t , ve kterých může dojít k přenosu nemoci, tudíž $\beta x_t y_t$ je počet dvojic v čase t , v nichž dojde k nakažení jednoho jedince od druhého. Podrobněji se tímto modelem zabývají např. [1], [2] nebo [4].

Předpoklad homogenity populace v tomto modelu není příliš realistický, neboť intenzita vzájemných kontaktů mezi jedinci v populaci je různá a rovněž stupeň imunity jednotlivých jedinců není stejný. Starší lidé a děti totiž bývají většinou náchylnější k nemocem než dospělí jedinci v produktivním věku. Předpoklad homogenity můžeme však odstranit, nahradíme-li konstantu β funkcí $\beta(z)$, jak je ukázáno v [4].

Pokud budeme uvažovat obecnější intenzitu přenosu nemoci β a přidáme do předchozího modelu přirozeným způsobem vakcinaci, dostaneme model určený diferenciální rovnicí

$$\begin{aligned} dx_t &= -\beta(z_t) y_t [x_t - \vartheta(z_t)]^+ dt, & x_0 &> 0, \\ dy_t &= \beta(z_t) y_t [x_t - \vartheta(z_t)]^+ dt - \gamma y_t dt, & y_0 &> 0, \\ dz_t &= \gamma y_t dt, & z_0 &= 0, \end{aligned} \quad (\text{R2})$$

kde $\beta(z_t)$ je intenzita přenosu nemoci v čase t , $\gamma > 0$ je intenzita uzdravování a $\vartheta(z_t)$ je počet navakcinovaných jedinců v čase t . Výraz $y_t [x_t - \vartheta(z_t)]^+$ tedy značí počet všech dvojic tvořených zdravým nenavakcinovaným jedincem a jedincem nakažených, tj. všech dvojic, ve kterých může dojít k přenosu nemoci. Jelikož lidé s nižším stupněm imunity bývají nakaženi jako první, může se

v průběhu epidemie průměrný stupeň imunity jednotlivých zdravých jedinců zvětšovat, a tudíž by se mohla zmenšovat intenzita přenosu nemoci. Proto volíme intenzitu přenosu nemoci $\beta(z) : \mathbf{R} \rightarrow \mathbf{R}^+$ jako nerostoucí spojitou funkci imunních jedinců. Jelikož z_t , tj. počet jedinců, kteří nemoc prodělali, nám dává informaci o tom, jak moc je nemoc rozšířená, volíme počet navakcinovaných jedinců $\vartheta(z) : \mathbf{R} \rightarrow \mathbf{R}^+$ jako neklesající spojitou funkci imunních jedinců. Pokud například zvolíme ϑ lineární funkci, pak budeme vakcinovat nejvíce v situaci, kdy máme nejvíce nemocných jedinců, jelikož přírůstek imunních jedinců je přímo úměrný počtu nemocných jedinců.

Podrobněji je tento model popsán v [3].

3 Stochastický model

Tento model je přirozenou stochastickou verzí předchozího deterministického modelu. Uvažujeme stejně jako v předchozím modelu konstantní velikost populace N , která je rozdělena na zdravé (X_t), nemocné (Y_t) a imunní jedince (Z_t). Model je popsán stochastickou diferenciální rovnicí

$$\begin{aligned} dX_t &= -\beta(Z_t)Y_t[X_t - \vartheta(Z_t)]^+ dt + \sqrt{\beta(Z_t)Y_t[X_t - \vartheta(Z_t)]^+} dW_t, \\ dY_t &= \beta(Z_t)Y_t[X_t - \vartheta(Z_t)]^+ dt - \gamma Y_t dt - \sqrt{\beta(Z_t)Y_t[X_t - \vartheta(Z_t)]^+} dW_t, \\ dZ_t &= \gamma Y_t dt, \end{aligned} \tag{R3}$$

s počáteční podmínkou $X_0 > 0$, $Y_0 > 0$, $Z_0 = 0$.

Intenzity γ i funkce $\beta(z)$ a $\vartheta(z)$ zde mají stejný význam jako v předchozím deterministickém modelu a W_t značí Wienerův proces.

K volbě difúzního koeficientu ve tvaru $\sqrt{\beta(Z_t)Y_t[X_t - \vartheta(Z_t)]^+}$ vedla následující úvaha. Pokud máme rozsah populace veliký a předpokládáme, že každý zdravý nenavakcinovaný jedinec může být nakažen se stejnou pravděpodobností, pak má počet nově nakažených jedinců během časového intervalu $[t, t + \Delta]$ přibližně Poissonovo rozdělení s parametrem $\lambda = Y_t[X_t - \vartheta(Z_t)]^+ \Delta$. Difúzní koeficient byl proto volen tak, aby rozptyl počtu nově nakažených jedinců v daném časovém intervalu byl také přibližně λ .

Označme

$$\begin{aligned} \tau_X &:= \inf\{t \geq 0 : X_t = \vartheta(Z_t)\}, \\ \tau_Y &:= \inf\{t \geq 0 : Y_t = 0\}, \\ \tau &:= \min\{\tau_X, \tau_Y\}, \end{aligned}$$

kde τ_X je první čas, kdy jsou všichni zdraví jedinci vakcinováni, τ_Y je první čas, kdy v celé populaci není žádný nemocný jedinec, a τ je první čas, v němž se nemoc nemůže dále šířit.

Pokud uvažujeme $\beta(z)$ a $\vartheta(z)$ lipschitzovské funkce, pak lze při drobné úpravě důkazu věty 3.1 v [5] dokázat, že rovnice (R3) má jediné silné řešení (X_t, Y_t, Z_t) , které splňuje, že pro každé $t \geq \tau$

$$\begin{aligned} X_t &= X_\tau, \\ Y_t &= Y_\tau e^{-\gamma t} - \tau, \\ Z_t &= Z_\tau. \end{aligned}$$

Uvažujme nulovou vakcinační funkci ϑ , konstantní intenzitu přenosu nemoci β , (X_t, Y_t, Z_t) řešení rovnice (R3) a označme $(\tilde{X}_t, \tilde{Y}_t, \tilde{Z}_t) = (\frac{X_t}{N}, \frac{Y_t}{N}, \frac{Z_t}{N})$ proces poměrů velikostí jednotlivých skupin v čase. Pak

$$\begin{aligned} \tilde{X}_t &= \frac{x_0}{N} - \int_0^t \beta N \left[\frac{X_s}{N} \frac{Y_s}{N} \right]^+ ds + \int_0^t \sqrt{\beta \left[\frac{X_s}{N} \frac{Y_s}{N} \right]^+} dW_s \\ &= \tilde{x}_0 + \int_0^t \tilde{\beta} [\tilde{X}_s \tilde{Y}_s]^+ ds + \int_0^t \sqrt{\frac{\tilde{\beta}}{N} [\tilde{X}_s \tilde{Y}_s]^+} dW_s. \end{aligned}$$

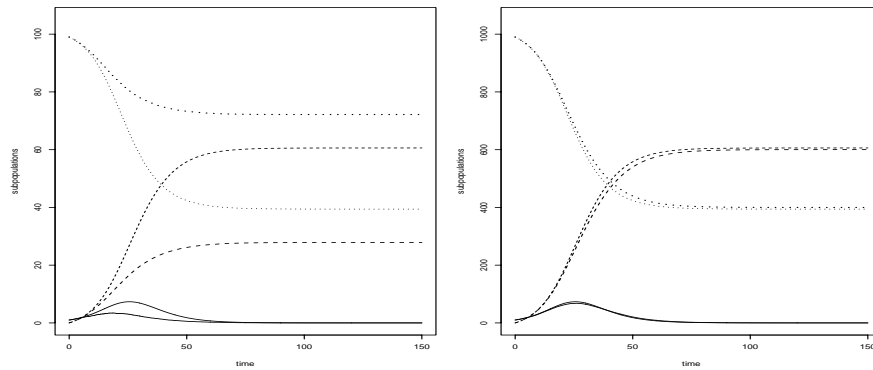
Obdobně lze přepsat také rovnice pro \tilde{Y}_t a \tilde{Z}_t . Podíváme-li se na tyto rovnice, vidíme, že proces $(\tilde{X}_t, \tilde{Y}_t, \tilde{Z}_t)$ není řešením rovnice (R3), neboť difúzní koeficient je oproti trendovému koeficientu \sqrt{N} -krát menší, než by měl být. Tento fakt lze interpretovat tak, že se zvětšující se populací vliv difúzního koeficientu slábne.

Na Obrázcích 1 a 2 vidíme, při jakých velikostech populace jsou ještě rozdíly mezi deterministickým modelem a jeho stochastickou verzí podstatné a kdy již začínají řešení těchto modelů splývat.

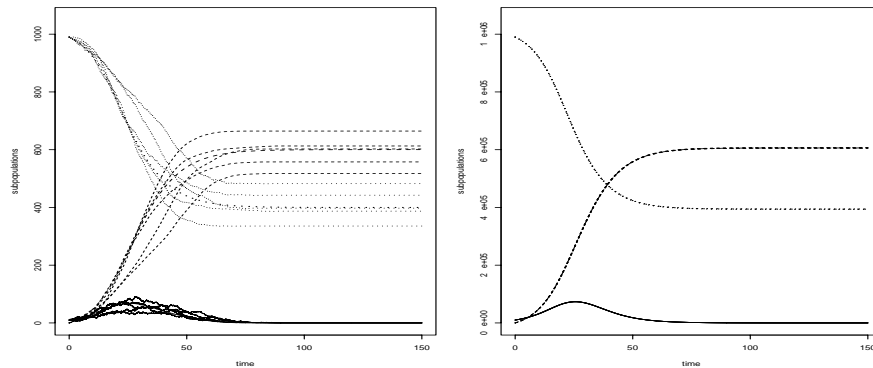
4 Simulace

V následujících ukázkách simulací bylo zvoleno $\gamma = 0.25$, $\beta = 0.38$ (Obrázek 1 a Obrázek 2), resp. $\beta = 0.3$ (Obrázek 3), a počáteční podmínka $Y_0 = 0.01 * N$ (tj. na počátku trpí chorobou jedno procento populace). Doba, pro kterou byla simulace provedena, je 150 časových jednotek (např. dní). V prvních dvou obrázcích byla (pro větší názornost) zvolena nulová vakcinace.

První dva obrázky ukazují zmenšující se rozdíl mezi řešením deterministického modelu a jeho stochastické verze se zvětšující se velikostí populace. V posledním obrázku jsou znázorněny střední hodnoty nákladů v závislosti na koeficientech lineární vakcinace vzhledem k dvěma různým penalizačním funkcím při velikosti populace $N = 1000$. V tomto příkladu jsme hledali minimální náklady v případě, kdy nás každá léčebná procedura stojí finanční jednotku (f.j.), předvakcinování jednoho jedince nás stojí 0.3 resp. 0.6 f.j., navakcinování jednoho jedince v průběhu epidemie nás stojí 0.4 resp. 0.304 f.j. V levém obrázku navíc uvažujeme, že každý den, kdy je v populaci více nemocných než 0.3%, nás stojí 0.5 f.j. Zatímco koeficienty u penalizační funkce vlevo by mohly být zvoleny realisticky, volba koeficientů penalizační funkce vpravo, kdy předvakcinace stojí dvakrát více než vakcinace v průběhu nemoci, zřejmě moc realistická není. Tato volba ale prezentuje zajímavý hraniční stav,

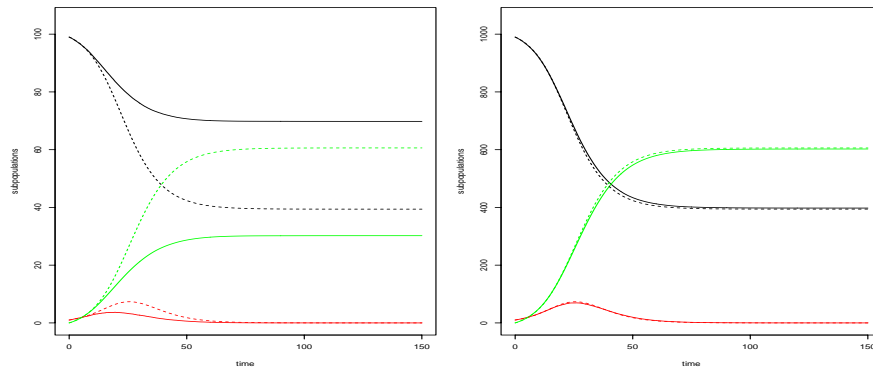


Obrázek 1: Rozdíl mezi řešením deterministického modelu (slabá čára) a střední hodnotou řešení stochastického modelu (silná čára). Zatímco pro populaci čítající 100 jedinců je rozdíl zřetelný (graf vlevo), pro větší populaci složenou z 1000 jedinců (graf vpravo) je již rozdíl zanedbatelný. Počet zdravých jedinců je znázorněn tečkovanou čarou, nemocných jedinců plnou čarou a imunních jedinců čárkovanou čarou.



Obrázek 2: Střední hodnota řešení stochastického modelu (silná čára) a pět realizací tohoto modelu (slabá čára). Zatímco pro menší populaci (1000 jedinců, graf vlevo) se průběh jednotlivých realizací podstatně liší od jejich střední hodnoty, pro velkou populaci (1 milion jedinců, obrázek vpravo) jsou všechny průběhy podobné a téměř splývají s řešením deterministického modelu.

kdy se optimální vakcinační strategie, tj. volba koeficientů ϑ_0 a ϑ_1 tak, aby střední hodnota penalizační funkce byla co nejmenší, přesouvá z osy $\vartheta_1 = 0$ na osu $\vartheta_0 = 0$.



Obrázek 3: Různá efektivita lineární vakcinace v závislosti na zvolených parametrech. Funkci počtu navakcinovaných jedinců volíme lineární $\vartheta(z) = v_0 + v_1 * z$ a penalizační funkci volíme ve tvaru $f = E(Z_{150} + 0.3 * v_0 + 0.4 * v_1 * Z_{150} + 0.5 * T)$ (vlevo) a $f = E(Z_{150} + 0.6 * v_0 + 0.304 * v_1 * Z_{150})$ (vpravo), tj. je penalizován počet jedinců, kteří nemoc prodělali, doba trvání epidemie, předvakcinace a vakcinace v průběhu epidemie z důvodů různých nákladů na způsoby vakcinace a léčby. Z_{150} je počet jedinců, kteří do stanoveného času prodělali nemoc, v_0 je počet předvakcinovaných jedinců, $v_1 * Z_{150}$ je počet jedinců navakcinovaných v průběhu času a T je délka doby, po kterou bylo nemocných více než 0.3% populace.

5 Závěr

V uvedené studii jsme si ukázali, že při vzrůstající velikosti populace se vliv náhody v našem modelu zmenšuje. Tento model má tedy smysl pro modelování vývoje epidemie v malých populacích. Pro velké populace (1 000 000 jedinců) stačí použít deterministický model, neboť stochastický model nám oproti deterministickému modelu žádné nové informace nepřináší.

Literatura

- [1] Daley D.J., Gani J. (1999). *Epidemic modelling: an introduction*. Cambridge University Press.
- [2] Kermack W.O., McKendrick A.G. (1927). *Contributions to the mathematical theory of epidemics*. Proc. Roy. Soc. Lond. A **115**, 700–721.
- [3] Staněk J. (2008). *Kermack-McKendrick epidemics vaccinated*. Kybernetika, přijato k publikaci.
- [4] Štěpán J., Hlubinka D. (2007). *Kermack-McKendrick epidemic model revisited*. Kybernetika **43**, 395–414.
- [5] Štěpán J., Staněk J. (2008). *Absorption in stochastic epidemics*. Kybernetika, v recenzním řízení.

Poděkování: Tato práce byla podporována grantem GAČR 201/05/H007.

Adresa: J. Staněk, MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8 – Karlín

E-mail: stanekj@karlin.mff.cuni.cz

JAK POSUZOVAT SPOLEHLIVOST SOFTWARE

Jan A. Strouhal

Klíčová slova: Spolehlivost softwaru, klasifikační strom, metoda plovoucích mezí, pětistupňová škála.

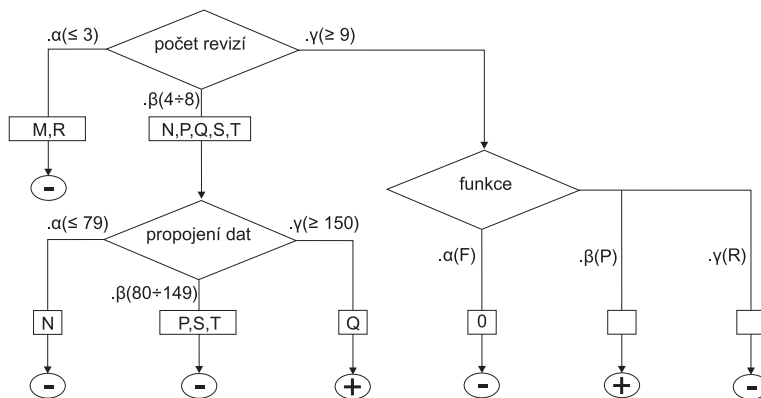
Abstrakt: Aplikace klasifikačních a regresních stromů v oblasti zvyšování spolehlivosti softwaru poskytuje především nástroj pro zvýšení efektivity tohoto procesu. Činí to vhodným výběrem modulů s vysokým očekávaným počtem chyb. Odhadem jejich rizikovosti lze snížit náklady na testování. V předchozích pracích jsou popsány nové vyvíjené metody, které usnadní a zlepší práci při tvorbě klasifikačního stromu pro softwarovou spolehlivost. Jsou to metody rozšíření základní klasifikační škály na více skupin a metoda vyhledávání optimálních hranic dělení použitých softwarových metrik. Přínos těchto metod je ukázán na konečném příkladu. Pro snadnější výpočty byl vyroben program pro hledání optimálních kořenů metrik při konstrukci těchto klasifikačních stromů. Důležitými vstupními údaji tohoto programu jsou softwarové metriky, hodnoty metrik, moduly a třídy modulů. Uživatelé jsou umožněni získat mezivýsledky, jako jsou funkce výběru metrik F a konečný výsledek, kterým je nalezení minimálního E (jakési míry chaosu) sloužícího k rozhodnutí o optimálním kořenu metriky. Nastavení programu je dostatečně variabilní. Program mi velice pomohl při simulacích i výpočtech na skutečných datech.

Abstract: Application of classification and regression trees in software reliability area give us a tool for increase efficiency in software reliability process. It has been done by proper choosing modules with expecting large faults count. Test cost can be reduce by estimation of risk. Previous articles describes new developed methods, which can make building software reliability classification trees easy. These methods are: extension of basic classification class to more groups method and finding optimal bounds for used software metrics split method. Contribution of these methods are shown on final example. We develop software utility for compute optimal metric for root. Utility has input variables as software metrics, metric values, modules and module classes. Utility has also semioutput variables as metric selection function F and output variable minimum E (chaos degree). Minimum E is used for optimal metric root decision. Utility options are sufficiently variable. This software was helpful in simulations as well as calculations on real data.

1 Úvod

V poslední době se využívá výpočetní technika téměř všude kolem nás. Existuje nepřeberné množství počítačových programů od jednoduchých her až po

složité programové systémy používané v jaderném inženýrství. Všechny programy procházejí svými životními fázemi od analýzy, návrhu, vývoje, ladění, používání až po opravy a dodatky. Koncový uživatel v závěrečné fázi, ve fázi užívání, požaduje, aby tento program byl maximálně spolehlivý, neboli aby bezchybně vykonával požadované funkce po danou dobu. A tak výrobce programu musí vyprodukovat natolik spolehlivý software, aby byl zákazník spokojen. Vyrobit naprosto bezchybný program je téměř nemožné, ale jsou cesty, které umožňují chybovost softwaru snížit. Jedna z cest je otestovat (někdy i opakovaně) všechny funkce programu za všech možných podmínek a při jakékoli opravě programu znovu vše podrobně otestovat. Tento postup je náročný nejen na pracovní síly a na prostředky, ale je zejména náročný na čas a toho je v dnešní době stále méně (zákazník na zadaný program většinou spěchá). A tak nastává otázka, zda by nebylo možné proces testování (popřípadě i ladění) urychlit. Odpověď je jednoduchá. Vždyť přece není nutné testovat části programu, které se dle předchozích zkušeností jeví jako části *nenáchylné k chybám* neboli bylo v nich detekováno jen málo chyb a naopak zacílit úsilí na ty části programu, které se dle předchozích zkušeností jeví jako části, ve kterých byla spousta chyb, tedy části *náchylné k chybám*. Pokud tedy zjistíme charakter (vlastnosti) uvedených dvou skupin a podle těchto vlastností dokážeme určit náchylnost k chybám, nemusíme pak testovat části, které se jeví jako nenáchylné k chybám. Vezmeme-li části programu (moduly), které jsme již v minulosti testovali, můžeme z jejich vlastností (metrik - termín ve spolehlivosti softwaru označuje něco jako měřitelnou vlastnost, atribut) zkonstruovat klasifikační strom, kterým pak můžeme klasifikovat části programu, u kterých chceme zjistit, zda je máme či nemusíme testovat. Takto zkonstruovaný strom nám může pomoci zkrátit dobu testování a v důsledku i zvýšit spolehlivost, protože se při testování zaměříme více na části, které jsou skutečně potřeba otestovat a odladit.



Obrázek 1: Příklad klasifikačního stromu.

2 Příklad klasifikačního stromu

Ukážeme názorný příklad klasifikace modulu Q v již vytvořeném stromě (viz obrázek obsahující moduly M až T). V příkladu jsou použity tři softwarové metriky (atributy): počet revizí (v rozsahu 0 až 15), propojení dat (v rozsahu 0 až 200) a funkce modulu (možnosti R - uživatelské rozhraní, P - řízení procesu a F - práce se soubory). Strom je v příkladu znázorněn vývojovým diagramem, čitelným pro softwarové vývojáře, jednotlivé větve jsou označeny α , β , γ . Modul Q má tyto atributy: počet revizí 7, propojení dat 172 a funkci P (řízení procesu). V prvním kroku jej tedy klasifikujeme do větve β dle metriky počet revizí a následně do větve γ dle metriky propojení dat.

Pomocí klasifikace jsme zjistili, že klasifikovaný modul Q je pravděpodobně náchylný k chybám (skončil v + listu) a měli bychom jej testovat.

3 Metodika tvorby klasifikačního stromu

Při vytváření klasifikačních stromů používáme údaje o vlastnostech modulů zjištěné na základě pozorování podobných objektů. Na rozdíl od regresních stromů zde nepoužíváme binární dělení, ale uzel se může dle hodnot jednoho atributu rozdělit na více poduzlů tím, že je použito více hranic dělení. Obecný postup při konstrukci klasifikačních stromů začíná rozkladem systému na moduly, ohodnocením do skupin, volbou množiny metrik (atributů), volbou hranic metrik. Pak definujeme optimalizační pravidlo.

Pomocí funkce výběru metrik a váhy na poduzlech najdeme vhodné dělení. *Funkce výběru metrik* F nabývá hodnot v rozsahu $(0, 1)$ a je pro dvouskupinové dělení definována vztahem

$$F(p_i, n_i) = -\frac{p_i}{p_i + n_i} \log_2 \frac{p_i}{p_i + n_i} - \frac{n_i}{p_i + n_i} \log_2 \frac{n_i}{p_i + n_i}$$

kde p_i je počet pozitivních modulů (náchylných k chybám), a n_i je počet negativních modulů v i -tém uzlu. Čím je F menší, tím větší je homogenita v uzlu. Pro každý uzel dělení dále vypočítáme váhu $w_i = \frac{p_i + n_i}{|C|}$, kde $|C|$ je počet všech modulů. Atribut A , pro nějž je $E(C, A) = \sum_{i=1}^v w_i \cdot F(p_i, n_i)$ minimální použijeme jako nový kořen pro další zpracování. Dělení provádíme pro každý uzel až do kritéria terminace nebo do vyčerpání všech atributů.

3.1 Rozklad na moduly

Softwarový systém rozložíme na řadu modulů. Lepší je větší počet modulů, ale dbáme na to, aby moduly stále byly dostatečně samostatnými částmi softwarového systému.

3.2 Ohodnocení do skupin

Pak také zvolíme ohodnocení do skupin (v tomto případě dvou: náchylné a nenáchylné k chybám) libovolným spolehlivostním kritériem (faktorem), například určitým počtem chyb. Pokud bylo v modulu při testování zjištěno méně chyb, zařadíme jej do skupiny nenáchylné k chybám, pokud bylo zjištěno alespoň tento určitý počet chyb při testování, zařadíme jej do skupiny náchylné k chybám.

3.3 Zavedení, výběr a volba metrik

Další částí tvorby klasifikačního stromu je, že na všech modulech zavedeme množinu metrik, podle kterých bude prováděna klasifikace. Změříme hodnoty všech metrik pro všechny moduly. Správná volba metrik, ze kterých bude strom tvořen je důležitou částí. Všechny metriky, které bychom chtěli použít, by měli mít vždy nějaký vliv na počet chyb v modulu. Pokud použijí například počet úhozů do klávesnice při psaní programového kódu, logicky použijí nesprávnou metriku, protože tato metrika nemá na počet chyb v modulu žádný vliv. Závislost nemusí být samozřejmě přímo odvoditelná, to bychom to měli jednoduché a mohli spočítat spolehlivost software přímo z této závislosti, ale měli bychom přemýšlet o závislosti metriky na možný výskyt chyb. Pro některé typy úloh a programů se může zdát vhodné použít některé metriky, pro jiné typy úloh a programů zase jiné. Výběr se ponechává na člověku tvořícímu analýzu softwarové spolehlivosti. Lépe je upřednostňovat ty metriky, které lze snadno a dostupnými prostředky přesně zjistit (počet řádků programového kódu v modulu). Počet zvolených a naměřených metrik by neměl být malý, ale neměly by také být podobné ve své podstatě. Nedá se předem zjistit, jaký je optimální počet metrik, které bychom měli ke tvorbě stromu použít, ale prakticky je téměř vždy možné použít všechny metriky, které jsme na modulech měřili. Výjimku tvoří navzájem redundantní metriky. Pokud jsou dvě metriky navzájem zastupitelné vychází při tvorbě stromu stejné nebo hodně podobné $E()$ pro obě metriky u všech výpočtů a obě metriky vždy rozdělují moduly do stejných nebo hodně podobných skupin. Tím je i naznačena určitá korelace mezi těmito metrikami, jedna metrika je redundantní a je nutné ji ze seznamu vyřadit. Většinou jsou tyto metriky vyřazeny předem dle logické úvahy o jejich fyzické nebo matematické podstatě. Příkladem může být metrika počtu řádků kódu v modulu a počet znaků kódu v modulu. Protože bývá počet znaků kódu na řádek kódu průměrně ve všech modulech stejné číslo, je mezi těmito dvěma metrikami jistá lineární závislost a není nutné používat k tvorbě stromu obě tyto metriky.

3.4 Volba hranic

Abychom mohli moduly rozdělovat pomocí určité metriky z hlavního uzlu do poduzlů, je nutné si pro tuto metriku stanovit také počet dělení a hranice dělení. Počet dělení (počet hranic, kterými rozdělíme moduly) se většinou

stanovil na dvě, tím pádem se vždy jeden uzel dělil na tři poduzly. Expertní hodnoty hranic se většinou stanovovaly z předchozích zkušeností nebo dle nějakého obecného principu. Příkladem si uveďme volbu hranic pro metriku propojení dat. Na všech modulech jsme spočítali, že metrika propojení dat nabývá hodnot od 0 do 200, tudíž si můžeme například stanovit hranice 65 a 130 (tak abychom rozdělili interval 0 200 na přibližně tři stejné intervaly).

3.5 Vhodné dělení

V každém kroku tvorby stromu se snažíme vyhledat vhodné dělení uzlu na poduzly. Pomáhá nám k tomu funkce výběru metrik, která nám ukazuje na homogenitu v jednotlivých uzlech. Vynásobíme-li výsledky této funkce váhou (poměr počtu modulů v uzlu z celkového počtu zbývajících modulů) získáme $E()$. Metriku, pro niž je $E()$ minimální použijeme jako nový kořen pro další zpracování, tuto metriku vyřadíme ze seznamu metrik a pro všechny vzniklé poduzly opakujeme postup dělení. Dělení provádíme až do vyčerpání všech metrik nebo do splnění *kritéria terminace*.

4 Optimalizace při tvorbě klasifikačního stromu

Při vytváření klasifikačního stromu předem dané (expertní) hranice metrik nemusí být vždy optimální a mohou způsobovat vytvoření stromu se *slepými* koncovými listy (bez určení jakékoli výstupní skupiny) nebo s koncovými listy, které nejsou stále *homogenní* (s několika výslednými skupinami), přestože jsou již vyčerpány všechny metriky. Proto byly navrženy nové metody, které optimalizují tvorbu klasifikačního stromu.

4.1 Optimalizace počtu hranic výběrem nejlepšího počtu

Myšlenka této optimalizace je v zmenšení počtu hranic v okamžiku, kdy není vyšší počet nutný. Například při okamžiku slepého listu můžeme sousední hranice, kde slepý list vznikl sloučit do jedné. Pro tuto metodu máme jednoduchý základní princip. Vypočteme stromy pro všechny možné počty hranic a pozorujeme maximum $E()$ pro celý strom, počet použitých metrik a pozorujeme homogenitu v koncových listech (poměr homogenních listů ku všem listům).

- Neprovádíme v případech, kdy máme dva (a více) modulů se stejnými hodnotami všech metrik, ale s odlišnou skupinou.
- Pro všechny metriky volíme stejný počet hranic. Na počátku tvorby zvolíme jednu hranici a tento počet postupně zvyšujeme.
- Tvorbu stromů ukončíme ihned, jakmile budeme mít v právě prováděném stromu všechny koncové listy homogenními.
- Dodržujeme podmínku, že metrika nemůže mít více hranic, nežli je počet naměřených odlišných hodnot mínus jedna pro tuto metriku.

Pokud nastane tato podmínka, pro tuto metriku již počet hranic ne zvyšujeme.

- Pokud již nemáme možnost zvýšit počet hranic u žádné z metrik, tvorbu stromů ukončíme.

Na všech získaných stromech spočítáme použitelnost stromu jako počet použitých metrik, dělených násobkem $E()$ a homogenity v koncových listech (násobek poměrů vyššího zastoupení skupiny ke všem modulům v listu). Ze všech stromů pak vezmeme ten, který má tuto použitelnost stromu nejvyšší.

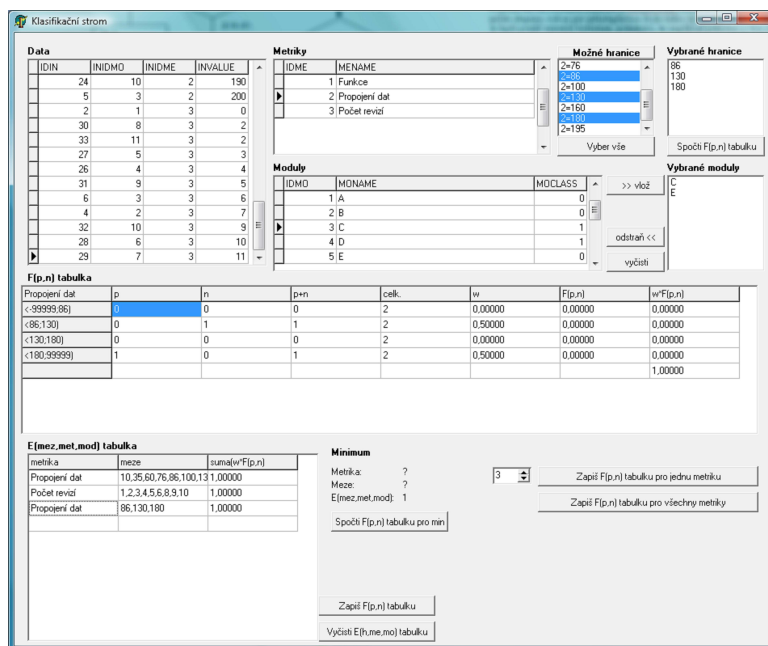
4.2 Optimalizace počtu hranic postupným zvyšováním

Tento postup je jakousi obdobou předchozího. Budeme postupně vytvářet klasifikační strom běžným způsobem s proměnným počtem hranic a budeme zaznamenávat maximum $E()$ pro celý strom a pozorovat homogenitu v koncových listech.

- Neprovádíme v případech, kdy máme dva (a více) modulů se stejnými hodnotami všech metrik, ale s odlišnou skupinou.
- Na počátku tvorby zvolíme jednu hranici pro všechny metriky.
- Tvorbu stromu ukončíme ihned, jakmile budeme mít všechny koncové listy homogenními.
- Pokud bude existovat nehomogenní koncový list, zvýšíme počet hranic o jedničku tam, kde nám vychází nejvyšší $E()$.
- Pokud bude existovat nehomogenní koncový list a ve stejné metrice jako v bezprostředně předchozí tvorbě stromu se nám objeví maximum $E()$, které je vyšší nebo rovno maximu při bezprostředně předchozí tvorbě stromu, vrátíme počet hranic v této metrice na bezprostředně předchozí. V této metrice již nikdy počet hranic nezvyšujeme a zvyšujeme počet hranic u metriky, která má druhé (popř. další) nejvyšší $E()$.
- Pokud bude existovat nehomogenní koncový list a ve stejné metrice, jako v bezprostředně předchozí tvorbě stromu se nám objeví maximum $E()$, které je nižší než při bezprostředně předchozí tvorbě stromu, zvýšíme počet hranic o jedničku u této metriky jen tehdy, pokud nám nevyšel stejný strom, jinak vrátíme počet hranic v této metrice na bezprostředně předchozí. V této metrice již nikdy počet hranic nezvyšujeme a zvýšíme počet hranic u metriky s druhým (popř. dalším) nejvyšším $E()$.
- Pokud při zvyšování počtu hranic v dané metrice narazíme na počet všech možných (odlišných) hodnot pro danou metriku mínus jedna, počet hranic u této metriky nezvyšujeme a zvyšujeme počet hranic u jiné.
- Nemáme-li již možnost zvýšit počet hranic u žádné z metrik, tvorbu stromu ukončíme.

5 Program - metoda plovoucích mezí

Při vytváření klasifikačního stromu předem dané (expertní) hranice metrik nejsou vždy optimální a způsobují vytvoření stromu se slepými koncovými listy nebo s koncovými listy, které jsou stále nehomogenní, přestože jsou již vyčerpány všechny metriky. Slepé listy můžeme přiřadit k vedlejší klasifikaci (jako by se vynechala hranice mezi listem a slepým listem) a tudíž použít méně hranic. Slepé listy můžeme také odstranit tím, že při vytváření klasifikačního stromu procházíme všechny možné meze (také i počet mezí) a dostaneme menší hodnoty $E()$ a tím i lepší dělení jednotlivých větví stromu. Tuto metodu nazveme metoda plovoucích mezí. Protože by bylo náročné procházet ručně všechny možné meze a zjišťovat vhodnou metriku a meze vytvořili jsme pomocný program, který propočítává minimum $E()$ přes všechny metriky. Program generuje tyto meze v polovinách mezi všemi naměřenými hodnotami pro danou metriku. Výsledkem (výstupem) programu je vyhledání optimální metriky pro kořen (přes minimum $E()$). Následně je pak možné vybrat pouze moduly jdoucí určitou větví a znovu určit optimální metriku pro další kořen.



Obrázek 2: Hlavní okno programu pro metodu plovoucích mezí.

F5(a,b,c,d,e) tabulka										
Propojení dat	a	b	c	d	e	a+b+c+d+e	celk.	w	F(a,b,c,d,e)	wF(a,b,c,d,e)
<-99999,195)	2	0	0	0	0	10	11	0,90909	0,72193	0,65630
<195,99999)	0	1	0	0	0	1	11	0,09091	0,00000	0,00000
										0,65630

E(mez,met,mod) tabulka		
metrika	meze	sumávF(p,n)
Propojení dat	10	0,65630
Propojení dat	35	0,80707
Propojení dat	60	0,84047
Propojení dat	76	0,84427
Propojení dat	96	0,82914

Minimum		1	
Metrika	Propojení d.	F(p,n) tabulku pro min.	Zapiš F(p,n) tabulku pro všechny metiky
Meze:	10	F(a,b,c,d,e) tab. pro min.	Zapiš F5(a,b,c,d,e) tabulku pro jednu metiku
E(mez,met,mod):	0,65630		Zapiš F5(a,b,c,d,e) tabulku pro všechny metiky

Obrázek 3: Část okna programu při rozšíření klasifikační šklály na 5 skupin.

6 Program - rozšíření klasifikační šklály na 5 skupin

Klasifikace na "pozitivní" a "negativní" moduly, uvažovaná v literatuře se v praxi ukazuje jako nedostatečná. Vhodným vylepšením klasifikace je rozšíření klasifikační šklály na více skupin podle důležitosti modulu a závažnosti následků jeho selhání. Takovým rozšířením může být například pětistupňová šklála (pět skupin):

1. netestovat
2. testovat jen namátkově; projít
3. testovat na hlavní funkce, zkráceně
4. testovat po určitou dobu
5. důkladně otestovat, dokud nedosáhneme požadované spolehlivosti, případně do stanovené doby

Na konci klasifikace samozřejmě nemusí být jen informace o nutnosti testování v její pětistupňové šklále, ale i informace o rozdělení závažnosti chyb. Další verzí pomocného programu pro výpočet je tedy jeho rozšířená verze, která umožňuje zadat klasifikaci modulů v pětistupňové šklále. Třidu modulu tedy do programu zadáváme v hodnotách 0,1,2,3,4, kde 0 přiřadíme netestovat až ke 4 přiřadíme důkladné otestování.

7 Závěr

V příspěvku jsem se zabýval konstrukcí klasifikačních stromů pro použití ve spolehlivosti softwaru. V první části jsem popsal základní konstrukci, která počítá s expertním počtem i hodnotami hranic metrik, v další části jsem popsal vylepšení této základní konstrukce vyhledáním optimálních hranic metrik. V příspěvku jsem také popsal rozšíření základní klasifikační šklály na více skupin. Program vytvořený pro tyto rozšíření přinesl ulehčení výpočtů.

Literatura

- [1] Strouhal J. (2008). *Spolehlivostní modely při tvorbě softwaru*. Disertační práce 2008.
- [2] Dohnal G. (2006). *Markovské modely spolehlivosti software*. Sborník mezinárodní konference ROBUST'06, JČMF, ISBN 80-7015-073-4, 453 – 460.

Poděkování: Tato práce byla řešena v rámci projektu MŠMT 1M06047 Centrum pro jakost a spolehlivost výroby.

Adresa: J.A. Strouhal, Centrum pro jakost a spolehlivost výroby ČVUT FSI, Ústav technické matematiky, 121 35 Praha 2, Karlovo náměstí 13

E-mail: jas@cqr.cz

MAXIMÁLNĚ VĚROHODNÉ ODHADY A LINEÁRNÍ REGRESE VE VÝBĚROVÝCH ŠETŘENÍCH

Michaela Šedová, Michal Kulich

Klíčová slova: Výběrová šetření, maximálně věrohodný odhad, lineární model.

Abstrakt: V klasické teorii výběrových šetření jsou předmětem studia parametry charakterizující konečnou populaci, jako např. úhrn nebo průměr N pevných hodnot. Příspěvek stručně popisuje situaci, kdy je vhodnější považovat pozorování za náhodné veličiny a zároveň je nutné brát v úvahu, že není k dispozici prostý náhodný výběr. Věnuje se výpočtu střední hodnoty v takovém případě, modifikuje metodu maximálně věrohodných odhadů a konkrétněji se zabývá lineárním modelem. Výsledky jsou ilustrovány na malé simulační studii.

Abstract: This paper assumes rather unusual situation in the survey sampling when observations are random variables and one does not have available simple random sample. Attention is concentrated on the calculation of the mean in linear model when maximum likelihood estimates are used. Obtained results are illustrated on small simulation study.

1 Motivace

Motivace zabývat se tímto tématem vyvstala na základě statistické analýzy projektu Accept [1], studie zaměřené na výzkum prevence šíření HIV v Africe a Thajsku. Jedná se o skupinově randomizovanou studii, při níž byla 48 komunitám náhodně přiřazena buď experimentální intervence (program intenzivního poradenství a testování na HIV) anebo intervence standardní (testování na klinikách). Jedním z vedlejších cílů studie je vyhodnocení rizikového chování v komunitách na základě průřezových dat získaných výběrovým šetřením. Nejprve byly náhodně vybrány domácnosti (každá se stejnou pravděpodobností), a poté byl ze všech členů vybrané domácnosti splňujících stanovená kritéria náhodně vybrán jeden respondent a zařazen do studie.

Příkladem rizikového chování může být konzumace alkoholu. Představme si, že bychom chtěli odhadnout průměrnou měsíční spotřebu alkoholu na osobu v dané komunitě. Je pravděpodobné, že členové početnějších rodin pocházejí z jiného sociálního prostředí než ti z rodin menších, což může mít na konzumaci alkoholu vliv. Abychom tento jev ve statistické analýze zohlednili, je nutné vzít v úvahu výběrové schéma, které vlastně nadhodnocuje počet členů malých domácností ve výběru a naopak podhodnocuje zastoupení členů domácností velkých.

V klasické teorii výběrových šetření jsou předmětem studia převážně parametry charakterizující konečnou populaci, jako např. úhrn nebo průměr N pevných hodnot. Avšak v projektu Accept, stejně tak jako v jiných studiích, bychom rádi výsledky zobecnili (na jiné populace, nebo i tutéž populaci v jiném čase). V takových situacích je vhodnější považovat pozorování za realizace náhodných veličin.

Tento přístup v teorii výběrových šetření není zcela běžný. V následujícím textu se pokusíme zformulovat a popsat běžné statistické úlohy, jako je např. odhad parametrů nebo testování hypotéz, je-li nutno brát v úvahu výběrové schéma. Více se budeme věnovat modifikaci maximálně věrohodných odhadů, zvláště pak lineárnímu modelu. Výběrové schéma, které uvažujeme, je oproti motivačnímu projektu Accept poněkud zjednodušené.

Mějme populaci o velikosti N . Poznamenejme, že tuto populaci považujeme za výsledek prostého náhodného výběru. Její rozsah N je pevná veličina. Výběrové schéma specifikujeme následujícím způsobem. Nechť je Y sledovaná náhodná veličina a W diskrétní náhodná veličina, která nabývá hodnot $\{1, 2, \dots, K\}$. V našem případě odpovídá W stratu v populaci. Píšeme

$$I_{ik} = \begin{cases} 1 & \text{je-li } W_i = k \\ 0 & \text{jinak.} \end{cases}$$

Označme N_k počet jedinců ve stratu k , $\sum_{k=1}^K N_k = N$. Nechť ξ_i jsou dichotomické náhodné veličiny značící, zda jedinec i byl zahrnut do výběru nebo ne a π_k pravděpodobnosti výběru jedince, který patří do strata k . Předpokládáme, že veličiny ξ_i jsou nezávislé, v každém stratu k tedy provádíme tzv. bernoulliůvský výběr (viz [2], str. 62) s konstantními pravděpodobnostmi

$$E(\xi_i | W_i = k) = P(\xi_i = 1 | W_i = k) = \pi_k, \quad \text{pro } i = 1, 2, \dots, N.$$

Z toho vyplývá, že velikost výběru je náhodná. Předpokládejme také, že W_i jsou pozorované pro všech N členů populace, zatímco Y_i pozorujeme pouze pro jedince z výběru, tj. pokud $\xi_i = 1$.

2 Odhad střední hodnoty pro dané výběrové schéma

Abychom ilustrovali základní úvahy, věnujme se nejprve stručně odhadu střední hodnoty náhodné veličiny Y . Značíme

$$EY = \theta = \sum_{k=1}^K P(W = k)\theta_k, \quad \text{kde } \theta_k = E(Y|W = k).$$

Uvažujme odhad střední hodnoty definovaný vztahem

$$\hat{\theta} = \frac{1}{N} \sum_{i=1}^N \sum_{k=1}^K \left(\frac{1}{\hat{\pi}_k} \xi_i Y_i \right) I_{ik}, \quad \text{kde } \hat{\pi}_k = \frac{1}{N} \sum_{i=1}^N \xi_i I_{ik}. \quad (1)$$

Odhad $\hat{\theta}$ je vážený průměr pozorovaných hodnot, kde váhy jsou převrácenou hodnotou empirických pravděpodobností výběru. Pro dané stratum k je pravděpodobnost výběru odhadnutá jako počet jedinců ve výběru, kteří patří do strata k , dělený celkovým počtem jedinců ve stratu k . Jedná se tedy vlastně o poměrový odhad (viz [2], str. 176).

Tvrzení 1. Předpokládejme, že vektory (Y_i, W_i, ξ_i) jsou nezávislé, stejně rozdělené (iid) a ξ_i je podmíněně nezávislé na Y_i , je-li dáno W_i , pro $i = 1, 2, \dots, N$. Předpokládejme, že $\text{var} Y_1 < \infty$. Potom

$$\sqrt{N}(\hat{\theta} - \theta) \xrightarrow{d} N(0, \Sigma), \quad (2)$$

kde

$$\Sigma = \text{var} Y_i + \sum_{k=1}^K P(W_i = k) \frac{1 - \pi_k}{\pi_k} \text{var}_k Y_i \quad (3)$$

a

$$\text{var}_k(Y_i) = \text{var}(Y_i | W_i = k).$$

Důkaz viz [3].

3 Maximálně věrohodné odhady

Nechť Y_i , $i = 1, 2, \dots, N$, jsou iid náhodné veličiny s hustotou $f(y, \theta)$, $\theta = (\theta_1, \dots, \theta_p) \in \Theta$. Klasický maximálně věrohodný odhad parametru θ se získá maximalizací logaritmu věrohodnostní funkce

$$L_N(\theta) = \sum_{i=1}^N L_i(\theta | y_i).$$

To vede na soustavu rovnic

$$\frac{\partial L_N(\theta)}{\partial \theta} = \sum_{i=1}^N \mathbf{U}_i(\theta) = \mathbf{0}, \quad (4)$$

kde

$$\mathbf{U}_i(\theta) = \left(\frac{\partial L_i(\theta | y_i)}{\partial \theta_j} \right)_{j=1}^p.$$

Nyní se znovu vraťme k výše popsanému výběrovému schématu. Na základě stejné úvahy jako při odhadu střední hodnoty modifikujeme soustavu rovnic (4) na

$$\mathbf{V}(\theta) = \sum_{i=1}^N \sum_{k=1}^K \frac{\xi_i}{\hat{\pi}_k} \mathbf{U}_i(\theta) \mathbf{I}_{ik} = \mathbf{0}. \quad (5)$$

Řešením soustavy (5) získáme odhad parametru θ .

Tvrzení 2. Předpokládejme, že (Y_i, W_i, ξ_i) jsou iid veličiny a ξ_i je podmíněně nezávislé na Y_i , je-li dáno W_i , pro $i = 1, 2, \dots, N$. Označme $\hat{\boldsymbol{\theta}}$ řešení soustavy rovnic (5). Potom

$$\sqrt{N}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} N(0, J(\boldsymbol{\theta})^{-1}\Sigma J(\boldsymbol{\theta})^{-1}), \quad (6)$$

kde

$$\Sigma = J(\boldsymbol{\theta}) + \sum_{k=1}^K P(W_i = k) \frac{1 - \pi_k}{\pi_k} J_k(\boldsymbol{\theta}),$$

$J(\boldsymbol{\theta})$ je Fisherova informace $\text{var} \mathbf{U}_i(\boldsymbol{\theta})$ a $J_k(\boldsymbol{\theta}) = \text{var}(\mathbf{U}_i(\boldsymbol{\theta})|W_i = k)$.

Náznak důkazu Nejprve ukážeme, že

$$\frac{1}{\sqrt{N}} \mathbf{V}(\boldsymbol{\theta}) \xrightarrow{d} N(0, \Sigma). \quad (7)$$

Taylorovým rozvojem $\frac{1}{\hat{\pi}_k}$ kolem $\frac{1}{\pi_k}$ dostaneme

$$\sqrt{N_k} \left(\frac{1}{\hat{\pi}_k} - \frac{1}{\pi_k} \right) = \frac{-\sqrt{N_k}}{\pi_k^2} (\hat{\pi}_k - \pi_k) + o_p(1) = -\frac{1}{\pi_k^2 \sqrt{N_k}} \sum_{i=1}^N (\xi_i - \pi_k) \mathbf{I}_{ik} + o_p(1)$$

a můžeme psát

$$\begin{aligned} \frac{1}{\sqrt{N}} \mathbf{V}(\boldsymbol{\theta}) &= \frac{1}{\sqrt{N}} \left[\sum_{i=1}^N \sum_{k=1}^K \frac{\xi_i \mathbf{U}_i(\boldsymbol{\theta})}{\pi_k} \mathbf{I}_{ik} - \right. \\ &\quad \left. - \sum_{k=1}^K \frac{1}{\pi_k N_k} \left(\sum_{i=1}^N \frac{\xi_i \mathbf{U}_i(\boldsymbol{\theta})}{\pi_k} \mathbf{I}_{ik} \right) \left(\sum_{j=1}^N (\xi_j - \pi_k) \mathbf{I}_{jk} \right) \right] + o_p(1). \end{aligned}$$

Označíme-li $\mathbf{S}_k = E(\mathbf{U}_i(\boldsymbol{\theta})|W_i = k)$, dostaneme

$$\frac{1}{\sqrt{N}} \mathbf{V}(\boldsymbol{\theta}) = \frac{1}{\sqrt{N}} \left[\sum_{i=1}^N \sum_{k=1}^K \frac{\xi_i \mathbf{U}_i(\boldsymbol{\theta})}{\pi_k} \mathbf{I}_{ik} - \sum_{i=1}^N \sum_{k=1}^K \frac{(\xi_i - \pi_k)}{\pi_k} \mathbf{S}_k \mathbf{I}_{ik} \right] + o_p(1).$$

Tedy

$$\frac{1}{\sqrt{N}} \mathbf{V}(\boldsymbol{\theta}) = \frac{1}{\sqrt{N}} \sum_{i=1}^N \mathbf{Q}_i(\boldsymbol{\theta}) + o_p(1),$$

kde

$$\mathbf{Q}_i(\boldsymbol{\theta}) = \sum_{k=1}^K \left[\frac{\xi_i \mathbf{U}_i(\boldsymbol{\theta})}{\pi_k} - \frac{(\xi_i - \pi_k)}{\pi_k} \mathbf{S}_k \right] \mathbf{I}_{ik}$$

jsou iid veličiny a $E\mathbf{Q}_i = E\mathbf{U}_i(\boldsymbol{\theta}) = \mathbf{0}$. Podle centrální limitní věty pro iid náhodné veličiny

$$\frac{1}{\sqrt{N}}\mathbf{V}(\boldsymbol{\theta}) \xrightarrow{d} N(\mathbf{0}, \text{var } \mathbf{Q}_i).$$

Po krátkém výpočtu dostaneme

$$\text{var } \mathbf{Q}_i = E\mathbf{Q}_i\mathbf{Q}_i^T = J(\boldsymbol{\theta}) + \sum_k P(W_i = k) \frac{1 - \pi_k}{\pi_k} J_k(\boldsymbol{\theta}).$$

Taylorovým rozvojem $\mathbf{V}(\hat{\boldsymbol{\theta}})$ kolem $\mathbf{V}(\boldsymbol{\theta})$

$$\mathbf{V}(\hat{\boldsymbol{\theta}}) - \mathbf{V}(\boldsymbol{\theta}) = \left(\frac{\partial}{\partial \theta_j} \mathbf{V}(\boldsymbol{\theta})^T \right)_{j=1}^p (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) + o_p\left(\frac{1}{\sqrt{N}}\right),$$

a tudíž

$$\sqrt{N}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) = - \left[\frac{1}{N} \left(\frac{\partial}{\partial \theta_j} \mathbf{V}(\boldsymbol{\theta})^T \right)_{j=1}^p \right]^{-1} \frac{1}{\sqrt{N}} \mathbf{V}(\boldsymbol{\theta}) + o_p(1).$$

Máme

$$-\frac{1}{N} \left(\frac{\partial}{\partial \theta_j} \mathbf{V}(\boldsymbol{\theta})^T \right)_{j=1}^p = -\frac{1}{N} \sum_{i=1}^N \sum_{k=1}^K \frac{\xi_i}{\hat{\pi}_k} \left(\frac{\partial}{\partial \theta_j} \mathbf{U}_i(\boldsymbol{\theta})^T \right) \mathbf{I}_{ik} + o_p(1) \xrightarrow{P} J(\boldsymbol{\theta}). \quad (8)$$

Z (8) a (7) dostaneme

$$\sqrt{N}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} N(0, J(\boldsymbol{\theta})^{-1} \Sigma J(\boldsymbol{\theta})^{-1}).$$

4 Lineární model

Předpokládejme, že (Y_i, \mathbf{x}_i) , $i = 1, 2, \dots, N$, jsou iid náhodné vektory a že Y_i za podmínky \mathbf{x}_i má rozdělení se střední hodnotou $\mathbf{x}_i^T \boldsymbol{\beta}$ a rozptylem σ^2

$$Y_i | \mathbf{x}_i \sim (\mathbf{x}_i^T \boldsymbol{\beta}, \sigma^2), \quad i = 1, 2, \dots, N. \quad (9)$$

Soustava rovnic pro odhad parametru $\boldsymbol{\beta}$ (neuvažujeme-li výběrové schéma) je potom

$$\sum_{i=1}^N \mathbf{x}_i (Y_i - \mathbf{x}_i^T \boldsymbol{\beta}) = 0. \quad (10)$$

Předpokládejme nyní opět výběrové schéma popsané výše. Připomeňme, že příslušnost k jednomu z K strat (jež reprezentuje náhodná veličina W_i) je pozorovaná pro všech N členů populace, zatímco Y_i a \mathbf{x}_i pozorujeme pouze pro jedince z výběru, tj. pokud $\xi_i = 1$.

Uvažujme nejprve situaci, kdy v každém stratu platí model (9), tj.

$$(Y_i | \mathbf{x}_i, W_i) \sim (\mathbf{x}_i^T \boldsymbol{\beta}, \sigma^2), \quad i = 1, 2, \dots, N. \quad (11)$$

Tedy střední hodnota a rozptyl Y_i jsou nezávislé na W_i . V takovém případě nemusíme výběrové schéma zohledňovat a klasická teorie lineárních modelů platí i pro data sesbíraná stratifikovaným výběrovým schématem.

Obecně však v každém stratu platí jiný vztah než (11). Nicméně v popředí našeho zájmu zůstává stále marginální model (9), někdy též nazývaný populační (v anglické literatuře „population averaged“).

Příklad Uvažujme, že vektor \mathbf{x}_i zahrnuje všechny prediktory Y_i (kromě W_i) a že ve stratu k platí

$$(Y_i | \mathbf{x}_i, W_i = k) \sim (\mathbf{x}_i^T \boldsymbol{\beta}_k, \sigma_k^2) \quad i = 1, 2, \dots, N.$$

Potom

$$\begin{aligned} E(Y_i | \mathbf{x}_i) &= \sum_{k=1}^K P(W_i = k) \mathbf{x}_i^T \boldsymbol{\beta}_k = \mathbf{x}_i^T \boldsymbol{\beta} \\ \text{var}(Y_i | \mathbf{x}_i) &= E \sigma_{W_i}^2 + \text{var} \mathbf{x}_i^T \boldsymbol{\beta}_{W_i} = \\ &= \sum_{k=1}^K P(W_i = k) \sigma_k^2 + \mathbf{x}_i^T \sum_{k=1}^K P(W_i = k) (\boldsymbol{\beta}_k - \boldsymbol{\beta})^2 \mathbf{x}_i, \end{aligned}$$

kde $\sigma_{W_i} = \sigma_k$ a $\boldsymbol{\beta}_{W_i} = \boldsymbol{\beta}_k$ pro $W_i = k$.

Z toho vidíme, že platí-li v každém stratu lineární model (homoskedastický) a jsou-li $\boldsymbol{\beta}_k$ stejná ve všech stratech, jsou předpoklady lineární regrese splněny i v modelu pro celou populaci. V takové situaci ale nemusíme výběrové schéma zohledňovat. Pokud se však $\boldsymbol{\beta}_k$ alespoň ve dvou stratech liší, populační model nemůže splňovat předpoklad homoskedasticity, neboť rozptyl Y_i závisí na prediktorech \mathbf{x}_i . \square

Problém zohlednění výběrového schématu v lineární regresi má tedy smysl uvažovat pouze v situaci, kdy nepředpokládáme homoskedasticitu. Pro rozptyl odhadu $\boldsymbol{\beta}$ pak můžeme použít např. robustní tzv. „sandwichový“ odhad.

V klasickém přístupu by $\hat{\boldsymbol{\beta}}$ byl stále řešením soustavy rovnic (10), avšak bereme-li v úvahu výběrové schéma, je nutné tuto soustavu modifikovat podle (5) a dostáváme

$$\sum_{i=1}^N \sum_{k=1}^K \frac{\xi_k}{\hat{\pi}_i} \mathbf{x}_i (Y_i - \mathbf{x}_i^T \boldsymbol{\beta}) \mathbf{I}_{ik} = 0. \quad (12)$$

Nechť i_1, \dots, i_n jsou všechna i taková, že $\xi_{i_l} = 1$ pro $l = 1, \dots, n$. Označme $d_{i_l} = \sum_{k=1}^K \frac{1}{\hat{\pi}_k} \mathbf{I}_{i_l k}$, $D = \text{Diag}(d_{i_1}, \dots, d_{i_n})$, $\mathbf{X} = (\mathbf{x}_{i_l j})_{n \times p}$ a $\mathbf{Y} = (Y_{i_l})_{n \times 1}$, pro $l = 1, \dots, n$. Řešením soustavy rovnic (12) je

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T D \mathbf{X})^{-1} (\mathbf{X}^T D \mathbf{Y}).$$

Pro $\hat{\boldsymbol{\beta}}$ platí:

$$\sqrt{N}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \xrightarrow{d} N(0, A^{-1} \Sigma A^{-1}), \quad (13)$$

$$\text{kde } A = E \mathbf{x}_i \mathbf{x}_i^T \quad (14)$$

$$\Sigma = \text{var } \mathbf{U}_i + \sum_{k=1}^K P(W_i = k) \frac{1 - \pi_k}{\pi_k} \text{var}(\mathbf{U}_i | W_i = k) \quad (15)$$

pro $\mathbf{U}_i = \mathbf{x}_i(Y_i - \mathbf{x}_i^T \boldsymbol{\beta})$.

Konzistentní odhad rozptylu $A^{-1} \Sigma A^{-1}$ získáme nahrazením neznámých hodnot jejich konzistentními odhady. Empirický odhad $\text{var } \mathbf{U}_i$ resp. $\text{var}(\mathbf{U}_i | W_i = k)$ označme jako $\text{var}_E \mathbf{U}_i$ resp. $\text{var}_E(\mathbf{U}_i | W_i = k)$. Zde je opět nutné vzít v úvahu výběrové schéma, tedy

$$\text{var}_E \mathbf{U}_i = \frac{1}{N} \sum_{i=1}^N \sum_{k=1}^K \frac{\xi_i}{\hat{\pi}_i} \mathbf{U}_i \mathbf{U}_i^T \mathbf{I}_{ik} \quad (16)$$

$$\text{var}_E(\mathbf{U}_i | W_i = k) = \frac{\sum_{i=1}^N \xi_i (\mathbf{U}_i - \hat{\mathbf{S}}_k)(\mathbf{U}_i - \hat{\mathbf{S}}_k)^T \mathbf{I}_{ik}}{\sum_{i=1}^N \xi_i \mathbf{I}_{ik}}, \quad (17)$$

kde $\hat{\mathbf{S}}_k$ je průměr \mathbf{U}_i ve stratu k .

5 Ilustrace

Výsledky ilustrujeme na malé simulační studii. Zajímá nás výše měsíčního platu v závislosti na vzdělání a délce praxe dané osoby. Necht výše platu závisí ale také na pohlaví. Předpokládejme, že platí lineární model s parametrem $\boldsymbol{\beta} = (\beta_0, \beta_z, \beta_p, \beta_{pz}, \beta_s, \beta_m, \beta_v, \beta_{vz})$

$$plat = \beta_0 - \beta_z * z + \beta_p * p - \beta_{pz} * z * p + \beta_s * S + \beta_m * M + \beta_v * V - \beta_{vz} * z * V + e, \quad (18)$$

kde z je dichotomická proměnná značící pohlaví (žena), p je délka praxe v letech, a S , M a V jsou dichotomické proměnné značící, zda má daná osoba pořadě střední vzdělání bez maturity, střední vzdělání s maturitou nebo vysokoškolské vzdělání (pro osoby se základním vzděláním jsou všechny tyto proměnné rovny nule) a e je náhodná veličina $e \sim N(0, 5000^2)$. Mějme populaci, ve které jsou hodnoty parametrů následující

$$\begin{aligned} \beta_0 &= 15\,000 & \beta_z &= -5\,000 & \beta_p &= 370 & \beta_{pz} &= -70 \\ \beta_s &= 3\,000 & \beta_m &= 9\,000 & \beta_v &= 27\,000 & \beta_{vz} &= -9\,000. \end{aligned} \quad (19)$$

Necht je podíl žen mezi výdělečně činnými osobami 0,4. Potom je marginální model popisující závislost platů pouze na délce praxe a vzdělání

$$\begin{aligned} plat &= [0,6\beta_0 M + 0,4(\beta_0 - \beta_z)] + [0,6\beta_p + 0,4(\beta_p - \beta_{pz})]p + \\ &+ \beta_s * S + \beta_m * M + [0,6\beta_v + 0,4(\beta_v - \beta_{vz})]V + e. \end{aligned}$$

Simulace byla provedena ve statistickém balíku R, verze 2.6.1. Nejprve jsme vygenerovali populaci o velikosti 10 000 podle modelu (18) s hodnotami parametru (19) a podílem žen 0,4. Proměnné o vzdělání a délce praxe byly generovány nezávisle na sobě a bez ohledu na pohlaví tak, aby přibližně odpovídaly reálnému rozložení těchto veličin v populaci. Potom jsme z této populace

	Skut. hodnota	Klas. odhad*	Odhad RVS*	Odhad SE*	SE** odhadu
Konstanta	13 000	14 101	13 018	532	542
Praxe	342	358	343	26	26
Vzděl. bez mat.	3 000	2 980	2 977	552	571
Vzděl. s mat.	9 000	8 985	8 987	554	556
Vzděl. VŠ	23 400	25 333	23 374	782	798

*průměrný; **empirická, SE... směrodatná chyba

Tabulka 1: Průměrné výsledky simulace (1000 opakování).

provedli náhodný výběr s pravděpodobností zahrnutí žen 0,1 a mužů 0,3. Parametr β jsme odhadli jak klasickým způsobem, tak zde popsaným postupem, který respektuje výběrové schéma (RVS). Tento proces jsme zopakovali 1000krát (Tabulka 1). Zatímco klasický odhad parametru je v těch složkách, které jsou odlišné pro obě pohlaví, vychýlený, průměrný odhad RVS parametru je blízko skutečným hodnotám. Stejně tak průměrný odhad směrodatné chyby odhadu RVS je blízky empirické směrodatné chybě odhadů, což je v souladu s výsledkem (13). V tomto příspěvku jsme se věnovali situaci, kdy je potřeba klasické statistické metody zobecnit tak, aby zohledňovaly dané výběrové schéma. Metodu maximální věrohodnosti jsme modifikovali přidáním vah, definovaných jako převrácená hodnota empirických pravděpodobností výběru. Podrobněji jsme se zabývali lineárním modelem, kde jsme pro odhad rozptylu použili tzv. „sandwichový“ odhad.

Literatura

- [1] Project Accept Study Group (2007). *Project Accept: A Phase III Randomized Controlled Trial of Community Mobilization, Mobile Testing, Same-Day Results, and Post-Test Support for HIV in Sub-Saharan Africa and Thailand*. <http://www.cbvct.med.ucla.edu/protocol.pdf>
- [2] Särndal C.E., Swensson B., Wretman J. (1991). *Model Assisted Survey Sampling*. Springer-Verlag, New York.
- [3] Šedová M., Kulich M. (2007). *Statistical Methods for Analysis of Survey Data*. In WDS'07 Proceedings of Contributed Papers: Part I – Mathematics and Computer Sciences Prague, Matfyzpress, 181 – 186.

Poděkování: Tato práce byla částečně podpořena výzkumným záměrem MSM 0021620839. Poděkování také patří KPMS, ČSOB a OMI ÚI AV ČR, v.v.i., za umožnění účasti M. Šedové na konferenci Robust 2008.

Adresa: M. Šedová, M. Kulich, MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8 – Karlín

E-mail: sedova@karlin.mff.cuni.cz

TESTY ADITIVITY V ANALÝZE DVOJNÉHO TRŽDĚNÍ BEZ OPAKOVÁNÍ

Marie Šimečková, Petr Šimeček, Dieter Rasch

Klíčová slova: Testy aditivity, Tukeyho test, Johnsonův – Graybillův test.

Abstrakt: Předpokládáme smíšený model analýzy dvojného třídění s pouze jedním pozorováním pro každou kombinaci efektů. Zajímá nás test přítomnosti interakce. Článek shrnuje vlastnosti pěti takových testů aditivity: Tukeyho testu, Mandelova testu, Johnsonova – Graybillova testu, Tusselova testu a lokálně nejlepšího invariantního testu. Tyto testy byly odvozeny pro model s pevnými efekty. Je navržena modifikace Tukeyho testu, která zvyšuje sílu tohoto testu pro obecnou interakci. Dále uvedeme empirickou formuli pro výpočet minimálního rozsahu výběru pro Johnsonův – Graybillův test.

Abstract: The two-way ANOVA mixed model with one observation for each row – column combination is considered and tests of interaction in this model are studied. The paper summarizes properties of five tests of additivity: Tukey test, Mandel test, Johnson – Graybill test, Tussel test and locally best invariant (LBI) test. These tests were developed for models with fixed effects. A modification of Tukey test to increase the power of this test for general interaction is proposed. Finally, an empirical formula for determination of the size of an experiment for the Johnson – Graybill test is introduced.

1 Úvod

Analýza dvojného třídění je dobře známá třída lineárních modelů, která umožňuje odhad a testování dvou hlavních efektů a jedné interakce. Obvykle je počet pozorování pro každou kombinaci faktorů větší než jedna, což umožňuje odhad hlavních efektů a interakce zároveň. Je-li počet opakování pro každou kombinaci faktorů právě jedna, klasické metody odhadů nelze použít.

Například si představme, že chceme testovat několik druhů léků. Je možné, že každý pacient bude na daný lék reagovat odlišně, ale nelze jednomu pacientovi podat více léků. Test významnosti interakce pro takový případ budeme nazývat testem aditivity.

První test aditivity pro model bez opakování navrhl John Tukey v [11]. Tento test byl odvozen pro speciální typ interakce a v ostatních případech má malou sílu (viz [4] a [5]). Další navržené testy jsou například Mandelův [7], Johnsonův – Graybillův [6], Tusselův [12] a lokálně nejlepší invariantní test (locally best invariant test, LBI) [3], tyto testy jsou shrnuty v [1]. Všechny tyto testy byly navrženy pro model s pevnými efekty.

Nyní popíšme situaci detailněji. Předpokládáme model analýzy dvojného třídění. Nejprve se zaměříme na model s pevnými efekty. Pak sledovaná veličina v i -tém řádku a j -tém sloupci je modelována jako

$$y_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ij}, \quad i = 1, \dots, a, \quad j = 1, \dots, b, \quad (1)$$

kde μ , α_i , β_j a $(\alpha\beta)_{ij}$ jsou reálné konstanty splňující $\sum_i \alpha_i = \sum_j \beta_j = \sum_i (\alpha\beta)_{ij} = \sum_j (\alpha\beta)_{ij} = 0$ a e_{ij} jsou nezávislé normálně rozdělené náhodné veličiny s nulovou střední hodnotou a rozptylem σ^2 .

Ve smíšeném modelu předpokládáme, že řádkový efekt je pevný, sloupcový efekt je náhodný a interakce je náhodná. Závislá proměnná je modelována jako

$$y_{ij} = \mu + \alpha_i + b_j + (ab)_{ij} + e_{ij}, \quad i = 1, \dots, a, \quad j = 1, \dots, b, \quad (2)$$

kde μ , α_i jsou reálné konstanty, b_j , $(ab)_{ij}$ a e_{ij} jsou normálně rozdělené náhodné veličiny, všechny s nulovou střední hodnotou a rozptyly postupně σ_B^2 , σ_{AB}^2 a σ^2 .

Chceme testovat hypotézu, že v modelu není interakce. V modelu s pevnými efekty (1) to znamená, že testujeme hypotézu $H_0: (\alpha\beta)_{ij} = 0, i = 1, \dots, a, j = 1, \dots, b$ proti alternativě $H_A: (\alpha\beta)_{ij} \neq 0$ pro alespoň jeden pár (i, j) . Ve smíšeném modelu (2) testujeme hypotézu $H_0: \sigma_{AB}^2 = 0$ proti alternativě $H_A: \sigma_{AB}^2 > 0$.

1.1 Testy aditivity

V této části stručně popíšeme pět testů aditivity odvozených pro model (1) s pevnými efekty s jedním pozorováním v podtřídě. Nejdříve si zavedeme značení: $\bar{y}_{..} = \sum_i \sum_j y_{ij}/ab$ označuje průměr závislé veličiny, $\bar{y}_{i.} = \sum_j y_{ij}/b$ řádkový průměr, $\bar{y}_{.j} = \sum_i y_{ij}/a$ sloupcový průměr. Nechť matice R je matice reziduí vzhledem k hlavním efektům s prvky $r_{ij} = y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}$. Nerosoutící posloupnost vlastních čísel matice RR^T budeme značit $\kappa_1 \geq \kappa_2 \geq \dots$ a jejich standardizovanou verzi $\omega_i = \frac{\kappa_i}{\sum_k \kappa_k}, i = 1, 2, \dots$.

Tukeyho test: Poprvé publikován v [11]. Tukey navrhl nejprve odhadnout řádkové a sloupcové efekty pomocí řádkových a sloupcových průměrů a pak testovat přítomnost interakce typu $(\alpha\beta)_{ij} = k \cdot \alpha_i \cdot \beta_j$, kde k je konstanta ($k = 0$ znamená nepřítomnost interakce). Testová statistika je rovna

$$S_T = MS_{int}/MS_{error},$$

kde

$$MS_{int} = \frac{\left(\sum_i \sum_j y_{ij} (\bar{y}_{i.} - \bar{y}_{..}) (\bar{y}_{.j} - \bar{y}_{..}) \right)^2}{\sum_i (\bar{y}_{i.} - \bar{y}_{..})^2 \sum_j (\bar{y}_{.j} - \bar{y}_{..})^2}$$

a

$$MS_{error} = \frac{\sum_i \sum_j (y_{ij} - \bar{y}_{..})^2 - a \sum_j (\bar{y}_{.j} - \bar{y}_{..})^2 - b \sum_i (\bar{y}_{i.} - \bar{y}_{..})^2 - MS_{int}}{(a-1)(b-1) - 1}.$$

Platí-li hypotéza aditivity, S_T má F -rozdělení s 1 a $(a-1)(b-1)-1$ stupni volnosti.

Mandelův test: Poprvé publikován v [7]. Mandel zobecnil Tukeyho test a testuje interakci typu $(\alpha\beta)_{ij} = \gamma_i \cdot \beta_j$, kde γ_i jsou konstanty. Testová statistika Mandelova testu je rovna

$$S_M = \frac{\sum_i (z_i - 1)^2 \sum_j (\bar{y}_{.j} - \bar{y}_{..})^2}{a-1} \bigg/ \frac{\sum_i \sum_j ((y_{ij} - \bar{y}_{i.}) - z_i(\bar{y}_{.j} - \bar{y}_{..}))^2}{(a-1)(b-2)},$$

kde

$$z_i = \frac{\sum_j y_{ij}(\bar{y}_{.j} - \bar{y}_{..})}{\sum_j (\bar{y}_{.j} - \bar{y}_{..})^2}.$$

Platí-li hypotéza aditivity, S_M má F -rozdělení s $a-1$ a $(a-1) \cdot (b-1)$ stupni volnosti.

Johnsonův – Graybillův test: Poprvé publikován v [6]. Autoři zvolili odlišný přístup a testují přítomnost interakce typu $(\alpha\beta)_{ij} = k \cdot \gamma_i \cdot \delta_j$, kde γ_i , δ_j a k jsou konstanty. Testová statistika Johnsonova – Graybillova testu je

$$S_J = \omega_1.$$

Hypotéza je zamítnuta, když S_J je vysoká.

Tusselův test: Viz [12]. Tusselův přístup je obdobný jako v Johnsonově – Graybillově testu. Bez újmy na obecnosti předpokládejme, že $a \leq b$. Navržená testová statistika je rovna

$$S_U = (a-1)^{(a-1)(b-1)/2} \left(\prod_{i=1}^{a-1} \omega_i \right)^{(b-1)/2}.$$

Hypotéza aditivity je zamítnuta, když S_U je nízká.

Lokálně nejlepší invariantní (LBI) test: Viz [3]. Tento test byl navržen tak, aby měl lokálně vyšší sílu než Tusselův test. LBI testová statistika je rovna

$$S_L = \frac{1}{a-1} \cdot \frac{1}{\sum_1^{a-1} \omega_i^2}.$$

Hypotéza aditivity je zamítnuta při nízké hodnotě S_L .

V naší práci byly kritické hodnoty pro poslední tři testy určovány pomocí simulace pro daná a a b .

Všechny zmiňované testy byly implementovány v prostředí R [8] v balíčku `AdditivityTests`. Tento balíček lze stáhnout z

http://5r.matfyz.cz/skola/AdditivityTests/additivityTests_0.3.zip.

2 Vlastnosti testů aditivity

V [9] bylo v simulační studii ukázáno, že testy aditivity popsané v předchozí sekci drží hladinu významnosti i při použití pro model se smíšenými efekty (2). Tamtéž je ukázáno, že síla těchto testů ve smíšeném modelu je obdobná jako v modelu s pevnými efekty. To znamená, že Tukeyho test a Mandelův test mají dobrou sílu pro interakci, která je součinem řádkového a sloupcového efektu, ale selhávají při detekci obecnější interakce. Síla Johnsonova – Graybillova testu, Tusselova testu a LBI testu je pro tento speciální případ o něco nižší, ale zato je dostačující i pro obecnější typ interakce.

Nyní představíme modifikaci Tukeyho testu, která zvyšuje sílu tohoto testu pro obecnou interakci. Pak uvedeme vzorec pro odhad počtu pozorování potřebných pro dosažení dané síly pro Johnsonův – Graybillův test.

2.1 Modifikovaný Tukeyho test

V klasickém Tukeyho testu popsáném výše testujeme model

$$\mathbf{y}_{ij} = \mu + \alpha_i + \beta_j + k \cdot \alpha_i \cdot \beta_j + \mathbf{e}_{ij} \quad (3)$$

proti podmodelu

$$\mathbf{y}_{ij} = \mu + \alpha_i + \beta_j + \mathbf{e}'_{ij}.$$

Odhady řádkových efektů $\hat{\alpha}_i$ a sloupcových efektů $\hat{\beta}_j$ jsou počítány stejným způsobem v obou modelech, ačkoli v plném modelu není závislost \mathbf{y}_{ij} na těchto parametrech lineární.

Hlavní myšlenka naší modifikace je, že plný model (3) je fitován nelineární regresí a k testu podmodelu

$$\mathbf{y}_{ij} = \mu + \phi_i + \psi_j + \mathbf{e}'_{ij}$$

je použit test poměrem věrohodnosti. Odhady řádkových a sloupcových efektů se v obou modelech liší.

Neadjustovaný test

Platí-li nulová hypotéza, odhadujeme parametry metodou maximální věrohodnosti jako $\hat{\mu} = \bar{y}_{..}$, $\hat{\alpha}_i^{(0)} = \bar{y}_{i.} - \bar{y}_{..}$ a $\hat{\beta}_j^{(0)} = \bar{y}_{.j} - \bar{y}_{..}$. Reziduální součet čtverců je roven

$$RSS^{(0)} = \sum_i \sum_j \left(y_{ij} - \hat{\mu} - \hat{\alpha}_i^{(0)} - \hat{\beta}_j^{(0)} \right)^2 = \sum_i \sum_j (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2.$$

V plném modelu (3) je v prvním kroku parametr k odhadnut jako

$$\hat{k}^{(0)} = \frac{\sum_i \sum_j \left(y_{ij} - \hat{\alpha}_i^{(0)} - \hat{\beta}_j^{(0)} - \hat{\mu}^{(0)} \right) \cdot \hat{\alpha}_i^{(0)} \cdot \hat{\beta}_j^{(0)}}{\sum_i \sum_j \left(\hat{\alpha}_i^{(0)} \right)^2 \cdot \left(\hat{\beta}_j^{(0)} \right)^2},$$

tedy stejně jako v Tukeyho testu. Pak pokračujeme iterativně a odhady vylepšujeme:

$$\begin{aligned}\hat{\alpha}_i^{(n)} &= \frac{\sum_j (y_{ij} - \hat{\mu} - \hat{\beta}_j^{(n-1)}) \cdot (1 + \hat{k}^{(n-1)} \cdot \hat{\beta}_j^{(n-1)})}{\sum_j (1 + \hat{k}^{(n-1)} \cdot \hat{\beta}_j^{(n-1)})^2}, \\ \hat{\beta}_j^{(n)} &= \frac{\sum_i (y_{ij} - \hat{\mu} - \hat{\alpha}_i^{(n-1)}) \cdot (1 + \hat{k}^{(n-1)} \cdot \hat{\alpha}_i^{(n-1)})}{\sum_i (1 + \hat{k}^{(n-1)} \cdot \hat{\alpha}_i^{(n-1)})^2}, \\ \hat{k}^{(n)} &= \frac{\sum_i \sum_j (y_{ij} - \hat{\alpha}_i^{(n-1)} - \hat{\beta}_j^{(n-1)} - \hat{\mu}) \cdot \hat{\alpha}_i^{(n-1)} \cdot \hat{\beta}_j^{(n-1)}}{\sum_i \sum_j (\hat{\alpha}_i^{(n-1)})^2 \cdot (\hat{\beta}_j^{(n-1)})^2}.\end{aligned}$$

Definujeme

$$RSS^{(n)} = \sum_i \sum_j (y_{ij} - \hat{\mu} - \hat{\alpha}_i^{(n)} - \hat{\beta}_j^{(n)} - k^{(n)} \hat{\alpha}_i^{(n)} \hat{\beta}_j^{(n)})^2.$$

Iteraci ukončíme, je-li rozdíl $|RSS^{(n)} - RSS^{(n-1)}|$ malý; označme pak $RSS = RSS^{(n)}$. Iterace ve většině případů konverguje velmi rychle, často se dostatečně blízko maxima dostaneme již po jednom kroku.

Při známém rozptylu σ^2 by testová statistika testu podílem věrohodností byla rovna

$$(RSS^{(0)} - RSS)/\sigma^2.$$

Tato statistika má asymptoticky χ^2 -rozdělení s 1 stupněm volnosti (viz například [13]).

Konzistentní odhad rozptylu reziduí σ^2 je roven $s^2 = \frac{RSS}{ab-a-b}$ a $\frac{RSS}{\sigma^2}$ má aproximativně χ^2 -rozdělení s $ab - a - b$ stupni volnosti. Tedy použijeme-li lineární aproximaci nelineárního modelu (3), statistika modifikovaného Tukeyho testu je rovna

$$S_{MT} = \frac{RSS^{(0)} - RSS}{\frac{RSS}{ab-a-b}}. \quad (4)$$

Statistika S_{MT} má aproximativně F -rozdělení s 1 a $ab - a - b$ stupni volnosti.

V [10] bylo ukázáno, že pro interakci typu $k \cdot \alpha_i \cdot \beta_j$ je síla modifikovaného testu téměř stejná jako síla Tukeyho testu. Pro obecnější interakci $\alpha_i \cdot \delta_j$ (δ_j konstanta) je síla modifikovaného testu výrazně vyšší.

Adjustace pro malé výběry

V [10] bylo ukázáno, že pro malé počty pozorování (nízký počet řádků a nebo sloupců b) je chyba prvního druhu místo 5% přibližně 6%. Je to proto, že statistika testu poměrem věrohodností konverguje k χ^2 -rozdělení pomalu (viz [2]) a pro nízké rozsahy výběru je třeba korekce. Zde ukážeme dva možné

přístupy pro případ, že počet řádků či sloupců je nižší než 20 (empirický odhad založený na našich simulacích).

První možností, jak vyřešit problém s malými výběry, je metoda bootstrap. Zabývejme se testovou statistikou $S = RSS^{(0)} - RSS$. Pak generujeme $N^{(boot)}$ -krát data podle modelu

$$y_{ij}^{(boot)} = \hat{\mu} + \hat{\alpha}_i^{(0)} + \hat{\beta}_j^{(0)} + r_{\pi_{ij}},$$

kde π je náhodná permutace indexů matice reziduí R . Pro každá generovaná data spočteme testovou statistiku $S^{(boot)} = RSS^{(0)(boot)} - RSS^{(boot)}$. Kritická hodnota modifikovaného Tukeyho testu je potom $(1 - \alpha) \cdot 100\%$ -kvantil z množiny statistik $S^{(boot)}$.

Druhá možnost je odhadnout rozptyl σ^2 náhodného šumu e_{ij} jako $s^2 = \frac{RSS}{ab-a-b}$ a pak generovat $N^{(sample)}$ -krát množinu dat podle modelu

$$y_{ij}^{(sample)} = \hat{\mu} + \hat{\alpha}_i^{(0)} + \hat{\beta}_j^{(0)} + e_{ij}^{(NEW)},$$

kde $e_{ij}^{(NEW)}$ jsou nezávislé normálně rozdělené náhodné veličiny s nulovou střední hodnotou a rozptylem s^2 . Protože za platnosti hypotézy aditivity je parametr k roven nule, vezmeme za testovou statistiku odhad tohoto parametru $\hat{k}^{(n)}$. Stejně jako pro bootstrap pak pro každý z $N^{(sample)}$ generovaných výběrů spočteme hodnotu této statistiky a hypotézu aditivity zamítneme, když více než $(1 - \alpha) \cdot 100\%$ z těchto statistik bude menších než statistika $\hat{k}^{(n)}$ spočtená pro pozorovaná data.

2.2 Rozsah výběru pro Johnsonův – Graybillův test

V této sekci budeme studovat další problém a navrhujeme empirickou formuli pro určení rozsahu experimentu pro Johnsonův – Graybillův test tak, aby bylo dosaženo dané síly testu. V modelu (2) předpokládáme interakci tvaru

$$(\mathbf{ab})_{ij} = k \cdot \alpha_i \cdot \mathbf{c}_j, \quad (5)$$

kde α_i jsou řádkové efekty v modelu (2), \mathbf{c}_j jsou nezávislé normálně rozdělené náhodné veličiny s nulovou střední hodnotou a rozptylem σ_B^2 , nezávislé na náhodných veličinách \mathbf{b}_j a náhodných chybách e_{ij} , k je reálná konstanta.

Interakce (5) je náhodná veličina s nulovou střední hodnotou a rozptylem $var(\mathbf{ab})_{ij} = k^2 \cdot \alpha_i^2 \cdot \sigma_B^2$.

Pomocí simulace jsme odhadli sílu Johnsonova – Graybillova testu pro různé hodnoty parametrů $a, b, \alpha_1, \dots, \alpha_a, k$ a σ_B v modelu (2). Předpokládali jsme počet úrovní pevného faktoru a roven 10, 20, 30, 40 nebo 50 a pět různých rozložení $\alpha_1, \dots, \alpha_a$: ekvidistantní, náhodný výběr z normálního rozdělení, náhodný výběr z t_3 -rozdělení, polovina α_i kumulovaná v jednom bodě a druhá polovina v jiném bodě a nakonec dva body α_i ve stejné vzdálenosti od 0 a ostatních $a - 2$ bodů rovno 0. Ve všech případech byly $\alpha_1, \dots, \alpha_a$ transformovány tak, aby jejich průměr byl roven 0. Bylo pozorováno, že

síla testu závisí na $\alpha_1, \dots, \alpha_a$ pouze prostřednictvím jejich součtu čtverců $\sum_{i=1}^a \alpha_i^2$. Předpokládali jsme tři různé hodnoty $\sum_{i=1}^a \alpha_i^2$: 296, 665 a 1496.

Rozptyl náhodného efektu \mathbf{b}_j jsme předpokládali roven 1, $\sqrt{2}$ nebo 2. Parametr k (který kontroluje rozptyl náhodné interakce $(\mathbf{ab})_{ij}$) nabýval hodnot 0.03, 0.05, 0.07 a 0.1. Rozptyl náhodného šumu \mathbf{e}_{ij} byl vždy roven 1, v případě odlišné hodnoty model může být přeskálován.

Síla každého testu roste se vzdáleností alternativní hypotézy od nulové hypotézy. Na základě průzkumu výsledků simulací jsme navrhli následující empirickou formuli pro závislost síly π Johnsonova – Graibillova testu na parametrech předpokládaného modelu (chyba 1. druhu rovna 5 %):

$$\pi(b) = 1 - \frac{1}{a \cdot b \cdot k^4 \cdot \sigma_B^4 \sum_{i=1}^a \alpha_i^2}. \quad (6)$$

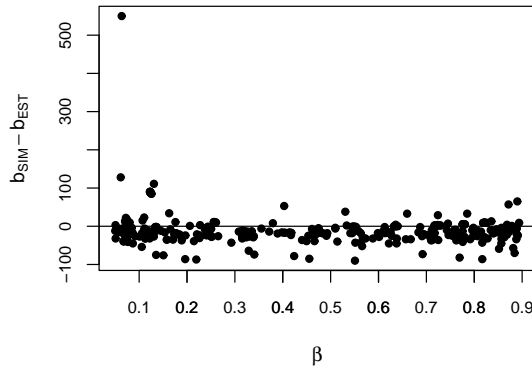
Formule je platná pouze pro sílu v intervalu $\langle 0.10, 0.95 \rangle$.

Připomeňme, že v praxi obvykle počet úrovní pevného faktoru a je pevně dán a velikost experimentu můžeme ovlivňovat pouze prostřednictvím počtu úrovní náhodného faktoru b . Jednoduchou úpravou formule (6) získáme vzoreček pro výpočet potřebného počtu úrovní náhodného faktoru

$$b(\beta) = \left\lceil \frac{1}{\beta \cdot a \cdot k^4 \cdot \sigma_B^4 \sum_{i=1}^a \alpha_i^2} \right\rceil, \quad (7)$$

kde $\lceil x \rceil$ znamená nejmenší celé číslo rovné x nebo vyšší než x . Pro $\sigma^2 \neq 1$ je třeba model přeskálovat.

Na obr. 1 je zakreslen rozdíl skutečného počtu úrovní náhodného faktoru b zjištěného simulací a odhadnutého pomocí formule (7). Formule (7) dává uspokojivé výsledky, ač je zde vidět několik odlehlých pozorování.



Obrázek 1: Závislost rozdílu skutečného počtu úrovní náhodného faktoru b (odhadnutého simulací) a počtu úrovní odhadnutého pomocí formule (7) na chybě 2. druhu β (vodorovná čára je nulová hladina).

Literatura

- [1] Alin A., Kurt S. (2006). *Testing non-additivity (interaction) in two-way ANOVA tables with no replication*. Statistical Methods in Medical Research, **15**, 63–85.
- [2] Bartlett M.S. (1937). *Properties of sufficiency and statistical tests*. Statistical Methods in Medical Research, **160**, 268–282.
- [3] Boik R.J. (1993). *Testing additivity in two-way classifications with no replications: the locally best invariant test*. Journal of Applied Statistics, **20**, 41–55.
- [4] Ghosh M.N., Sharma D. (1963). *Power of Tukey's test for non-additivity*. Journal of the Royal Statistical Society B, **25**, 213–219.
- [5] Hegeman V., Johnson D.E. (1976). *The power of two tests for nonadditivity*. Journal of the American Statistical Association, **71**, 945–948.
- [6] Johnson D.E., Graybill F.A. (1972). *An analysis of a two-way model with interaction and no replication*. Journal of the American Statistical Association, **67**, 862–868.
- [7] Mandel J. (1961). *Non-additivity in two-way analysis of variance*. Journal of the American Statistical Association, **56**, 878–888.
- [8] R Development Core Team (2008). *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org>.
- [9] Rusch T., Šimečková M., Kubinger K.D, Moder K., Šimeček P., Rasch D. (submitted) *Test of additivity in mixed and fixed effects two-way ANOVA models with single subclass numbers*. In Proceedings of The International Conference on trends and Perspectives in Linear Statistical Inference LINSTAT 2008, Bedlewo, Poland, April 21-25, 2008. Springer, special issue of Statistical Papers.
- [10] Šimeček P., Šimečková M. (submitted) *Modification of Tukey's additivity test*. Journal of Statistical Planning and Inference.
- [11] Tukey J.W. (1949). *One degree of freedom for non-additivity*. Biometrics, **5**, 232–242.
- [12] Tusell F. (1990). *Testing for interaction in two-way ANOVA tables with no replication*. Computational Statistics and Data Analysis, **10**, 29–45.
- [13] Zvára K. (2008). *Regrese*. Matfyzpress, Praha.

Poděkování: Práce byla podpořena výzkumným záměrem Ministerstva zemědělství České republiky MZE 0002701403. K simulacím byly využity výpočetní prostředky projektu METACentrum, podporovaného výzkumným záměrem MSM6383917201.

Adresa: M. Šimečková, P. Šimeček, D. Rasch, Výzkumný ústav živočišné výroby, Přátelství 815, 104 00 Praha

E-mail: simeckova.marie@vuzv.cz, simecek@gmail.com,
dieter.rasch@boku.ac.at

MATHEMATICAL MODELS FOR EPIDEMICS

Josef Štěpán

Keywords: Stochastic models for epidemics, Stochastic and partial differential equations.

Abstract: The paper reviews Kermack- Mc Kendrick deterministic classics (1927), its extensions and recent developments in the field of stochastic models for epidemics via the stochastic and partial differential equations theory. Theorems 1 and 2 in Section 6 suggest a possible interaction between these two ways of the modeling.

Dedicated to the memory of my tutor and friend professor Josef Machek.

Abstrakt: Článek je věnován Kermack–Mc Kendrickově klasickému deterministickému modelu vývoje epidemií. Po shrnutí výsledků známých z literatury je pozornost soustředěna na nejnovější výsledky z dané oblasti využívající stochastické a parciální diferenciální rovnice. Věty 1 a 2 v sekci 6 přitom ukazují na možné interakce mezi těmito dvěma přístupy k modelování používané v dané oblasti.

1 Kermack-Mc Kendrick model – a 1927 classics

The model is given formally by the following 2-dimensional differential equation

$$\dot{X}_t = -\beta \cdot X_t \cdot Y_t, \quad \dot{Y}_t = \beta \cdot X_t \cdot Y_t - \gamma \cdot Y_t, \quad \dot{Z}_t = \gamma \cdot Y_t \quad (1)$$

with initial conditions

$$X_0 = x_0 > 0, \quad Y_0 = y_0 = n_0 - x_0 > 0, \quad Z_0 = 0. \quad (2)$$

Obviously, $X_t + Y_t + Z_t = n_0$, and X_t and Z_t being a decreasing and increasing function, respectively, provided that $\beta > 0$ and $\gamma > 0$.

On the language of epidemics theory, $X_0 + Y_0 + Z_0 = n_0$ is a fixed size of population, X_t is the size of population of the individuals exposed to the infection (susceptibles), Y_t the infected individuals who may spread the disease (infectives) and finally Z_t the individuals restored to health who are not able not to spread the infection or get themselves infected again (removals).

A constant $\beta > 0$ is interpreted as the intensity of infection, while $\gamma^{-1} > 0$ is chosen to be proportional to the average duration of being infected status.

The model is said to be adequate in case of a highly infectious disease with a fast recovery in an homogeneous population, where the product $X_t \cdot Y_t$ counts possible contacts between the susceptibles and infected individuals contacts, $\beta \cdot X_t \cdot Y_t$ the number of “fatal” contacts.

The equation (1) has a unique solution (X_t, Y_t, Z_t) given the initial conditions (2). To recover its properties we reason as follows. Obviously,

$$\frac{\dot{X}}{X} = -\frac{\beta}{\gamma} \dot{Z} \quad \Rightarrow \quad X_t = x_0 \cdot \exp \left\{ -\frac{\beta}{\gamma} \cdot Z_t \right\} = x_0 \cdot \exp \left\{ -\int_0^t \beta \cdot Y_u \, du \right\}.$$

Similarly,

$$Y_t = y_0 \cdot \exp \left\{ \int_0^t \beta \cdot X_u \, du - \gamma t \right\}$$

Hence, $Z_t = \gamma \cdot \int_0^t Y_u \, du$ is an increasing function such that $0 \leq Z_t < n_0$ holds, while X_t decreases on $[0, \infty)$ and $Y_\infty = 0$.

Observing that

$$\dot{Y}_t = \beta \cdot X_t \cdot Y_t - \gamma \cdot Y_t,$$

and assuming $\frac{\gamma}{\beta} < x_0$, we conclude that $\dot{Y}_t > 0$ if and only if $X_t > \frac{\gamma}{\beta}$ holds and therefore Y_t has a unique maximum $Y_{t^+} = n - z^+ - \frac{\gamma}{\beta}$ where z^+ and t^+ are to be specified in Section 2. in a more general setting.

Also remark that Y_t is decreasing on $[0, \infty)$ if and only if $x_0 \leq \frac{\beta}{\gamma}$ and observing that

$$X_t = x_0 \cdot \exp \left\{ -\frac{\beta}{\gamma} \cdot Z_t \right\},$$

we get the equation for Z_∞

$$Z_\infty = n_0 - X_\infty = n_0 - x_0 \cdot \exp \left\{ -\frac{\beta}{\gamma} \cdot Z_\infty \right\}$$

that has a unique solution on $[0, \infty)$.

The transition $X \rightarrow Y$ is performed with the rate β , $Y \rightarrow Z$ with the rate γ and finally $X \rightarrow Z$ with the rate $\rho := \frac{\beta}{\gamma}$. The parameters β , γ and ρ are the important inner characteristics of the epidemics, while Z_∞ , $Y^+ = \max Y_t$, $t^+ = \arg \max Y_t$ list its pivotal outputs.

The equation (1) has no explicit solution available. The approximation $e^{-u} \sim 1 - u + \frac{1}{2}u^2$ delivers approximation

$$Z_t \sim \frac{\rho^2}{x_0} \left(\frac{x_0}{\rho} - 1 \right) + \frac{\alpha \rho^2}{x_0} \tanh \left(\frac{1}{2} \gamma \alpha t - \phi \right), \quad (3)$$

where

$$\alpha = \left[\frac{2x_0}{\rho^2} (n - x_0) + \left(\frac{x_0}{\rho} - 1 \right)^2 \right]^{\frac{1}{2}}, \quad \phi = \tanh^{-1} \left[\frac{1}{\alpha} \left(\frac{x_0}{\rho} - 1 \right) \right]$$

and

$$X_t \sim x_0 \cdot \exp \{ -\rho Z_t \}, \quad Y_t \sim n_0 - Z_t - X_t.$$

Having a population with $n_0 = 10000$ and $x_0 = 9995$ we get the following outputs:

ρ	Z_∞	Y^+	t^+
$8995 \sim x_0$	7995	1540	22.33
$4995 \sim \frac{1}{2}x_0$	948	57	98.29

The choice of $\rho \sim \frac{\beta}{\gamma} \sim x_0$ (the first line) provides a model for a pandemics and that with $\rho \sim \frac{1}{2}x_0$ (the second line) starts a mild epidemics.

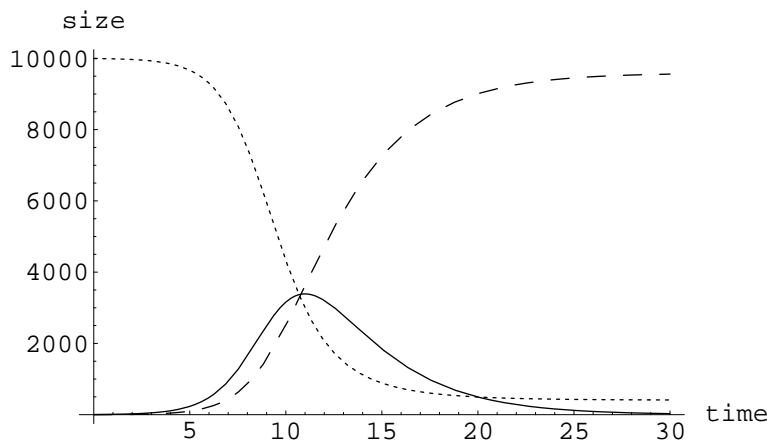


Figure 1: A mild epidemics.

We refer to Kermack, Mc Kendric [1], Daley, Gani [2], Štěpán, Hlubinka [4] and Hurt, Hlubinka [5] for a more detailed account on the content of the present section.

2 An extension

An obvious extension to the model (1) is given as

$$\dot{X}_t = -\beta(Z_t) \cdot X_t \cdot Y_t, \quad \dot{Y}_t = \beta(Z_t) \cdot X_t \cdot Y_t - \gamma \cdot Y_t, \quad \dot{Z}_t = \gamma \cdot Y_t \quad (4)$$

with the initial conditions (2). We assume $\beta(z)$ nonnegative, bounded and locally Lipschitz on \mathbb{R}^+ . Theorem 1 in [4] states that the equation has a unique solution $(X, Y, Z) \in C^1(\mathbb{R}^+, \mathbb{R}^3)$ that is positive on $(0, \infty)$ and Z_t is received as $t^{-1}(Z)$, where

$$t(Z) = \frac{1}{\gamma} \int_0^Z \left[n_0 - u - x_0 \cdot \exp \left\{ \frac{-1}{\gamma} \int_0^u \beta(v) dv \right\} \right]^{-1} du. \quad (5)$$

Hence,

$$X_t = x_0 \cdot \exp \left\{ \frac{-1}{\gamma} \int_0^{Z_t} \beta(u) du \right\}, \quad Y_t = n - Z_t - X_t.$$

Again, X_t and Z_t is a decreasing and increasing function, respectively, Y_t positive with $Y_\infty = 0$. Moreover, $z = Z_\infty$ is found to be a solution to

$$n_0 - z = x_0 \cdot \exp \left\{ \frac{-1}{\gamma} \int_0^z \beta(u) du \right\}$$

and it is a unique solution in the interval $[0, n_0]$ if $\beta(z)$ is chosen to be non increasing.

Assuming $\beta(0) > 0$, $\frac{\gamma}{\beta(0)} < x_0$ and $\beta(z)$ non increasing, then Y_t has a unique maximum $Y^+ = Y_{t^+}$, where

$$Y^+ = n_0 - Z^+ - \frac{\gamma}{\beta(Z^+)}, \quad t^+ = \frac{1}{\gamma} \int_0^{Z^+} \frac{1}{Y_u} du,$$

and Z^+ is a unique $0 < Z < Z_\infty$ such that $\beta(Z)X(Z) = \gamma$. Obviously, Y_t is increasing on $[0, t^+]$ and decreasing on $[t^+, \infty)$.

Let us borrow, as an illustration, [4, Example 2]: Consider $\beta(Z)$ non increasing and supported by $[0, Z_1]$ defined as

$$\beta(Z) = \beta_0 \left(1 - \frac{Z}{Z_1} \right) \quad \text{if } Z < Z_1 \quad \beta(Z) = 0 \quad \text{elsewhere,}$$

where $\beta_0 > 0$, $Z_1 > 0$ and $\rho_0 = \frac{\gamma}{\beta_0} < x_0$. Note that $\dot{Y}_{0^+} = \rho_0 < x_0$ is a necessary and sufficient condition for the outbreak of epidemics. Moreover, $\dot{\beta}(Z_t) = -\frac{\beta_0 \gamma}{Z_1} \dot{Y}_t$ and $t^+ < t_1$ where t_1 is the time of the first entry of Z_t to Z_1 .

The choice

$$n_0 = 10000, \quad x_0 = 9995, \quad \gamma = \frac{1}{3}, \quad \rho_0 = 4995, \quad Z_1 = 4000$$

delivers

$$Z_\infty = 3303, \quad Y^+ = 1133, \quad t^+ = 19.04$$

and the total number of infected Z_∞ that is much smaller than that in our first example with the identical n_0, x_0, γ and $\rho \sim \rho_0$.

The numerical calculations employ a Mathematica program proposed by [5].

A more sophisticated choice for $\beta(z)$ is suggested in [3], see also, [2]:

$$\beta(Z) = \begin{cases} \frac{2\beta_0}{\left(1 - \frac{Z}{\rho_0}\right) + \left(1 - \frac{Z}{\rho_0}\right)^{-1}}, & 0 \leq Z \leq \rho_0 \\ 0 & \text{elsewhere} \end{cases}$$

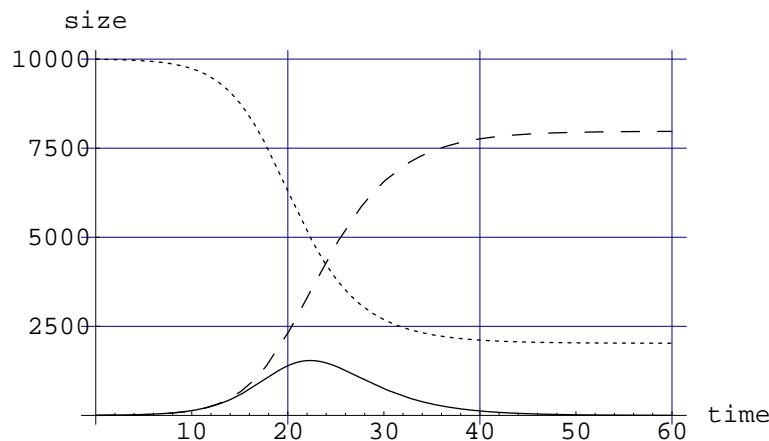


Figure 2:

where

$$\gamma, \beta_0 > 0, \quad \rho_0 = \frac{\gamma}{\beta_0}.$$

As a unique solution to (4) we recover (X_t, Y_t, Z_t) defined by (3), that we mentioned before to be an approximate solution to the classical Kermack-McKendrick equation (1).

3 Ad hoc stochastic models

History and present state of stochastic modeling of epidemics is treated in [6, 2, 7, 8].

Close to our setting are the models provided by [9] where $Y_t - \int_0^t (\beta X_u Y_u - \gamma Y_u) du$ is assumed to be martingale and a method how to estimate β suggested. See also, [10, 11, 12].

In [13] a diffusion model is proposed on the top of a deterministic infection in a population that consists only of susceptibles X_t and infectives Y_t , denoting by $N_t = X_t + Y_t$ the time dependent size of the population. The deterministic part is given by the differential equation

$$\dot{X}_t = -\frac{\beta}{N_t} X_t Y_t, \quad \dot{Y}_t = \frac{\beta}{N_t} X_t Y_t - \gamma Y_t \quad (\Rightarrow \quad \dot{N}_t = -\gamma Y_t).$$

and its diffusion extension as

$$\begin{aligned} dX_t &= -\beta \frac{X_t Y_t}{N_t} dt + b_{11}(t) dW_t^1 + b_{12}(t) dW_t^2 \\ dY_t &= \beta \frac{X_t Y_t}{N_t} dt - \gamma Y_t dt + b_{21}(t) dW_t^1 + b_{22}(t) dW_t^2. \end{aligned} \tag{6}$$

Here, (W^1, W^2) is a 2-dimensional Brownian motion and

$$\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} \frac{\beta}{N}XY & -\frac{\beta}{N}XY \\ -\frac{\beta}{N}XY & \frac{\beta}{N}XY + \gamma Y \end{pmatrix}^{\frac{1}{2}}.$$

There is a unique strong solution to (6) by Philis-Saranson theorem, see [14, 12.12 Theorem, p. 134]. The size of population N_t is received as a supermartingale here. The model is considered as a good device to study infections as HIV-AIDS.

Štěpán and Hlubinka [4] propose Kermack-Mc Kendrick model in the form

$$\begin{aligned} \dot{\mathbf{x}}_t &= -\alpha(\mathbf{x}_t, \mathbf{y}_t, \mathbf{z}_t, N_t)\mathbf{x}_t\mathbf{y}_t, & \dot{\mathbf{y}}_t &= \alpha(\mathbf{x}_t, \mathbf{y}_t, \mathbf{z}_t, N_t)\mathbf{x}_t\mathbf{y}_t - \gamma\mathbf{y}_t, & \dot{\mathbf{z}}_t &= \gamma\mathbf{y}_t \\ \mathbf{x}_0 &= \frac{x_0}{n_0}, & \mathbf{y}_0 &= \frac{y_0}{n_0}, & \mathbf{z}_0 &= 0, \end{aligned} \quad (7)$$

where N_t is a time dependent size of a population that starts at $N_0 = n_0 := x_0 + y_0$ and $\mathbf{x}_t + \mathbf{y}_t + \mathbf{z}_t = 1$ holds. The coordinates of the solution $(\mathbf{x}_t, \mathbf{y}_t, \mathbf{z}_t)$ are now the proportions of the susceptibles, infectives and removals subpopulations, respectively.

The coefficient $\alpha(x, y, z, n)$ is assumed to be a function on $\{(x, y, z) : x + y + z = 1\}$ for arbitrary $n \in [a, b]$, where $0 \leq a < n_0 < b < \infty$ are prescribed bounds for N_t that registers the time dependent size of population as a result of immigration - emigration processes. This might be a model for an epidemics or even pandemic with negligible fatal consequences. See also [16] where a special case is treated.

Thus, $X_t = \mathbf{x}_t \cdot N_t$, $Y_t = \mathbf{y}_t \cdot N_t$, $Z_t = \mathbf{z}_t \cdot N_t$ where $X_t + Y_t + Z_t = N_t$ are the sizes of the corresponding subpopulations and we search for a stochastic differential equation (SDE) to recover for $L = (X, Y, Z)$. First, however, we must specify N_t .

We suggest N_t to solve the Engelbert-Schmidt equation

$$dN_t = N_t\sigma(N_t) dW_t, N_0 = n_0, \quad \sigma(n) \geq 0 \quad \text{bounded and supported by } [a, b] \quad (8)$$

which keeps the martingale solution N_t inside of a prescribed interval $[a, b]$.

Rescaling $\alpha(x, y, z, n) = n \cdot \beta(nx, ny, nz)$ we get

$$\begin{aligned} dX_t &= -\beta(X_t, Y_t, Z_t)X_tY_t dt + X_t\sigma(N_t) dW_t, & X_0 &= x_0 \\ dY_t &= \beta(X_t, Y_t, Z_t)X_tY_t dt - \gamma Y_t dt - Y_t\sigma(N_t) dW_t, & Y_0 &= y_0 \\ dZ_t &= \gamma Y_t + Z_t\sigma(N_t) dW_t, & Z_0 &= 0 \end{aligned} \quad (9)$$

that combined with (8) is a 3-dimensional SDE.

Assume that $\beta(x, y, z) \geq 0$ is bounded and $\sigma(n)$ supported by (a, b) .

Then [4, Lemma 1 and Theorem 2] state that

$$\begin{aligned} N_t &= n_0 \exp \left\{ \int_0^t \sigma(N_u) dW_u - \frac{1}{2} \int_0^t \sigma^2(N_u) du \right\} \\ X_t &= \frac{x_0}{n_0} \exp \left\{ - \int_0^t \beta(X_u, Y_u, Z_u) Y_u du \right\} N_t \\ Y_t &= \frac{y_0}{n_0} \exp \left\{ \int_0^t \beta(X_u, Y_u, Z_u) X_u du - \gamma t \right\}, \quad Z_t = \int_0^t \frac{Y_u}{N_u} du \cdot N_t. \end{aligned}$$

Moreover, N_t is a bounded martingale, $N_t \in [a, b]$ and N_∞ exists while $0 < X_t \leq b$ is a supermartingale and $0 \leq Z_t \leq b$ a submartingale. Finally, outside a P -null set, X_∞, Z_∞ exist, $Y_\infty = 0$ and the implications

$$X_\infty > 0 \iff N_\infty > 0 \quad \text{and} \quad a > 0 \Rightarrow X_\infty > 0$$

hold.

Assuming that

$$\int_a^{a^+} \frac{1}{u^2 \sigma(u)^2} du = \int_b^{b^+} \frac{1}{u^2 \sigma^2(u)} du = \infty \tag{10}$$

then $P[N_\infty = a] = \frac{b-n_0}{b-a}$ and $P[N_\infty = b] = \frac{n_0-a}{b-a}$.

According to [4, Theorem 3] the equation (9) has a unique strong solution if $\beta(x, y, z) \geq 0$ is a Lipschitz continuous function on

$$\Delta_{ab} = \{(x, y, z) \in [0, b]^3 : a \leq x + y + z \leq b\}$$

and $\sigma(n)$ is globally Lipschitz.

It follows by [4, Theorem 4], that that there is a unique weak solution to (9) if $\beta(x, y, z)$ is a globally Lipschitz and Engelbert-Schmidt equation (8) has the property of the unique weak existence, that is known to be equivalent to the condition (10) according to [15, Theorem 23.1,p.451].

How to solve the equation (9)? We have “the deterministic part” of the equation

$$\begin{aligned} \dot{\mathbf{x}}_t &= -\alpha(\mathbf{x}_t, \mathbf{y}_t, \mathbf{z}_t, N_t) \mathbf{x}_t \mathbf{y}_t, \quad \dot{\mathbf{y}}_t = \alpha(\mathbf{x}_t, \mathbf{y}_t, \mathbf{z}_t, N_t) \mathbf{x}_t \mathbf{y}_t - \gamma \mathbf{y}_t, \quad \dot{\mathbf{z}}_t = \gamma \mathbf{y}_t \\ \mathbf{x}_0 &= \frac{x_0}{n_0}, \quad \mathbf{y}_0 = \frac{y_0}{n_0}, \quad \mathbf{z}_0 = 0, \end{aligned}$$

where N_t is a random time dependent size of the population driven by

$$dN_t = N_t \sigma(N_t) dW_t, \quad N_0 = n_0,$$

and where $\alpha(x, y, z, n) = n \cdot \beta(nx, ny, nz)$.

It can be proved that the steps

- (1) construct a trajectory $N_t(\omega)$ of a solution N_t to (8),

(2) substitute $N_t(\omega)$ to (7) to get an ordinary differential equation, solve it to get a solution $(\mathbf{x}_t(\omega), \mathbf{y}_t(\omega), \mathbf{z}_t(\omega))$,

(3) put $X_t(\omega) = \mathbf{x}_t(\omega)N_t(\omega)$, $Y_t(\omega) = \mathbf{y}_t(\omega)N_t(\omega)$, $Z_t(\omega) = \mathbf{z}_t(\omega)N_t(\omega)$

exhibit a solution to our original equation (9). This is a unique strong solution, for example, if $\alpha(x, y, z, n) \geq 0$ is Lipschitz on $\{(x, y, z) \in [0, 1]^3 : x + y + z = 1\}$ uniformly for $n \in [a, b]$ and $\sigma(n)$ is a Lipschitz continuous function supported by $[a, b]$.

[4, Example 4] treats $\alpha(x, y, z, n) = \frac{B}{A+z}$ that is invariant of the size n to get the equation (7) in the form

$$\begin{aligned} \dot{\mathbf{x}} &= -\frac{B}{A+\mathbf{z}}\mathbf{x}\mathbf{y}, & \dot{\mathbf{y}} &= \frac{B}{A+\mathbf{z}}\mathbf{x}\mathbf{y} - \gamma\mathbf{y}, & \dot{\mathbf{z}} &= \gamma\mathbf{y} \\ \mathbf{x}_0 &= p_0 = \frac{x_0}{n_0}, & \mathbf{y}_0 &= 1 - p_0, & \mathbf{z}_0 &= 0. \end{aligned}$$

Choosing $B = \gamma$ and $A = \frac{p_0}{2}$, the equation (7) is solved, according to (5), by the differential equation

$$\dot{\mathbf{z}} = \gamma \left(1 - \mathbf{z} - \frac{p_0^2}{p_0 + 2\mathbf{z}} \right), \quad \mathbf{z}_0 = 0$$

that has no explicit solution. A numerical solution for $\gamma = \frac{1}{3}$, $p_0 = 0.9995$ and $A = \frac{p_0}{2}$ is visualized by

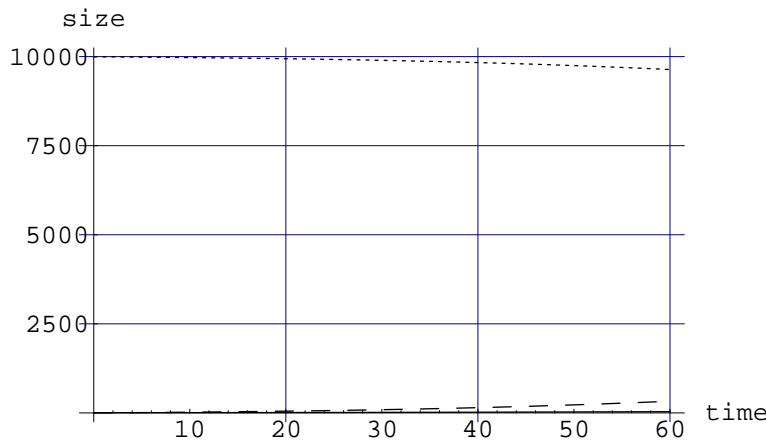


Figure 3:

The above choice of $\alpha(x, y, z, n)$ is rescaled to $\beta(x, y, z) = \frac{2\gamma}{p_0(x+y+z)+2z}$ to get SDE (9) in the form

$$\begin{aligned}
 dN_t &= N_t \sigma(N_t) dW_t, & N_0 &= n_0 \\
 dX_t &= -\beta(X_t, Y_t, Z_t) X_t Y_t dt + X_t \sigma(N_t) dW_t, & X_0 &= x_0 \\
 dY_t &= \beta(X_t, Y_t, Z_t) X_t Y_t dt - \gamma Y_t dt + Y_t \sigma(N_t) dW_t, & Y_0 &= y_0 \\
 dZ_t &= \gamma Y_t + Z_t \sigma(N_t) dW_t, & Z_0 &= 0.
 \end{aligned}$$

Choosing $\sigma(n) = 6\sqrt{\frac{(n-a)(b-n)}{a+b}}$, $a = 1000$, $b = 1050$, $n_0 = 1025$ we get the population size N_t as in Figure 4, X_t and Z_t as in Figure 5. The time dynamics of Y_t is recorded by Figure 6.

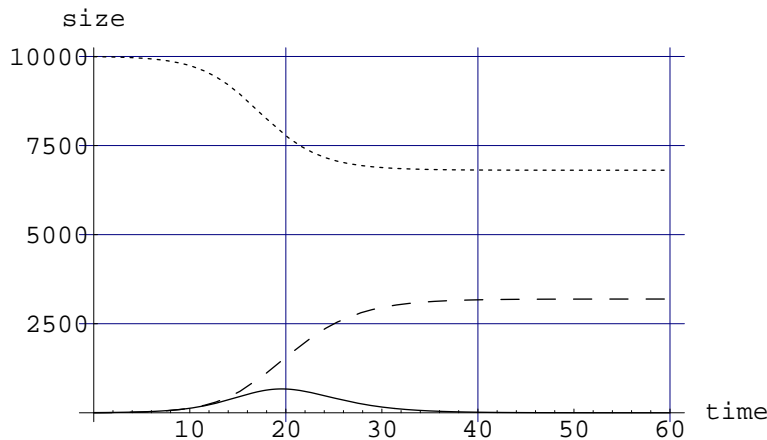


Figure 4:

The graphics and numerical calculations are borrowed from [5], again.

4 Stochastic models - a general theory via SDE

A general stochastic model of the Kermack-McKendrick epidemics is proposed in [17] by the following stochastic differential equation:

$$\begin{aligned}
 dX_t &= -\Phi(X_t, Y_t) dt + \Psi(X_t, Y_t) dW_t, & X_0 &= x_0 > 0 \\
 dY_t &= (\Phi(X_t, Y_t) - \gamma Y_t) dt - \Psi(X_t, Y_t) dW_t, & Y_0 &= y_0 > 0.
 \end{aligned} \tag{11}$$

Denote again $n_0 = x_0 + y_0$ and assume

$$\begin{aligned}
 &\Phi, \Psi : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad \text{Borel bounded on } [0, n_0]^2, \\
 &\Phi(x, y) = \Psi(x, y) = 0 \quad \text{outside } (0, \infty)^2, \quad \Psi(x, y) \geq 0 \quad \text{on } [0, n_0]^2.
 \end{aligned} \tag{12}$$

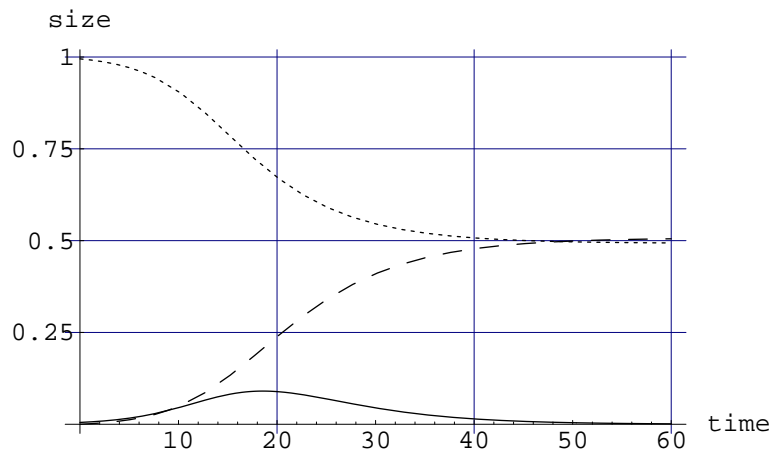


Figure 5:

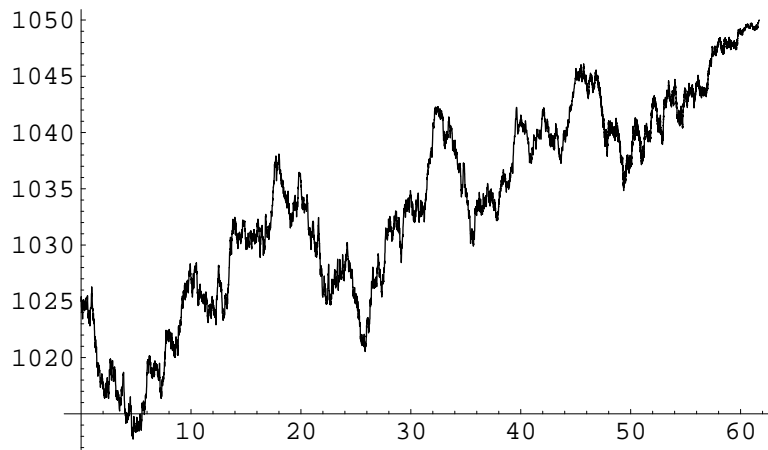


Figure 6:

For any solution (X, Y) to (11) X_t is a supermartingale and

$$(X, Y, Z) \in [0, n_0]^3, \quad X + Y + Z = n_0, \quad \text{where } Z_t = \gamma \int_0^t Y_s ds$$

$$Y_\infty = 0, \quad X_\infty \text{ exists}$$

according to [17, Lemma 2.1, Lemma 2.2 and Theorem 2.3].

An important feature of the model is a one to one correspondence $X \leftrightarrow Y$ $X_t = n_0 - Y_t - \gamma \int_0^t Y_s ds$ that may be inverted to

$$Y_t = -X_t + e^{-\gamma t}n_0 + \gamma e^{-\gamma t} \int_0^t e^{\gamma s} X_s ds$$

by means of Theorem 21.2 on Doleáns equation in [15, p. 414].

Define the stopping times

$$\tau_X = \inf\{t \geq 0 : X_t = 0\}, \quad \tau_Y = \inf\{t \geq 0 : Y_t = 0\}, \quad \tau = \tau_X \wedge \tau_Y$$

and observe that $X_t = 0$ on $[\tau_X, \infty)$ because X is a nonnegative supermartingale. Hence, a **solution** (X, Y) , is said to be **absorbed** if

$$t > \tau_Y \Rightarrow Y_t = 0 \quad \iff \quad (X_t, Y_t) = (X_\tau, Y_\tau \cdot e^{-\gamma(\tau_Y+t)}), \quad t > 0.$$

[4, Theorem 3.1 and Corollary 3.2] state that

if $\Phi(x, y), \Psi(x, y)$ are locally Lipschitz on $(0, n_0]^2$ then strong existence and both the strong and weak uniqueness hold for absorbed solutions.

and if Φ, Ψ are Lipschitz on $[0, n_0]^2$ then the strong and weak uniqueness hold, moreover, the unique solution is absorbed.

[17, Theorem 4.1 and 4.2] provide also sufficient conditions on the coefficients Φ and Ψ in order that (11) would produce only solutions (X, Y) that are absorbed.

The conditions read either as

$$\Phi(x, y) \leq \gamma \cdot \epsilon, \quad y \leq \epsilon, \quad x \in [0, n_0] \quad \text{for some } \epsilon > 0$$

or as

the weak uniqueness holds for the equation (11).

An obvious **criticism** of the model (11) may read as follows: A transfer $Y \rightarrow X$ is made possible by the stochastic differential $\Psi(X_t, Y_t) dW_t$. That its negative value only suggests a possibility to get infected again, or that of a false diagnosis is perhaps also an obvious answer.

5 Examples

The above theory was designed to cover the following examples all of them borrowed from [17]:

1. A naive extension of the Kermack-McKendrick model.

$$dX_t = -X_t Y_t dt + X_t Y_t dW_t, \quad dY_t = (X_t Y_t - \gamma Y_t) dt - X_t Y_t dW_t$$

The equation has a unique solution among $(X, Y) \geq 0$ and $\tau_X = \tau_Y = \infty$ almost surely.

2. A square root volatility model.

$$dX_t = -c\Phi(X_t, Y_t) dt + \sqrt{\Phi(X_t, Y_t)} dW_t$$

$$dY_t = (c\Phi(X_t, Y_t) - \gamma Y_t) dt - \sqrt{\Phi(X_t, Y_t)} dW_t$$

where $c > 0$ and $\Phi(x, y) > 0$ is a locally Lipschitz function on $[0, n_0]^2$ such that $\Phi(x, y) = 0$ holds on $[-\infty, 0]^2$.

There is a unique strong solution (X, Y) among $(X, Y) \geq 0$ and it is absorbed.

Moreover,

$$\tau_X < \infty \iff \tau_Y = \infty, \quad \tau = \tau_X \wedge \tau_Y < \infty.$$

The X_t -part of the solution is given as

$$X_t = x_0 - c\alpha(t) + B_{\alpha(t)},$$

where B is a Brownian motion and $\alpha(t)$ a random time change.

The specification

$$dX_t = -c \cdot X_t^+ Y_t^+ dt + \sqrt{d \cdot X_t^+ Y_t^+} dW_t$$

$$dY_t = c \cdot (X_t^+ Y_t^+ - \gamma t) dt + \sqrt{d \cdot X_t^+ Y_t^+} dW_t$$

with $X_0 = 970$, $Y_0 = 30$, $Z_0 = 0$, $c = 0.0005$, $d = 0.0015$, $\gamma = 0,25$ is visualized by

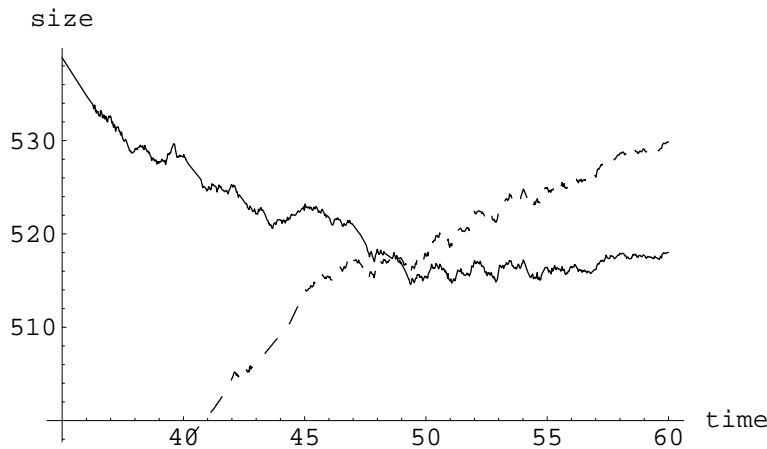


Figure 7:

3. A model that transfers the infection to at most cY_t individuals through the $[t, t + dt]$ time interval.

$$dX_t = -c \cdot Y_t \cdot I_{[X_t > 0]} dt + \sqrt{Y_t^+ \cdot I_{[X_t > 0]}} dW_t$$

$$dY_t = (c \cdot Y_t \cdot I_{[X_t > 0]} - \gamma Y_t) dt - \sqrt{Y_t^+ \cdot I_{[X_t > 0]}} dW_t.$$

If $0 < c \leq \gamma$ then there is unique solution among $(X, Y) \geq 0$, it is an absorbed solution and $P[\tau_Y < \infty] > 0$.

The unique strong solution is constructed as follows: Assume that $c \leq \gamma$, and the first thing, solve uniquely the equation

$$V_t = y_0 + (c - \gamma) \int_0^t V_s ds - \int_0^t \sqrt{V_s} dW_s$$

to get $V \geq 0$ with $\lambda_V = \inf\{t \geq 0 : V_t = 0\} < \infty$ such that $P[\lambda_V < \infty] = 1$ and such that V_t is killed at the time λ_V . A complete information on $\mathcal{L}(V)$ is available (see [19, Example 8.2, p. 221].

Putting

$$I_t = x_0 - c \int_0^t V_s ds + \int_0^t \sqrt{V_s^+} dW_s, \quad \lambda_I = \inf\{t \geq 0 : I_t = 0\}$$

and

$$\lambda = \lambda_V \wedge \lambda_I < \infty$$

we get

$$X_t = I_{t \wedge \lambda}, \quad Y_t = \begin{cases} V_t & \text{if } t < \lambda \\ V_\lambda e^{-\gamma(t-\lambda)} & \text{if } t \geq \lambda \end{cases}, \quad \lambda = \tau_X \wedge \tau_Y.$$

4. Choosing X_t as a martingale.

$$dX_t = Y_t \cdot I_{[X_t > 0]} dW_t, \quad dY_t = -\gamma Y_t dt - Y_t \cdot I_{[X_t > 0]} dW_t$$

The equation has a unique strong solution among $(X, Y) \geq 0$ with

$$0 < P[\tau_X < \infty] < 1$$

constructed as

$$X_t = x_0 + y_0 \cdot \int_0^t U_s dW_s, \quad Y_t = \begin{cases} y_0 \cdot U_t, & t < \lambda \\ y_0 \cdot U_\lambda \cdot e^{-\gamma(t-\lambda)}, & t \geq \lambda \end{cases}$$

where

$$U_t = \exp\{-\gamma + 1/2)t - W_t\}, \quad \lambda = \inf\left\{t \geq 0 : \int_0^t U_s dW_s = -\frac{x_0}{y_0}\right\}.$$

[17, Lemma 2.4], assuming (12), proclaims that outside a P -null set

$$\tau_Y < \infty \quad \Rightarrow \quad \tau_X < \tau_X$$

and

$$\tau_Y < \infty \quad \Rightarrow \quad \tau_X = \infty$$

hold for each absorbed solution (X, Y) .

Even the deterministic models can not be expected to provide a better behavior.

5. A deterministic counterpart to the Example 4.

The equation

$$dX_t = -\gamma Y_t^+ \cdot I_{[X_t > 0]} dt, \quad dY_t = -\gamma Y_t^+ \cdot I_{[X_t \leq 0]} dt, \quad x_0 = y_0 = \gamma = 1.$$

has a unique solution

$$X_t = (1 - t)^+, \quad Y_t = \exp\{-(t - 1)^+\}, \quad \tau_X = 1, \quad \tau_Y = \infty.$$

6 Partial differential equations and epidemics

Fix n_0 as a size of population and consider $x_0, \geq 0$ such that $x_0 + y_0 \leq n_0$. Applying the transformation $(X, Y) \rightarrow (Y, Z = n_0 - X - Y)$ to the equation

$$\begin{aligned} dX_t &= -\Phi(X_t, Y_t) dt + \Psi(X_t, Y_t) dW_t, & X_0 &= x_0 \geq 0 \\ dY_t &= (\Phi(X_t, Y_t) - \gamma Y_t) dt - \Psi(X_t, Y_t) dW_t, & Y_0 &= y_0 \geq 0 \end{aligned} \quad (13)$$

we get, more generally, equations

$$dY_t = b(Y_t, Z_t) dt - \sigma(Y_t, Z_t) dW_t, \quad Y_0 = y \quad dZ_t = \gamma Y_t dt, \quad Z_0 = z, \quad (14)$$

where

$$b(y, z) = \Phi(n_0 - y - z, y) - \gamma y, \quad \sigma(y, z) = \Psi(n_0 - y - z, y).$$

It easily seen that if (Y_t, Z_t) solves (14) with the initial conditions

$$Y_0 = y_0, \quad Z_0 = n_0 - x_0 - y_0 \quad \text{then} \quad (X_t, Y_t) = (n_0 - Y_t - Z_t, Y_t) \quad \text{solves (13).}$$

Having such a solution (Y_t, Z_t) and denoting by T a triangle in the (y, z) -plane with the vertices $(0, 0)$, $(n_0, 0)$ and $(0, n_0)$ (see Figure 8) we observe that:

- (1) $(Y, Z) \in T$
- (2) If $(Y_t, Z_t) = (0, z)$ at $t = \tau_Y$ then it stays at $(0, z)$ forever.

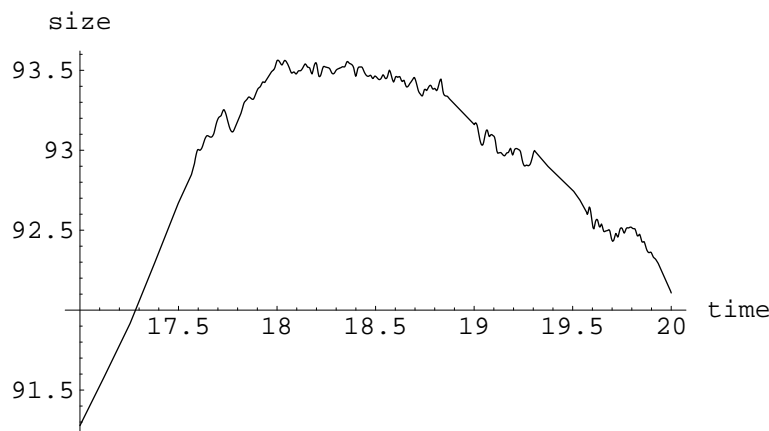


Figure 8:

(3) If $(Y_t, Z_t) \in D := [y + z = n_0]$ at $t = \tau_X$ then

$$(Y_{\tau_X+t}, Z_{\tau_X+t}) = (Y_{\tau_X} \cdot e^{-\gamma t}, n_0 - Y_{\tau_X} \cdot e^{-\gamma t}) \rightarrow (0, n_0).$$

(4) (Y_t, Z_t) never enters $z = 0$ for a $t > 0$.

The associated partial differential equation (PDE) reads as

$$\dot{u}(t, y, z) = b(y, z) \cdot u_y(t, y, z) + \gamma y \cdot u_z(t, y, z) + \frac{1}{2} \sigma^2(y, z) \cdot u_{yy}(t, y, z), \quad (15)$$

where

$$t \geq 0, \quad (y, z) \in \mathbb{R}^2, \quad u \in C^{1,2}(\mathbb{R}^+ \times \mathbb{R}^2).$$

Choosing a solution $u(t, y, z)$ to (15) that satisfies an initial condition $u(0, y, z) = f(y, z)$, where $f \in C(\mathbb{R}^2)$, we may establish, in some cases,

$$E_{y,z} f(Y_t, Z_t) = E[f(Y_t, Z_t) | Y_0 = y, Y_0 = z] \quad \text{as} \quad E_{y,z} f(Y_t, Z_t) = u(t, y, z),$$

where (Y, Z) is a solution to (14).

Come back to our original equation (13) and consider a function $g(x, y)$ in $C(\mathbb{R}^2)$. Putting $f(y, z) = g(n_0 - y - z, y)$ we get $f \in C(\mathbb{R}^2)$ and

$$Eg(X_t, Y_t) = u(t, y_0, n_0 - x_0 - y_0), \quad \text{holds for the solution } (X_t, Y_t) \text{ to (13)}$$

where $u(t, y, z)$ is a solution to (15) with $u(0, y, z) = f(y, z) = g(n_0 - y - z, y)$.

More precise information about the connection between the equations (14) and (15) is available by the next combination [18, Theorem 8.1.1, p. 139 and Lemma 7.5.4 (see also the remark preceding 7.5.2)]

Kolmogorov's backward equation theorem Let $f \in C^2(\mathbb{R}^2)$ be a function with a compact support and (Y_t, Z_t) be a solution to (13) where $b(y, z)$, $\sigma(y, z)$ are Lipschitz continuous functions on \mathbb{R}^2 .

- (1) If a function $w(t, y, z) \in C^{1,2}(\mathbb{R} \times \mathbb{R}^2)$ is a bounded solution to (15), then

$$w(t, y, z) := E_{y,z}f(Y_t, Z_t) \quad \text{on } \mathbb{R}^+ \times \mathbb{R}^2.$$

- (2) Put $u(t, y, z) := E_{y,z}f(Y_t, Z_t)$. Assume that $u(t, y, z)$ is a function in $C^{1,2}(\mathbb{R}^+ \times \mathbb{R}^2)$. Then $u(t, y, z)$ is a unique solution to (15).

To make the theorem to be useful in our situation we have to force $u(t, y, z)$ to be a function in $C^{1,2}$. Assuming that

$$\begin{aligned} \Phi(x, y) \text{ and } \Psi(x, y) & \text{ are bounded functions} \\ & \text{with a bounded continuous first and second derivation on } \mathbb{R}^2 \end{aligned} \quad (16)$$

we get that $b(y, z) = \Phi(n_0 - y - z, y) - \gamma y$, $\sigma(y, z) = \Psi(n_0 - y - z, y)$ are Lipschitz continuous on \mathbb{R}^2 and $u(t, y, z)$ as a function in $C^{1,2}(\mathbb{R}^+ \times \mathbb{R}^2)$ by [20, (9.4.4) Theorem].

The functions $u(t, y, z)$ are also easily forced to have a compact support $K_t \subset \mathbb{R}^2$ individually for arbitrary $t > 0$.

Consider $n_1 > 2n_0$ and compact triangles in the (x, y) -plane S_0 and S_1 with the vertices $(0, 0), (0, n_0), (n_0, 0)$ and $(0, 0), (0, n_1), (n_1, 0)$, respectively. Thus, we have $S_0 \subset [0, n_0]^2 \subset S_1$ and we shall assume that

$$\Phi(x, y) = \Psi(x, y) = g(x, y) = 0 \quad \text{outside } S_1. \quad (17)$$

Denote $f(y, z) = g(n_0 - y - z, y)$ and observe that

$$b(y, z) = -\gamma y, \quad \sigma(y, z) = f(y, z) = 0 \quad \text{for } (y, z) \notin T_1, \quad (18)$$

where $T_1 \supset T$ is an (y, z) -triangle with the vertices $(0, n_0 - n_1), (0, n_0), (n_1, n_0 - n_1)$.

It is not difficult to check that the equation (14) has a unique solution

$$Y_t = y \cdot e^{-\gamma t}, \quad Z_t = z + \gamma y \cdot \int_0^t e^{-\gamma s} ds = z + y(1 - e^{-\gamma t}), \quad t \geq 0$$

if either $y < 0$ or $y \geq 0$ and $y + z > n_0$. Also, for a fixed $t > 0$,

$$(Y_s, Z_s) = (y \cdot e^{-\gamma s}, z + y(1 - e^{-\gamma s})), \quad 0 \leq s \leq t$$

if $y \geq 0$ and $z + y(1 - e^{-\gamma t}) < n_0 - n_1$. Hence,

$$u(t, y, z) = E_{y,z}f(Y_t, Z_t) = 0 \quad \text{if } (y, z) \notin K_t,$$

where

$$K_t := \{(y, z) \in \mathbb{R}^2 : y \geq 0, y + z \leq n_0, z + y(1 - e^{-\gamma t}) \geq n_0 - n_1\}$$

that is a compact set because the lines $z + y = n_0$ and $z + y(1 - e^{-\gamma t}) = n_0 - n_1$ meet each other at the point (y^*, z^*) where $y^* = n_1 e^{\gamma t} > 0$ and $z^* = n_0 - n_1 e^{\gamma t} < 0$. Thus, by an application of Kolmogorov's backward equation theorem we prove.

Theorem 1. Let $b(y, z)$, $\sigma(y, z)$ and $f(y, z)$ are functions in $C^2(\mathbb{R}^2)$ that satisfy (18). Then

$$u(t, y, z) = E_{y,z} f(Y_t, Z_t), \quad \text{where } (Y_t, Z_t) \text{ solves (14)}$$

is a unique solution to the equation (15) with the initial condition $u(0, y, z) = f(y, z)$.

Moreover, if $\Phi(x, y)$, $\Psi(x, y)$ and $g(x, y)$ are functions in $C(\mathbb{R}^2)$ that satisfy (17) and (X_t, Y_t) solves our original equation (13), then

$$Eg(X_t, Y_t) = u(t, y_0, n_0 - x_0 - y_0), \quad t \geq 0,$$

where $u(t, y, z)$ is a solution to (15) such that $u(0, y, z) = g(n_0 - y - z, y)$ holds.

To solve PDE (15) we need obviously some tools. the Laplace transformation in order to remove the time variable is one of them:

If $u(t, y, z) = E_{y,z} f(Y_t, Z_t)$ is a bounded solution to (15) with the initial condition $u(0, y, z) = f(y, z)$, where $f(y, z)$ is a function in $C^2(\mathbb{R}^2)$ supported by a compact set, put

$$U(\lambda, y, z) = \int_0^\infty e^{-\lambda t} u(t, y, z) dt, \quad \lambda > 0.$$

Integrating per partes a we get the equation

$$\lambda U(\lambda, y, z) - b(y, z) \cdot U_y(\lambda, y, z) - \gamma y \cdot U_z(\lambda, y, z) - \frac{1}{2} \sigma^2(y, z) \cdot U_{yy}(\lambda, y, z) = f(y, z) \tag{19}$$

with the boundary conditions

$$U(\lambda, 0, z) = f(0, z) \cdot \frac{1}{\lambda}, \quad U(\lambda, y, n_0 - y) = \int_0^\infty e^{-\lambda t} f(y e^{-\lambda t}, n_0 - y e^{-\lambda t}) dt$$

to reflect the absorption rules (1), (2), (3) and (4) for the solution (Y_t, Z_t) to the equation (14). Rewriting (19) to

$$\begin{aligned} \lambda^2 U(\lambda, y, z) - b(y, z) \lambda U_y(\lambda, y, z) - \gamma y \lambda U_z(\lambda, y, z) - \frac{1}{2} \sigma^2(y, z) \lambda U_{yy}(\lambda, y, z) = \\ = \lambda f(y, z) \end{aligned}$$

and sending $\lambda \rightarrow 0^+$ we might get the equation

$$b(y, z) \cdot u_y(\infty, y, z) + \gamma y \cdot u_z(\infty, y, z) + \frac{1}{2} \sigma^2(y, z) \cdot u_{yy}(\infty, y, z) = 0,$$

since

$$\lim_{\lambda \rightarrow 0^+} \lambda \cdot U_y(\lambda, y, z) = u_y(\infty, y, z), \quad \lim_{\lambda \rightarrow 0^+} \lambda \cdot U_z(\lambda, y, z) = u_z(\infty, y, z).$$

Because $u(\infty, y, z) = E_{y,z}f(Y_\infty, Z_\infty) = E_{y,z}f(0, Z_\infty)$ we may venture to make a **conjecture** that $E_{y,z}f(0, Z_\infty) =: v(y, z)$ is a solution to the equation

$$b(y, z) \cdot v_y(y, z) + \gamma y \cdot v_z(y, z) + \frac{1}{2} \sigma^2(y, z) \cdot v_{yy}(y, z) = 0 \quad (20)$$

with the boundary conditions

$$v(0, z) = f(0, z) \quad \text{and} \quad v(y, n_0 - y) = f(0, n_0).$$

The contrary implication is obviously a true statement:

Theorem 2. Let $b(y, z)$, $\sigma(y, z)$ and $f(y, z)$ are functions in $C^2(\mathbb{R}^2)$ that satisfy (18). Let $v(y, z) \in C^2(\mathbb{R}^2)$ be a solution to (20) that satisfies the boundary condition $v(0, z) = f(0, z)$ and $u(t, y, z)$ a unique solution to (15) with the initial condition $u(t, y, z) = f(y, z)$. Then

$$v(y, z) = E_{y,z}f(0, Z_\infty) = u(\infty, y, z) \quad \text{holds for all} \quad (y, z) \in \mathbb{R}^2. \quad (21)$$

Moreover, if $\Phi(x, y)$, $\Psi(x, y)$ and $g(x, y)$ are functions in $C(\mathbb{R}^2)$ that satisfy (17) and (X_t, Y_t) solves our original equation (13), then

$$Eg(X_\infty, 0) = v(y_0, n_0 - x_0 - y_0),$$

where $v(y, z)$ is a solution to (20) such that $v(0, z) = g(n_0 - z, 0)$ holds.

Simply observe that by Theorem 1

$$u(t, y, z) = E_{y,z}f(Y_t, Z_t) \quad \forall (t, y, z) \in \mathbb{R}^+ \times \mathbb{R}^2$$

and by Itô formula and (20)

$$E_{y,z}v(Y_t, Z_t) = v(y, z) \quad \forall (t, y, z) \in \mathbb{R}^+ \times \mathbb{R}^2.$$

Sending $t \rightarrow \infty$ we prove (21).

7 Suggestions for a possible further research

There is much to be done in connection with the partial differential equations (15) and (20). Namely, Theorem 1 would deserve a further research with the aim to weaken the requirements on the coefficients $\Phi(x, y)$, $\Psi(x, y)$ and $f(x, y)$ to cover the models introduced by the examples in Section 5. Also, the heuristics presented in Section 6 crowned by a conjecture that $v(y, z) = E_{y,z}f(0, Z_\infty)$ is a solution to the equation (20) will need a nontrivial effort.

Our models are presented as one-dimensional ones. Their generalization to higher dimensions would allow to study stochastic epidemics with multiple pathogens.

Another open area of possible research are models with vaccination. The models presented up to now are general enough to include a vaccination as control variable. J. Staněk [21, 23] suggests a deterministic model given as

$$\begin{aligned} dX_t &= -\beta(Z_t) \cdot Y_t \cdot [X_t - v(Z_t)]^+ dt \\ dY_t &= \beta(Z_t) \cdot [X_t - v(Z_t)]^+ dt - \gamma \cdot Y_t \\ dZ_t &= \gamma \cdot Y_t, \\ X_0 &> 0, \quad Y_0 > 0, \quad Z_0 = 0, \end{aligned} \quad (22)$$

where $\beta(Z_t) \geq 0$ is a contact rate, $v(Z_t) \geq 0$ the size of vaccination assuming that

$\beta(z)$ and $v(z)$ are bounded and Lipschitz continuous functions
 $\beta(z)$ non increasing and $v(z)$ non decreasing function.

Then (22) has a unique solution such that

$$X_\infty, Y_\infty, Z_\infty \text{ exist and } Y_\infty = 0$$

holds. The paper provides a method how to compute Z_∞ and states conditions under which there is a unique $Y^+ = \max Y_t = Y_{t^+}$. A method how to compute Y^+ , t^+ is suggested. For an another approach we refer our reader to [22].

The first trial to incorporate a stochastic vaccination control to the general models introduced in Sections 5. and 6. of the present paper is made in [23]. An optimal Markov control of an epidemics by vaccination process should be the final goal of the research (the prize of vaccination versus the output quantity Z_t).

References

- [1] Kermack W.O., Kendrick A.G. (1927). *Contribution to mathematical theory of epidemics*. Proc. Roy. Soc. London Ser. A **115**, 700–721.
- [2] Daley D.J., Gani J. (1999). *Epidemic modelling: an introduction*. Cambridge University Press, Cambridge.
- [3] Kendall D.G. (1956). *Deterministic and stochastic epidemics in closed population*. Proc. Third Berkeley Symp. Math. Stat. Probab. **4**, 149-165.
- [4] Štěpán J., Hlubinka D. (2007). *Kermack-McKendrick epidemic model revisited*. Kybernetika **43**, 395–414.
- [5] Hurt J., Hlubinka D. (2006). *The Kermack-McKendrick model of epidemics and its extensions*. Wolfram Technology Conference 2006 (<http://library.wolfram.com/infocenter/Conferences/6462>).

- [6] Bailey N.T.J. *The mathematical theory of epidemics*. Hafner Publ. Company, New York.
- [7] Andersson H., Britton T. (2000). *Stochastic epidemics models and their statistical analysis (Lecture Notes in Statistics 151)*. Springer, New York.
- [8] Greenhalgh D. (2003). *Stochastic processes in epidemic modelling and simulation*. Handbook of Statistics **21**, 285–335.
- [9] Becker N.G., (1989). *Analysis of infectious data*. Chapman and Hall, London.
- [10] Allen E.J. (1999). *Stochastic differential equations and persistence time for two interacting populations*. Dynamics of Continuous, Discrete and Impulsive System **5**, 271–281.
- [11] Kiruparahan N., (2003). *Deterministic and stochastic epidemic models with multiple pathogens (PhD Thesis)*. Texas Tech. Univ., Lubbock.
- [12] Kiruparahan N., Allen L.S.J. (2004). *Consistence of multiple pathogen strains in stochastic epidemic models with density-dependent mortality*. Bull. Math. Biol. **66**, 841–864.
- [13] Allen L.S.J., Kiruparahan N. (2005). *Asymptotic dynamics of deterministic and stochastic epidemic models with multiple pathogens*. Internat. J. Num. Anal. Model. **2**, 329–344.
- [14] Rogers L.C.G., Williams D. (2000). *Diffusions, Markov processes and martingales*. Vol.2: Itô Calculus. Cambridge University Press, Cambridge.
- [15] Kallenberg O. (2002). *Foundations of modern probability (The 2nd Edition)*. Springer, New York.
- [16] Subramaniam R., Balachandran K., Kim J.K. (2000). *Existence of solution of a stochastic integral equation with an application from the theory of epidemics*. Nonlinear Funct. Anal. Appl. **5**, 23–29.
- [17] Štěpán J., Staněk J. (2008). *Absorption in stochastic epidemics*. To be published.
- [18] Øksendal B. (2003). *Stochastic differential equations. An introduction with applications*. Springer, New York.
- [19] Ikeda N., Watanabe S. (1989). *Stochastic differential equations and diffusion processes (The 2nd Edition)*. North Holland.
- [20] Arnold L. (1974). *Stochastic differential equations*. J. Wiley, New York.
- [21] Staněk J. (2008). *Kermack-McKendrick epidemics vaccinated*. Kybernetika **44(5)**, 705–714.
- [22] Wai-Yuan T., Hulin W. (1974). *Deterministic and stochastic models of AIDS epidemics and HIV infections with intervention*. World Scientific, Singapore.
- [23] Staněk J. (2008). *Stochastický model vývoje epidemie s vakcinací*. The Proceedings of Robust 2008, to be published.

Acknowledgement: The research was supported by the MSM 0021620839 grant.

Address: J. Štěpán, MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8

E-mail: josef.stepan@mff.cuni.cz

NELINEÁRNÍ STATISTICKÉ INFERENCE V NANOTECHNOLOGIÍCH

Pavel Tuček, Jiří Tuček, Jaroslav Marek,
Roman Procházka

Klíčová slova: Nanotechnologie, Mössbauerův jev, nelineární regresní modely s podmínkou.

Abstrakt: Tento příspěvek je zaměřen na statistické metody a vývoj filtrovacích procedur v Mössbauerově spektroskopii spojené s následným prokládáním očištěných dat příslušnou funkční závislostí. Statistické nástroje pro redukci šumu v měřeném spektru jsou užívány v mnoha oblastech. Šum v Mössbauerově spektroskopii je dán jako superpozice nerezonantních fotonů ve spojení s elektronickým šumem a rychlosti měřicího systému, která může být charakterizována jako vliv nelinearity zkoumaného procesu. Zcela nový matematicko-statistický přístup zlepšuje poměr signálu a šumu. Filtrovací procedura je založena na užití periodogramu a signifikantní hranice mezi šumem a užitečnými frekvencemi je zpětnovazebný kontrolní mechanismus založený na korelační analýze. Následně jsou ukázány možnosti přístupu finálního fitování. Korektnost tohoto řešení byla ověřena několika statistickými testy a bylo provedeno mnoho experimentálních měření s rostoucí kvalitou pro daný vzorek. Využitelnost této metody není univerzální, ale je vázána na některé specifické podmínky.

Abstract: This contribution focuses on the statistical methods and development of filtration procedures for Mössbauer spectra evaluation combined with subsequent fitting of the filtered data by the appropriate function dependence. Statistical tools for noise reduction in the measured spectrum are used in various fields. Noise in Mössbauer spectroscopy is given as a superposition of nonresonant photons in combination with electric noise and velocity of the measuring system, which can be characterized as an influence of nonlinearity of the investigated process. Absolutely new mathematical statistical approach improves the signal-to-noise ratio. The filtration procedure is based on the use of the periodogram and significant boundary between noise and useful frequencies is controlled by the feedback mechanism based on the correlation analysis. Moreover, we show the possibilities of the approach of the final fitting. The correctness of this solution has been verified by several statistical tests and we have carried out a series of experimental measurements with increasing quality for a given sample. The applicability of this method is not universal since it is connected to some specific conditions.

1 Motivace

Před několika lety bylo pod záštitou Univerzity Palackého v Olomouci založeno Centrum výzkumu nanomateriálů, jehož stěžejním cílem je syntetizovat a následně charakterizovat připravené systémy ve formě nanočástic. Z celé škály zde syntetizovaných nanočásticových systémů přitom vykazují největší aplikační potenciál magnetické nanočástice na bázi oxidu železitého. Aby bylo možné sofistikovaně rozhodnout, zdali syntetizovaný systém svými fyzikálně-chemickými vlastnostmi plně vyhovuje požadavkům kladených danou aplikací, je nutné jej charakterizovat pomocí řady experimentálních technik. Ve většině případů pak vhodnou matematickou analýzou získaných experimentálních dat obdržíme fyzikálně-chemické parametry, které jednoznačně popisují syntetizovaný nanočásticový systém a jsou tudíž klíčové pro jejich použití v praxi. Z hlediska praktického využití magnetických nanočástic na bázi oxidu železitého hrají nejdůležitější roli jejich magnetické vlastnosti. Ukazuje se, že k jejich magnetické charakterizaci lze s výhodou použít ^{57}Fe Mössbauerovu spektroskopii, která je založena na bezodrazové rezonanční emisi a absorpci γ -záření emitujícími a absorbujícími jádru atomů železa. Výsledkem obdržíme mössbauerovské spektrum, což je závislost počtu fotonů γ -záření, prošlých železo-obsahujícím vzorkem, na relativní rychlosti pohybu zářiče vůči absorberu. Je příhodné podotknout, že mössbauerovské spektrum je pro každý materiál jedinečné a lze z něj vyvodit parametry popisující magnetické chování zkoumaných magnetických nanomateriálů při různých teplotách. Jelikož jaderné procesy emise a absorpce vykazují stochastický charakter, nabízí se k vyhodnocení mössbauerovských spekter a odvození jejich příslušných parametrů použít teorii statistických modelů. Zde tedy navrhneme a prověříme statistický model, který bude sloužit k získání co možná nejpřesnějších odhadů parametrů příhodné aproximační funkce. Je zřejmé, že čím přesněji budeme schopni stanovit parametry aproximační funkce, tím komplexnější bude magnetická charakterizace různých nanočásticových systémů, jejichž rozlišení nám zpřesněné parametry umožní.

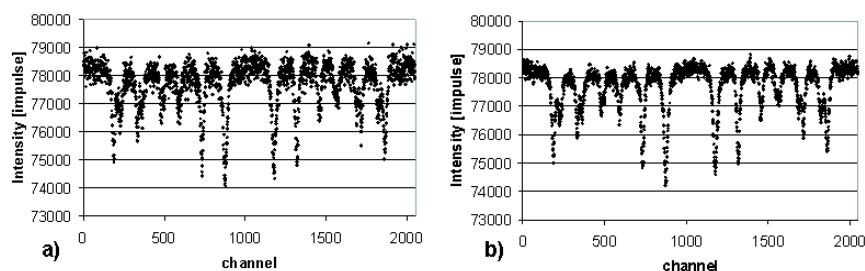
Primárním cílem tohoto modelu je tedy odhadnout parametry aproximační funkce pro popis pořízených mössbauerovských spekter, přičemž získaná experimentální data byla vyhlazena pomocí filtru navrhnutého v [4]. Tyto výsledky nacházejí dále využití pro porovnání s etalonem [1] a již provedenými studii.

2 Filtrace naměřených dat

Při analýze zkoumaného jevu byl registrován negativní vliv chyb měření na výsledné charakteristiky popisující vlastnosti měřeného materiálu. Z těchto důvodů byl navrhnut filtr, jehož účelem je odstranit rušivé vlivy při měření, které se negativně podílí na výsledku celé měřící procedury.

Analytické filtry pracují ve spektrální oblasti, kde jejich filtrovací vlastnosti vyjádřené pomocí matematických formulací mají jisté nevýhody pro

použití na poli filtrace mössbauerovských spekter. Nemožnost aplikace takových filtrů je limitována výkonností v určité spektrální oblasti. Na tyto nevýhody poukazují autoři v článku [4]. Druhý problém takových filtrů je změna fáze a deformuje výstupní informace, která se projevuje deformací spektrální pološírky, což je pro následnou charakterizaci zkoumaného materiálu nepřijatelné. Pro korektní užití vyvinuté filtrovací procedury vyžaduje nutnost možnosti redukovat nepodstatnou část informace ve vstupních datech. Navíc musí platit, že celá spektrální oblast musí zůstat bez ztráty informace v signálu. To jsou hlavní důvody, díky kterým byl vybrán statistický přístup. Za pomoci vývojového softwaru LabVIEW [2] byl vyvinut software, který obsahuje navržený filtr založený na periodogramu a zpětnovazební kontrole pomocí korelační analýzy. Výstupem námi navrženého softwaru je pak spektrum před filtrací, po filtraci a rovněž analýza odstraněného šumu.

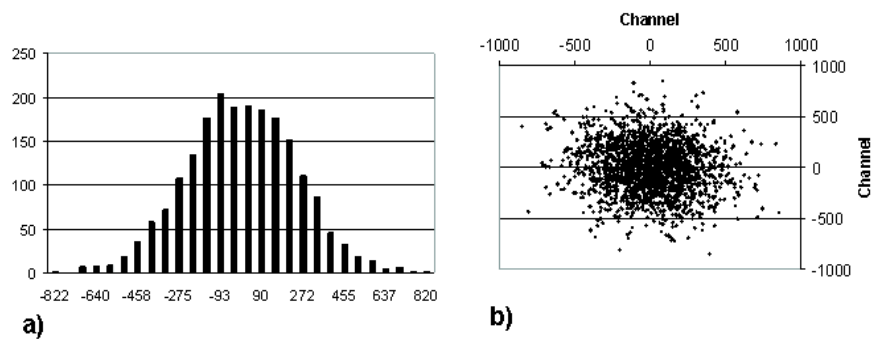


Obrázek 1: Sextet Mössbauerovského spektra před a po filtraci.

Na obr. 1 je zřetelně vidět ukázka spektra před filtrací a po filtraci. Vlevo je zřetelně vidět, jak spektrum obsahuje díky měřicím podmínkám a všem negativním okolnostem množství nepotřebné informace, která resultuje do nekvalitního obrazu spektra. Vpravo je vyobrazeno spektrum po aplikaci filtru, které již neobsahuje nepotřebnou informaci a rozostřenost dat je způsobena pouze faktem, že nebyly dosaženy teoretické podmínky měření, za kterých by byl signál zcela čitelný. Tyto podmínky jsou ale v experimentálním prostředí nerealizovatelné. Toto rozostření je tedy pro následnou analýzu potřebné. Samozřejmostí tohoto filtru taktéž je, že byly odstraněny odlehlá měření, která prokazatelně nesouvisí s měřenou veličinou.

Navržený statistický filtr je založen na použití periodogramu, který umožňuje odhad statisticky významných komponent signálu ve spektrální oblasti. Použití je pak doplněno o experimentální nastavení. Celý proces je navíc kontrolován zpětnovazebným procesem za použití analýzy korelace. Největší předností tohoto přístupu je fakt, že filtr redukuje množství nedůležité informace v datech a pracuje v celé spektrální oblasti.

Je zjevné, že chování statistického filtru nebude vždy uspokojivě účinné, protože účinnost závisí na statistické kvalitě experimentálních dat. Současné použití obdélníkového filtru jako zádržového filtru pak zajistí úplnou a uspokojivou filtraci.

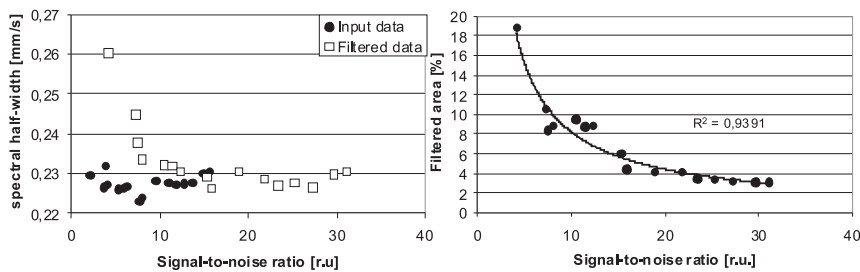


Obrázek 2: Histogram a korelogram šumu.

Obr. 2 představuje odfiltrovaný šum z výše uvedeného spektra, které bylo podrobeno filtrační proceduře za účelem odstranění nežádoucích měření. Vše, co bylo z naměřeného Mössbauerovského spektra odstraněno, je uloženo a podrobeno korelační analýze, kdy pro experimentátora nejpodstatnější částí je tzv. korelogram, který vizualizuje vlastnosti odstraněného šumu. Pokud je odstranění těchto částí spektra oprávněné, měl by se v korelogramu objevit „kruh“. V okamžiku, kdy korelogram utváří jinou strukturu, pak je ze spektra odstraňována užitečná informace. V průběhu filtrování probíhá zároveň testování významnosti korelačního koeficientu, které signalizuje experimentátorovi splnění dalšího požadavku, který je kladen na filtrovaná data a šum. Celá procedura je zastavena teprve tehdy, až dojde ke splnění všech požadavků, kladených na filtr.

Podmínky pro užitečnou filtraci. Prvotně je potřebné stanovit podmínky, za kterých bude zpracování signálu užitečné. Pro tento účel je užitečné zkoumat závislost mezi silou filtru a poměrem signálu a šumu (SNR). Závislost filtrační procedury a poměru signálu a šumu bylo testováno na experimentálních měřeních. Výsledky shrnují následující grafy.

Na obr. 3 vpravo je evidentně zřetelná klesající mocninná řada. Pro měření spektra s menším SNR je šum odfiltrován spolu s neidentifikovatelným signálem. Tohoto nemůže být dosaženo žádnou jinou metodou. Pro předchá-



Obrázek 3: Závislost spektrální pološířky a SNR. Závislost účinnosti a SNR.

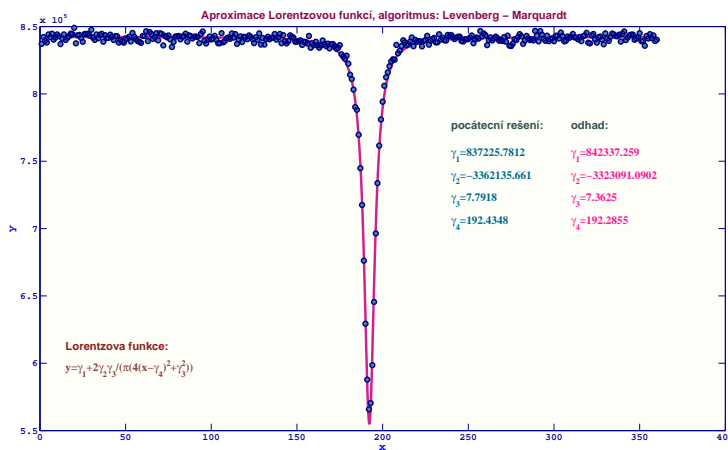
zení deformací ve spektrální oblasti (což je jeden z hlavních požadavků, které byly na filtrační proceduru kladeny) je nezbytné zvolit vhodnou oblast, kde filtrační účinnost vykazuje saturaci. Reciproká funkce poté ukazuje kvalitu filtru. Na obr. 3 vlevo jsou ukázány závislosti spektrálních pološírek před a po filtraci na SNR. Spektrální pološířky byly obdrženy za použití programu RECOIL bez fixovaných parametrů. Z této prokázané závislosti spektrálních pološírek na SNR pro filtrovaná data lze usoudit na vhodnost použití vyvinutého filtru.

3 Odhady parametrů pro fitování spekter

Cílem celého výzkumu nanomateriálů je charakterizovat a popsat vlastnosti, které zkoumané materiály mají. Za tímto účelem je potřebné přesné stanovení a proložení naměřeného Mössbauerovského spektra. Díky tomuto kroku se otevírá nová kapitola, která vyžaduje delšího testování. V této části poukážeme na zatím provedené možnosti fitování Mössbauerovského spektra. V budoucnosti, až budou provedeny všechny potřebné analýzy, bude celá metoda porovnána s již existujícím softwarem používaným pro tyto účely.

Na úvod této části musí být navíc podotknuto, že autoři zatím řeší úlohu zjednodušenou, kdy se snaží nafitovat spektrum obsahující pouze „jeden pík“. Úlohou složitější, která byla zadána jako řešení pozdější, je aproximace celé šíře spektra, které obsahuje „více píků“. Tuto úlohu zatím autoři tohoto příspěvku nezmiňují, ačkoliv jisté parciální výsledky již dosaženy byly.

Obecně lze tedy říci, že na fitování Mössbauerovských spekter bylo použito několik metod, z nichž v tomto příspěvku popíšeme zejména regresní model s podmínkou typu II, ve kterém bylo dosaženo nejhodnotnějších výsledků.



Obrázek 4: Aproximace dat Lorentzovou funkcí pomocí Levenberg–Marquardtova algoritmu.

3.1 Levenberg–Marquardt algoritmus

Pro fitování spekter existuje obecně řada metod, které lze použít a záleží pouze na experimentátorovi, kterou z nabízených metod zvolí. Prvně jsme zvolili Levenberg–Marquardtův algoritmus (viz např. [5]) — všeobecně používanou metodu na aproximaci měřených dat nelineární funkcí.

Levenberg–Marquardtův algoritmus se řadí do třídy nelineárních metod nejmenších čtverců, jehož cílem je odhadnout hodnoty neznámých parametrů zkoumané nelineární funkce. Tento algoritmus je kombinací dvou numerických metod (metoda největšího spádu a Gauss–Newtonova metoda).

Na obr. 4 jsou znázorněna naměřená data aproximovaná Lorentzovou funkcí

$$f(x, \gamma_1, \gamma_2, \gamma_3, \gamma_4) = \gamma_1 - \frac{2\gamma_2}{\pi} \frac{\gamma_3}{4(x - \gamma_4)^2 + \gamma_3^2}. \quad (1)$$

Dále je na obr. 4 vidět použité počáteční řešení i získaný odhad.

V další části odhadneme parametry pomocí nelineárních regresních modelů.

3.2 Model s podmínkou typu II

Máme měřené hodnoty Y_1, Y_2, \dots, Y_{360} odpovídající hodnotám Mössbauerovi spektroskopie $\beta_1, \beta_2, \dots, \beta_{360}$ v daných kanálech x_1, x_2, \dots, x_{360} .

Cílem je najít odhady neznámých parametrů $\gamma_1, \gamma_2, \gamma_3$ a γ_4 odpovídající skutečným hodnotám parametrů funkce (1) popisující daný jev. Naměřené hodnoty byly nejprve očištěny pomocí navrženého filtru, viz [4], a korespondují s navrženým teoretickým předpisem pro tento jev.

K nalezení odhadů neznámých parametrů použijeme regresní model neúplného přímého měření vektorového parametru se systémem podmínek typu II, viz [3], str. 189 (odvození odhadů a varianční matice pro tzv. 3. model).

Uvažujme model neúplného přímého měření $\mathbf{Y} \sim (\boldsymbol{\beta}, \boldsymbol{\Sigma})$ s podmínkou

$$\begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\gamma} \end{pmatrix} \in \left\{ \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} : \mathbf{b} + \mathbf{B}\mathbf{u} + \mathbf{G}\mathbf{v} = \mathbf{0} \right\}; \quad (2)$$

zde $\boldsymbol{\beta}$ je neznámý n -rozměrný parametr, $\boldsymbol{\gamma}$ je neznámý k -rozměrný parametr, $\boldsymbol{\Sigma}$ je dána $n \times n$ p. d. matice, \mathbf{B} je dána $n \times n$ matice a \mathbf{G} je dána $n \times k$ matice. Potom BLUE vektoru $\boldsymbol{\beta}$ a $\boldsymbol{\gamma}$ je

$$\begin{pmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\boldsymbol{\gamma}} \end{pmatrix} = \begin{pmatrix} \mathbf{I} - \boldsymbol{\Sigma}\mathbf{B}'\mathbf{Q}_{1,1}\mathbf{B} \\ -\mathbf{Q}_{2,1}\mathbf{B} \end{pmatrix} \mathbf{Y} + \begin{pmatrix} \mathbf{I} - \boldsymbol{\Sigma}\mathbf{B}\mathbf{Q}_{1,1} \\ -\mathbf{Q}_{2,1} \end{pmatrix} \mathbf{b} \quad (3)$$

kovarianční matice je

$$\text{Var} \begin{pmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\boldsymbol{\gamma}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Sigma} - \boldsymbol{\Sigma}\mathbf{B}'\mathbf{Q}_{1,1}\mathbf{B}\boldsymbol{\Sigma}, & -\boldsymbol{\Sigma}\mathbf{B}'\mathbf{Q}_{1,2} \\ -\mathbf{Q}_{2,1}\mathbf{B}\boldsymbol{\Sigma}, & \mathbf{Q}_{2,2} \end{pmatrix}, \quad (4)$$

kde

$$\begin{pmatrix} \mathbf{Q}_{1,1} & \mathbf{Q}_{1,2} \\ \mathbf{Q}_{2,1} & \mathbf{Q}_{2,2} \end{pmatrix} = \begin{pmatrix} \mathbf{B}\Sigma\mathbf{B}', & \mathbf{G} \\ \mathbf{G}', & 0 \end{pmatrix}^{-1}.$$

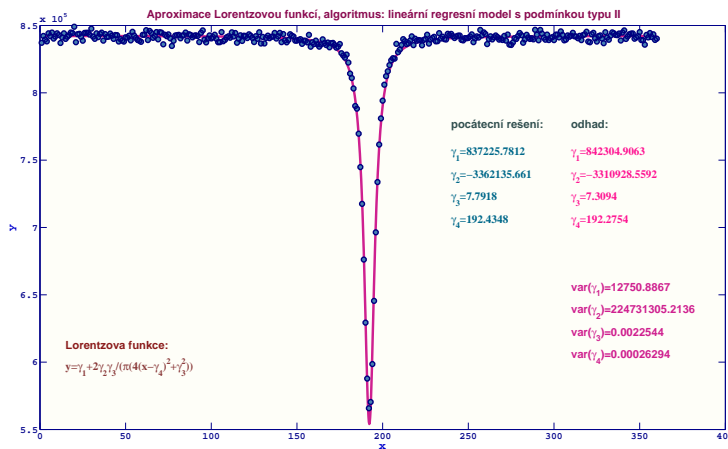
Parametry $\gamma_1, \gamma_2, \gamma_3$ a γ_4 musí splnit pro $x_i = i = 1, \dots, n$ podmínky

$$g_i(\boldsymbol{\beta}, \Theta) = \beta_i - \gamma_1 - \frac{2\gamma_2}{\pi} \frac{\gamma_3}{4(x_i - \gamma_4)^2 + (\gamma_3)^2} = 0. \quad (5)$$

Nelineární podmínky $\mathbf{g}(\boldsymbol{\beta}, \boldsymbol{\gamma}) = (g_1(\boldsymbol{\beta}, \boldsymbol{\gamma}), \dots, g_{360}(\boldsymbol{\beta}, \boldsymbol{\gamma}))' = \mathbf{0}$ je třeba linearizovat pomocí Taylorova rozvoje — v lineární verzi mají podmínky tvar $\mathbf{B}\delta\boldsymbol{\beta} + \mathbf{G}\delta\boldsymbol{\gamma} + \mathbf{b} = \mathbf{0}$, kde $\mathbf{B} = \frac{\partial \mathbf{g}(\boldsymbol{\beta}^0, \boldsymbol{\gamma}^0)}{\partial \boldsymbol{\beta}'}$, $\mathbf{G} = \frac{\partial \mathbf{g}(\boldsymbol{\beta}^0, \boldsymbol{\gamma}^0)}{\partial \boldsymbol{\gamma}'}$, a $\mathbf{b} = \mathbf{g}(\boldsymbol{\beta}^0, \boldsymbol{\gamma}^0)$ v přibližném řešení $(\boldsymbol{\beta}^0, \boldsymbol{\gamma}^0)$. Po dosazení do vzorců (3) a (4) dostáváme odhady parametrů $\boldsymbol{\gamma}$ a odhad jejich přesnosti:

3.3 Další výzkum

Atoři se v současnosti soustředí na ladění programových kódů a sestavení procedury, která by plně vyhovovala pro fitování Mössbauerovských spekter obsahujících pouze „jeden pík“. Dále autoři pokračují ve vývoji modelů umožňujících řešení modifikované složitější úlohy s „více píky“. Dalším cílem autorů bude pro potřeby nanocentra připravit software, který automaticky navrhne pro zadaná data a zadaný typ aproximující funkce vhodný algoritmus a nalezne odhady s co nejmenší disperzí. Autoři předpokládají, že řešení bude založeno na teorii nelineárních regresních modelů. Zde je ale důležité si uvědomit problémy, které při výpočtu nastávají. První problém nastává u linearizace. Vlastní výpočet parciálních derivací je možné provést pomocí symbolic computing. Závažnějším problémem ale je nalezení vhodného počátečního řešení, které leží v linearizační oblasti. Počáteční řešení může být hledáno



Obrázek 5: Aproximace dat Lorentzovou funkcí pomocí lineárního modelu s podmínkou typu II.

např. LM algoritmem. Skutečnost, že nalezený odhad v modelu s podmínkou typu II splňuje pouze linearizované podmínky a nikoliv původní nelineární podmínky, je možno řešit doiterováním např. pomocí heuristických algoritmů. Jednou z metod, kterou autoři připravují do vyvíjeného software, je ortogonální regrese. U dat aproximovaných Lorentzovou funkcí byl také navržen dvouetapový model s podmínkou typu II, kde nejprve dojde v 1. etapě k nalezení polohy „píku“ a ve druhé etapě jsou hledány ostatní neznámé parametry Lorentzovy funkce. V úlohách, kdy je třeba co nejpřesněji odhadnout jen některý parametr, lze využít \mathbf{H} -optimální odhady — viz [6]. Autoři očekávají, že se jim podaří ve vyvíjeném software vytvořit rozhodovací strom, který najde vhodný přístup k aproximaci na základě měř krivosti zadané funkce.

4 Závěr

Autoři tohoto příspěvku si kladli za cíl seznámit širokou veřejnost s problematikou materiálového výzkumu, se kterým je úzce spjata řada sofistikovaných problémů, které vyžadují interdisciplinární spolupráci za účelem jejich vyřešení. V průběhu práce byl vytvořen filtr vyhovující přesně podmínkám kladeným odborníky z nanomateriálového výzkumu a navrhnuty postupy, jak dále daná data fitovat za účelem přesného odhadu parametrů funkční závislosti.

Literatura

- [1] Stevens J.G., Khasanov A.M., Miller J.W., Pollak H., Li Z. (eds.) (2005). *Mössbauer mineral handbook*. MEDC, 624.
- [2] *Labview 8.20, User manual, National Instrument* (2006).
- [3] Kubáček L., Kubáčková L. (2000). *Statistics and metrology (in Czech)*. Olomouc. Publishing House of Palacký University.
- [4] Procházka R., Tuček P., Tuček J., Marek J., Mashlan P., Pechoušek J. *Statistical analysis and digital processing of Mössbauer spectra*, v rec. řízení.
- [5] Seber G.A.F., Wild C.J. (2003). *Nonlinear regression*. John Wiley & Sons, New Jersey.
- [6] Kubáček L., Marek J. (2004). *Partial optimum estimator in two stage regression model with constraints and a problem of equivalence*. Math. Slovaca **54**.

Poděkování: Práce na tomto článku byla podporována výzkumnými granty MSM6198959218, 1M6198959201 a MSM6198959214.

Adresa: P. Tuček, PřF UP, Katedra geoinformatiky, Tř. Svobody 26, 771 46 Olomouc; J. Tuček, R. Procházka, Centrum pro nanomateriálový výzkum, Šlechtitelů 11, 783 71 Olomouc; J. Tuček, PřF UP, Katedra experimentální fyziky, Tř. Svobody 26, 771 46, Olomouc; P. Tuček, J. Marek, PřF UP, Katedra matematické analýzy a aplikací matematiky, Tomkova 40, 779 00 Olomouc - Hejčín

E-mail: pavel.tucek@upol.cz

ADAPTIVNÍ STOCHASTICKÉ ALGORITMY V NELINEÁRNÍ REGRESI

Josef Tvrđík

Klíčová slova: Nelineární regrese, odhad parametrů, globální optimalizace, diferenciální evoluce, řízené náhodné prohledávání, adaptivní algoritmy.

Abstrakt: Příspěvek se zabývá využitím adaptivních stochastických algoritmů v odhadu parametrů nelineárních regresních modelů. Adaptivní algoritmus řízeného náhodného prohledávání (CRS) je experimentálně porovnán s novými adaptivními verzemi diferenciální evoluce (DE) na úlohách z referenční databáze NIST. Pouze některé verze DE jen na několika málo úlohách předčily algoritmus CRS.

Abstract: Self-adaptive stochastic algorithms are applied to the estimation of parameters in non-linear regression models. Self-adaptive variant of controlled random search (CRS) is compared experimentally with novel self-adaptive variants of differential evolution (DE) using tasks of NIST reference datasets. Several variants of DE outperformed CRS in a few tasks only.

1 Úvod

Aditivní nelineární regresní model můžeme zapsat v následující formě

$$Y_i = g(\vec{x}_i, \vec{\beta}) + \epsilon_i, \quad i = 1, 2, \dots, n, \quad (1)$$

kde g je daná funkce, $\vec{x}_i^T = (x_1, x_2, \dots, x_k)$ je i -tý řádek matice návrhu \vec{X} , která je typu $n \times k$, $\vec{\beta}^T = (\beta_1, \beta_2, \dots, \beta_d)$ je vektor parametrů a ϵ_i jsou nezávislé náhodné veličiny s nulovou střední hodnotou a shodným rozptylem, $E(\epsilon_i) = 0$, $var(\epsilon_i) = \sigma^2$. Odhad parametrů modelu metodou nejmenších čtverců znamená nalézt hodnoty $\vec{\beta}$, pro které je součet druhých mocnin residuí (2) minimální.

$$Q(\vec{\beta}) = \sum_{i=1}^n [Y_i - g(\vec{x}_i, \vec{\beta})]^2 \quad (2)$$

Na rozdíl od lineární regrese je nalezení správných hodnot odhadů obtížný algoritmický problém, neboť funkce (2) může mít více lokálních minim a její tvar je závislý na datech, což působí potíže při hledání odhadů standardně užívanými iterativními deterministickými algoritmy (Levenberg-Marquardt, Gauss-Newton). Tyto algoritmy hledají lokální minimum, takže výsledek záleží na volbě počátečních hodnot, v některých úlohách tvar funkce (2) může dokonce způsobit numerický kolaps iteračního procesu. Odhad parametrů nelineárních regresních modelů je problém globální optimalizace.

2 Problém globální optimalizace

Problém globální optimalizace lze jednoduše formulovat: Pro danou účelovou funkci

$$f : D \rightarrow \mathbb{R}, \quad D \subset \mathbb{R}^d,$$

hledáme bod $\vec{x}^* = \arg \min_{\vec{x} \in D} f(\vec{x})$, tj. bod, ve kterém má funkce nejmenší hodnotu. Bod \vec{x}^* je nazýván bodem globálního minima (global minimum point) a D je prohledávaný obor (doména).

V nelineární regresi řešíme úlohu, kdy doména je vymezena jako d -rozměrný kvádr $D = \prod_{j=1}^d [a_j, b_j]$, $a_j < b_j$, $j = 1, 2, \dots, d$, účelová funkce $f(\vec{x})$ je spojitá a umíme ji vyhodnotit v každém bodě $\vec{x} \in D$.

Neexistuje polynomiálně složitý deterministický algoritmus pro obecné řešení problému globální optimalizace. Řešení je nutno hledat heuristicky pomocí stochastických algoritmů. Tyto algoritmy jsou často inspirovány přírodou, např. ochlazením tuhého tělesa, evolucí populace jedinců, poznatky z genetiky nebo imunologie, chováním jedinců ve společenství apod. Algoritmům modelujícím vývoj populace se říká evoluční algoritmy. Tyto algoritmy používají tzv. evoluční operátory, zejména selekci, křížení a mutaci.

3 Diferenciální evoluce a adaptace řídicích parametrů

Algoritmus diferenciální evoluce (DE) navrhli Storn a K. Price v devadesátých letech [16] a od té doby se DE stala velmi užívaným algoritmem pro hledání globálního minima [11]. Princip tohoto algoritmu ukazuje jeho zápis v pseudokódu na obr. 1.

```

1  generate an initial population  $P = (\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N)$ ,  $\vec{x}_i \in D$ ;
2  repeat
3    for  $i := 1$  to  $N$  do
4      generate a new trial vector  $\vec{y}$ ;
5      if  $f(\vec{y}) < f(\vec{x}_i)$  then insert  $\vec{y}$  into new generation  $Q$ 
6      else insert  $\vec{x}_i$  into new generation  $Q$ 
7      endif
8    endfor
9     $P := Q$ ;
10 until stopping condition

```

Obrázek 1: Algoritmus diferenciální evoluce.

Nový zkusmý bod \vec{y} na řádce 4 se generuje využitím mutace a křížení. Selektce jedince do další generace probíhá turnajovou metodou, vítězí lepší z dvojice (řádky 5 až 7). Podle užití varianty mutace a křížení se označují varianty diferenciální evoluce zkratkou $DE/m/n/k$, kde m označuje variantu mutace, n počet dvojic bodů pro difference užití v mutaci a k užitý způsob křížení.

Bylo navrženo několik variant mutace pro diferenciální evoluci [16, 11, 3]. V této práci jsou užity dva typy mutace, a to nejobvyklejší klasická [16] označovaná zkratkou $DE/RAND/1$

$$\vec{v} = \vec{r}_1 + F(\vec{r}_2 - \vec{r}_3), \quad (3)$$

$\vec{r}_1 \neq \vec{r}_2 \neq \vec{r}_3 \neq \vec{x}_i$ jsou vybrány náhodně z populace P , $F > 0$ je vstupní řídicí parametr algoritmu ovlivňující mutaci. Druhou variantou mutace je *DE/RA-NDrl/1* navržená v [5]. Od předchozí varianty se liší jen tím, že bod \vec{r}_1 je bod s nejnižší funkční hodnotou mezi náhodně vybranými body $\{\vec{r}_1, \vec{r}_2, \vec{r}_3\}$, tj. $\vec{r}_1 = \arg \min_{i \in \{1,2,3\}} f(\vec{r}_i)$.

Bod \vec{y} vzniká křížením aktuálního bodu \vec{x}_i ze staré generace a bodu \vec{v} vzniklého mutací. Křížení je řízeno parametrem $CR \in [0, 1]$. V binomickém křížení se přepisují náhodně vybrané prvky vektoru \vec{x}_i s pravděpodobností CR . V exponenciálním křížení se přepisuje L sousedních souřadnic \vec{x}_i , sousedství se uvažuje cyklicky, tj. poslední souřadnice sousedí s první. Počet přepisovaných pozic L klesá exponenciálně v závislosti na CR [23]. V obou typech křížení je nejméně jedna souřadnice bodu \vec{x}_i přepsána odpovídající souřadnicí \vec{v} .

Diferenciální evoluce má oproti jiným evolučním algoritmům méně řídicích parametrů. Volíme jen velikost populace, typ mutace a křížení a hodnoty parametrů F a CR . Účinnost hledání globálního minima je však velmi silně závislá na vhodném nastavení hodnot řídicích parametrů, zejména F a CR . Obvykle užívaným postupem je hledání vhodných hodnot těchto parametrů pro řešený problém časově náročnou metodou pokus-omyl. Různá doporučení, většinou značně neurčitá, lze najít v [3, 4, 11, 16]. Proto byly zkoumány postupy adaptivního nastavení hodnot řídicích parametrů. Ali a Törn [1] navrhli jednoduché pravidlo měnící hodnoty F během hledání. Další možnosti adaptace hodnot řídicích parametrů jsou shrnuty v článku Liu a Lampinen [8]. Nedávno Quin a Suganthan [13] užíli adaptivní volbu typu mutace kombinovanou se řízeně náhodným nastavováním hodnot F a CR .

Evoluční adaptaci parametrů F a CR navrhli Brest et al. [2]. Hodnoty F a CR se inicializují náhodně pro každý bod populace a přežívají s jedinci populace. Mohou však být i změněny mutací v každé generaci, pravděpodobnost mutace se zadává jako vstupní parametr algoritmu. Adaptace hodnot F a CR soutěží několika variant jejich nastavení byla navržena v [18]. Z H různých nastavení vybíráme náhodně s pravděpodobnostmi q_h , $h = 1, 2, \dots, H$. Tyto pravděpodobnosti se mění podle dosavadní úspěšnosti nastavení. Za úspěšné je považováno h -té nastavení tehdy, když generuje nový zkusmý bod \vec{y} , pro který je $f(\vec{y}) < f(\vec{x}_i)$. Pravděpodobnost q_h se vyhodnotí jako relativní četnost

$$q_h = \frac{n_h + n_0}{\sum_{j=1}^H (n_j + n_0)}, \quad (4)$$

kde n_h je aktuální počet úspěchů h -tého nastavení a $n_0 > 0$. Volba $n_0 > 1$ zabraňuje dramatickým změnám q_h jedním náhodným úspěchem. Abychom předešli degeneraci procesu hledání, jsou aktuální hodnoty pravděpodobností změněny na jejich počáteční hodnoty $q_h = 1/H$ tehdy, když některá hodnota

pravděpodobnosti q_h klesne pod zadanou hodnotu $\delta > 0$. Několik variant takového soutěživého nastavení parametrů F a CR v diferenciální evoluci s binomickým křížením bylo porovnáno numericky na řadě testovacích úloh [19] a dvě nejlépejší varianty implementované v Matlabu a C++ jsou volně dostupné v knihovně programů [15].

Mutace podle (3) může způsobit, že nový zkušmý vektor \vec{y} je mimo doménu D . V takovém případě může být hodnota souřadnice, pro kterou je $y_j < a_j$ nebo $y_j > b_j$, nahrazena hodnotou překlopenou do D , souměrnou vzhledem k a_j nebo b_j [7]. Tento postup je užít v implementaci testovaných algoritmů v této práci i v programové knihovně [15].

4 Řízené náhodné prohledávání se soutěží heuristik

Algoritmus řízeného náhodného prohledávání (CRS, controlled random search) navrhl před více jak 30 léty W. L. Price [12]. Zápis algoritmu je na obr. 2. Na rozdíl od diferenciální evoluce, u CRS je v každé generaci nahrazen nejhorší bod populace. Pro generování zkušmého bodu \vec{y} Price užíval reflexi simplexu ($d + 1$ bodů náhodně vybraných z populace) podle Nelder a Meada [10], později byla navržena řada zlepšených variant algoritmu CRS, viz např. [1, 6].

```

1  generate  $P$  (population of  $N$  points in  $D$  at random);
2  find  $\vec{x}_{\max}$  (the point in  $P$  with the highest function value);
3  repeat
4    generate a new trial point  $\mathbf{y} \in D$  by using a local heuristic;
5    if  $f(\vec{y}) < f(\vec{x}_{\max})$  then
6       $\vec{x}_{\max} := \vec{y}$ ;
7      find new  $\vec{x}_{\max}$ ;
8    endif
9  until stopping condition;
```

Obrázek 2: Algoritmus řízeného náhodného prohledávání.

Adaptivní verze CRS založená na soutěži více lokálních heuristik pro generování zkušmého bodu \vec{y} byla navržena v [17], pravidla soutěže jsou podobná jako u soutěže nastavení v diferenciální evoluci s tím rozdílem, že pravděpodobnost výběru lokální heuristiky q_i je možno počítat i s využitím vážení velikosti zlepšení funkční hodnoty podle vztahů

$$w_i = \frac{f_{\max} - \max(f(\vec{y}), f_{\min})}{f_{\max} - f_{\min}}, \quad q_i = \frac{W_i + w_0}{\sum_{j=1}^H (W_j + w_0)},$$

kde f_{\max} a f_{\min} jsou největší a nejmenší funkční hodnoty v populaci, W_j je součet vah w_j v předcházejícím hledání, $w_0 > 0$ je vstupní parametr algoritmu.

5 Aplikace v odhadu regresních parametrů

Pro odhad parametrů nelineárních regresních modelů byla navržena adaptivní verze CRS se čtyřmi soutěžícími heuristikami [20]. V dalším textu je tato varianta algoritmu označena zkratkou *CRS4*. V tomto algoritmu je užitá i adaptivní podmínka ukončení hledání. Adaptace podmínky ukončení je založena na tom, aby rozdíl mezi největší a nejmenší hodnotou indexu determinace v populaci, tj. $R_{\max}^2 - R_{\min}^2$ byl výrazně menší než $1 - R_{\max}^2$. Algoritmus adaptace podmínky ukončení s řídicími vstupními parametry ε_0 a γ je zapsán na obr. 3.

```

1  generate initial population  $P$ ;
2   $\varepsilon := \varepsilon_0$ ;
3  pass := false;
4  while  $1 - R_{\max}^2 < \gamma \times \varepsilon$  & pass = false (outer loop)
5      while  $R_{\max}^2 - R_{\min}^2 > \varepsilon$  (inner loop)
6          stochastic optimization algorithm (CRS, DE, ...);
7          pass := true;
8      endwhile; (end of inner loop)
9      if pass = false then
10          $\gamma := \gamma/10$ ; (stopping condition softened)
11      endif;
12      if  $1 - R_{\max}^2 < \gamma \times \varepsilon$  & pass = true then
13          $\varepsilon := \varepsilon/10$ ; (stopping condition made more strict)
14         pass := false;
15      endif;
16  endwhile; (end of outer loop)

```

Obrázek 3: Adaptace podmínky ukončení pro nelineární regresi.

6 Numerické porovnání algoritmů

Algoritmy byly porovnávány na souboru úloh nelineární regrese NIST [14] obsahujícím 27 úloh tří úrovní obtížnosti. Úloha Lanczos1 s uměle generovanými daty nebyla do experimentálního porovnání zařazena, neboť u ní má náhodná složka modelu extrémně nízký rozptyl ($1 - R^2 \approx 1e - 26$), takže je spíše testem numerické odolnosti algoritmu než jeho schopnosti nalézt globální minimum. Pro každou úlohu a algoritmus bylo provedeno 100 opakování, v každém opakování byla sledována časová náročnost, vyjádřená počtem vyhodnocení účelové funkce (2), a kvalita nalezeného řešení hodnocená počtem významných číslic shodných s tzv. certifikovanou hodnotou residuálního součtu čtverců [9]. Pokud se nalezené řešení shodovalo alespoň ve čtyřech číslicích, bylo považováno za správné.

Pro porovnání algoritmů bylo užito jednoduché kritérium $Q_m = \overline{ne}/R$, kde R je procento běhů, ve kterých bylo nalezeno správné řešení a \overline{ne} je průměrný počet vyhodnocení účelové funkce v úspěšných bězích [3]. Hodnoty Q_m jsou uvedeny v tabulce 1. Nejmenší hodnota Q_m na každém řádku je podtržena. Pokud hodnota Q_m pro některou z variant DE byla nižší než pro *CRS4*, je uvedena tučně. Označení úloh je shodné s [14]. Algoritmy v ta-

	<i>CRS4</i>	<i>cb9rl</i>	<i>ce9rl</i>	<i>cb6e6rl</i>	<i>cb9e9rl</i>	<i>brest</i>	<i>brcmp</i>	<i>brcmprl</i>
Chwirut1	<u>19</u>	25	30	24	26	60	42	28
Chwirut2	<u>19</u>	24	30	24	27	58	43	27
Danwood	<u>12</u>	17	19	16	18	40	26	17
Gauss1	91	108	121	100	116	171	135	<u>90</u>
Gauss2	<u>96</u>	150	180	141	164	262	217	135
Lanczos3	<u>303</u>	1231	3606	∞	1368	47038	178050	20547
Misra1a	<u>18</u>	36	92	88	36	106	98	47
Misra1b	<u>15</u>	29	81	67	33	72	86	44
Enso	156	187	181	158	189	248	182	<u>136</u>
Gauss3	<u>105</u>	127	170	125	143	387	298	172
Hahn1	<u>132</u>	146	274	282	172	864	933	801
Kirby2	65	<u>59</u>	92	92	65	294	375	165
Lanczos2	<u>305</u>	1077	4004	229050	1268	33880	223740	231690
Mgh17	<u>90</u>	160	292	531	174	530	607	308
Misra1c	<u>19</u>	43	260	312	50	198	289	101
Misra1d	<u>18</u>	40	220	249	45	165	201	96
Nelson	<u>49</u>	138	1225	2245	147	310	1392	200
Roszman1	34	<u>26</u>	37	<u>26</u>	<u>29</u>	89	60	41
Bennett5	<u>366</u>	85944	8480	1420	2700	115621	8500	∞
Boxbod	<u>8</u>	13	14	12	14	23	16	12
Eckerle4	17	18	19	<u>16</u>	19	22	20	<u>14</u>
Mgh09	<u>88</u>	150	292	593	164	467	1140	248
Mgh10	<u>210</u>	924	64260	∞	1154	13027	100880	16454
Rat42	<u>19</u>	21	26	21	24	53	37	25
Rat43	29	<u>23</u>	33	<u>25</u>	<u>27</u>	74	55	36
Thurber	97	<u>96</u>	155	115	106	557	412	242

Tabulka 1: Porovnání algoritmů na úlohách NIST – hodnoty Q_m .

bulce jsou označeny zkratkami. *CRS4* je řízené náhodné prohledávání [20], ostatní algoritmy jsou adaptivní varianty DE. Počáteční písmeno *c* je užito pro varianty se soutěží nastavení, písmena *b* nebo *e* uvnitř zkratky vyznačují typ křížení, následující číslice je počet soutěžících nastavení a přípona *rl* vyznačuje užití mutace podle [5]. Označení *brest* je užito pro evoluční adaptaci řídicích parametrů navrženou Brestem et al. [2], verze evoluční adaptace řídicích parametrů kombinované se soutěží obou typů křížení jsou označeny *brcmp* a *brcmprl*. Všechny tyto adaptivní varianty DE jsou podrobněji popsány v [21], včetně užitých hodnot řídicích parametrů. U všech porovnávaných algoritmů byla užita stejná adaptivní podmínka ukončení s parametry $\varepsilon_0 = 1e - 9$, $\gamma = 1e + 7$. Velikost populace pro *CRS4* byla zvolena $N = 10d$, pro všechny varianty DE byla velikost populace $N = \min(\max(20, 5d), 60)$, neboť takto zvolená velikost populace se osvědčila v jiných úlohách [21].

Testované adaptivní varianty DE využívající současně obou typů křížení

prokázaly dobrou výkonnost na dvou obtížných sadách testovacích úloh [21], a tak se očekávalo, že budou úspěšné i odhadu parametrů nelineárních regresních modelů. Jak ukazují výsledky v tabulce 1, toto očekávání se nenaplnilo. Projevil se známý a nepříjemný důsledek tzv. No Free Lunch teorémů [22], tj. pořadí výkonnosti dvou algoritmů může být opačné pro různé skupiny problémů. Pouze u sedmi z 26 úloh byla některá z variant DE úspěšnější než CRS_4 . Mezi těmito variantami nebyla žádná, ve které se neuzívá binomické křížení a mutace $RANDr1$. Naopak, na řadě úloh byly hodnoty Q_m u některých variant DE řádově horší než u CRS_4 , ve třech případech dokonce ani v jednom ze 100 běhů nebylo nalezeno správné řešení (odpovídající hodnoty Q_m jsou vyznačeny symbolem ∞). CRS_4 měl ve 22 úlohách úspěšnost 100% a minimální hodnota úspěšnosti byla 86%, ve většině případů při nižších časových nárocích než kterákoliv varianta diferenciální evoluce.

7 Závěr

Testované varianty DE jen zřídka předčily algoritmus CRS_4 , navíc v řadě úloh byla jejich účinnost výrazně horší. Využití diferenciální evoluce s exponenciálním křížením zřejmě není cesta k rychlým a spolehlivým algoritmům pro odhad parametrů nelineárních regresních modelů. Pokud budeme hledat algoritmus efektivnější než CRS_4 , patrně bude nutno uvažovat o hybridních algoritmech kombinujících několik stochastických algoritmů k získání spolehlivých startovních hodnot a hledání dokončovat algoritmy konvergujícími k lokálnímu minimu.

Literatura

- [1] Ali M.M., Törn A. (2004). *Population set based global optimization algorithms: some modifications and numerical studies*. Computers and Operations Research **31**, 1703–1725.
- [2] Brest J., Greiner S., Boškovič B., Mernik M., Žumer V. (2006). *Self-adapting control parameters in differential evolution: a comparative study on numerical benchmark problems*. IEEE Transactions on Evolutionary Computation **10**, 646–657.
- [3] Feoktistov V. (2006) *Differential evolution in search of solution*. Springer.
- [4] Gämperle R. Müller S.D., Koumoutsakos P. (2002). *A parameter study for differential evolution*. In: Grmela A., Mastorakis N.E. (eds), *Advances in Intelligent Systems Fuzzy Systems, Evolutionary Computing*, WSEAS Press, Athens, 293–298.
- [5] Kaelo P., Ali M.M. (2006). *A numerical study of some modified differential evolution algorithms*. European J. Operational Research, **169**, 1176–1184.
- [6] Krivý I., Tvrdlík J. (1995). *The controlled random search algorithm in optimizing regression models*. Comp. Statist. & Data Anal. **20**, 229-234.
- [7] Kvasnička V., Pospíchal J., Tiňo P. (2000). *Evolučné algoritmy*. STU. Bratislava.

- [8] Liu J., Lampinen J. (2005). *A fuzzy adaptive differential evolution algorithm*. *Soft Computing* **9**, 448–462.
- [9] McCullough B.D., Wilson B. (2005). *On the accuracy of statistical procedures in Microsoft Excel 2003*. *Comput. Statist. and Data Anal.* **49**, 1244–1252.
- [10] Nelder J.A., Mead R. (1964). *A simplex method for function minimization*. *Computer J.* **7**, 308–313.
- [11] Price K.V., Storn R., Lampinen J. (2005). *Differential evolution: a practical approach to global optimization*. Springer.
- [12] Price, W.L., (1977) *A controlled random search procedure for global optimization*. *Computer J.* **20**, 367–370.
- [13] Quin A.K., Suganthan P.N. (2005). *Self-adaptive differential evolution for numerical optimization*. In: IEEE Congress on Evolutionary Computation, 1785–1791.
- [14] *Statistical Reference Datasets. Nonlinear Regression*. NIST Information Technology Laboratory. <http://www.itl.nist.gov/div898/strd/>.
- [15] *Stochastic Algorithms for Global Optimization - MATLAB and C++ library*. <http://albert.osu.cz/oukip/optimization/>
- [16] Storn R., Price K.V. (1997). *Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces*. *J. Global Optimization* **11**, 341–359.
- [17] Tvrđík, J., Křivý, I., Mišík, L. (2001). *Evolutionary algorithm with competing heuristics*. In: Ošmera, P. (ed) MENDEL 2001, 7th International Conference on Soft Computing. Technical University, Brno, 58–64.
- [18] Tvrđík J. (2006). *Competitive differential evolution*. In: Matoušek R. and Ošmera P. eds, MENDEL 2006, 12th International Conference on Soft Computing. University of Technology, Brno, 7–12.
- [19] Tvrđík J. (2007). *Differential evolution with competitive setting of its control parameters*. *TASK Quarterly* **11**, 169–179.
- [20] Tvrđík J., Křivý I., Mišík L. (2007). *Adaptive population-based search: application to estimation of nonlinear regression parameters*. *Comput. Statist. and Data Anal.* **52**, 713–724.
- [21] Tvrđík J. (2008). *Adaptive differential evolution and exponential crossover*. in: Proceedings of IMCSIT 2008, PTI, Wisla, 927–931.
- [22] Wolpert D.H., Macready W. (1997). *No free lunch theorems for optimization*. *IEEE Transactions on Evolutionary Computation*, **1**, 67–82.
- [23] Zaharie D. (2007). *A comparative analysis of crossover variants in differential evolution*. Proceedings of IMCSIT 2007, PTI, Wisla, 171–181.

Poděkování: Tato práce byla podporována grantem GAČR 201/08/0472 a MSM 6198898701.

Adresa: J. Tvrđík, PřF OU, KIP, 30. dubna 22, 701 03 Ostrava

E-mail: josef.tvrdik@osu.cz

METODA MLE PRO LOGARITMICKO-NORMÁLNÍ ROZDĚLENÍ S PRAHEM

Vladimír Václavík

Klíčová slova: logaritmicko-normální rozdělení; prahový parametr; metoda maximální věrohodnosti; LMLE.

Abstrakt: Za odhad parametrů metodou maximální věrohodnosti (MLE) logaritmicko-normálního rozdělení s prahovým parametrem se místo globálního maxima používá lokální maximum věrohodnostní funkce a tato metoda se většinou označuje LMLE. S nalezením lokálního maxima však mohou být spojeny obtíže. Předkládáme metodu, která odhady LMLE najde, pokud existují.

Abstract: In the paper is suggested a method that always finds MLE of the parameters of three parameter lognormal distribution which is based on local maximum likelihood function.

1 Logaritmicko-normální rozdělení

Náhodná veličina X , která se řídí tříparametrickým logaritmicko-normálním rozdělením $LN(\gamma, \mu, \sigma^2)$ s prahovým parametrem $\gamma \in \mathbb{R}$, parametrem $\mu \in \mathbb{R}$ (e^μ je parametrem měřítka) a parametrem tvaru $\sigma > 0$, má hustotu pravděpodobnosti

$$f(x) = \begin{cases} 0 & \text{pro } x \leq \gamma, \\ \frac{1}{\sigma\sqrt{2\pi}} \frac{1}{x-\gamma} \exp\left\{-\frac{1}{2\sigma^2}[\ln(x-\gamma) - \mu]^2\right\} & \text{pro } x > \gamma. \end{cases}$$

Pro takto definovanou veličinu X platí $\ln(X - \gamma) \sim N(\mu, \sigma^2)$.

2 Metoda maximální věrohodnosti

Popis metody MLE pro tříparametrické logaritmicko-normální rozdělení poskytl např. Calitz [1], uvádí i stručnou historii této metody. Metodě se věnovalo mnoho dalších autorů, např. Cohen a Whitten [3] a Kane [5].

Pro pevné hodnoty x_1, \dots, x_n má (logaritmus) věrohodnostní funkce tvar

$$\begin{aligned} L(\gamma, \mu, \sigma) &= \ln \prod_{i=1}^n f(x_i) \\ &= -\frac{n}{2} \ln 2\pi - n \ln \sigma - \sum_{i=1}^n \ln(x_i - \gamma) - \frac{1}{2\sigma^2} \sum_{i=1}^n [\ln(x_i - \gamma) - \mu]^2, \end{aligned}$$

kde $\gamma < x_{\min} = \min\{x_1, \dots, x_n\}$, $\mu \in \mathbb{R}$ a $\sigma > 0$.

Věrohodnostní funkce je shora neomezená, pro $\gamma \rightarrow x_{\min}^-$, $\mu \rightarrow -\infty$ a $\sigma \rightarrow +\infty$ je $L \rightarrow +\infty$, globální maximum tedy neexistuje (viz Hill [3]). Za odhad parametrů γ , μ a σ metodou maximální věrohodnosti se proto bere bod lokálního maxima věrohodnostní funkce (metoda LMLE), který budeme hledat řešením rovnic (nutná podmínka lokálního extrému na otevřené množině)

$$\frac{\partial L}{\partial \gamma} = \sum_{i=1}^n (x_i - \gamma)^{-1} + \frac{1}{\sigma^2} \sum_{i=1}^n \frac{\ln(x_i - \gamma) - \mu}{x_i - \gamma} = 0 \quad (1)$$

$$\frac{\partial L}{\partial \mu} = \frac{1}{\sigma^2} \sum_{i=1}^n [\ln(x_i - \gamma) - \mu] = 0 \quad (2)$$

$$\frac{\partial L}{\partial \sigma} = \frac{n}{\sigma} - \frac{1}{\sigma^3} \sum_{i=1}^n [\ln(x_i - \gamma) - \mu]^2 = 0. \quad (3)$$

Z rovnic (2) a (3) snadno vyjádříme odhady $\hat{\mu}$ a $\hat{\sigma}$ parametrů μ a σ

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n \ln(x_i - \hat{\gamma}), \quad (4)$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n [\ln(x_i - \hat{\gamma}) - \hat{\mu}]^2, \quad (5)$$

kde $\hat{\gamma}$ je odhadem parametru γ a získáme jej následujícím způsobem. Označme $\hat{\mu}_\gamma = \frac{1}{n} \sum_{i=1}^n \ln(x_i - \gamma)$ a $\hat{\sigma}_\gamma^2 = \frac{1}{n} \sum_{i=1}^n [\ln(x_i - \gamma) - \hat{\mu}_\gamma]^2$, $\hat{\sigma}_\gamma = \sqrt{\hat{\sigma}_\gamma^2}$. Dosazením $\hat{\mu}_\gamma$ a $\hat{\sigma}_\gamma^2$ za μ a σ^2 do věrohodnostní funkce snadnou úpravou získáme

$$\lambda(\gamma) = L(\gamma, \hat{\mu}_\gamma, \hat{\sigma}_\gamma) = -\frac{n}{2} (\ln 2\pi + 1 + \ln \hat{\sigma}_\gamma^2 + 2\hat{\mu}_\gamma),$$

tedy funkci jedné proměnné γ . Podobně dosazením $\hat{\mu}_\gamma$ a $\hat{\sigma}_\gamma^2$ za μ a σ^2 do rovnice (1) získáme rovnici o jedné neznámé γ . Nalezením lokálního maxima funkce $\lambda(\gamma)$ nebo z upravené rovnice (1) pak získáme odhad $\hat{\gamma}$ parametru γ a ze vztahů (4) a (5) odhady parametrů μ a σ .

Z literatury (např. [3, 4, 5]) vyplývá, že funkce λ má maximálně jeden bod lokálního maxima a pro $\gamma \rightarrow x_{\min}^-$ je $\lambda(\gamma) \rightarrow +\infty$. Je-li $\hat{\gamma}_0$ bodem maxima λ , pak má funkce i lokální minimum $\hat{\gamma}_1$ a platí $\hat{\gamma}_0 < \hat{\gamma}_1 < x_{\min}$.

S nalezením maxima věrohodnostní funkce L mohou být spojeny obtíže. Byly navrženy různé postupy. Harter a Moore [2] hledají maximum opakovaným iterováním rovnic (1)–(3). Neuvádějí však, že jedním z řešení těchto rovnic je i sedlový bod věrohodnostní funkce. Není tedy jasné, zda jejich postup vede vždy k nalezení odhadů LMLE.

Použití Newtonovy-Raphsonovy metody (metoda tečen) pro řešení rovnice $\lambda'(\gamma) = 0$ popisuje např. Calitz [1]. Tato metoda vede k řešení v případě

vhodné startovací hodnoty. Při nevhodné volbě je nalezen bod minima funkce λ , popř. získáme nepřipustnou hodnotu z oblasti $[x_{\min}, +\infty)$ nebo metoda diverguje k $-\infty$. V takovém případě je třeba zvolit jinou startovací hodnotu a postup opakovat.

Další možností je hledat lokální maximum funkce λ na zvoleném intervalu $[a, x_{\min})$, kde většinou $a = 0$, tj. uvažujeme $\gamma \geq 0$. Daný interval potom rozdělíme na (ekvidistantní) body $\gamma_i: a = \gamma_1, \dots, \gamma_k < x_{\min}$ a hledáme lokální maximum funkčních hodnot $\lambda(\gamma_1), \dots, \lambda(\gamma_k)$. Tento jednoduchý postup však nevyhovuje našemu předpokladu $\gamma \in \mathbb{R}$ a může ztroskotat na nevhodné volbě bodů γ_i .

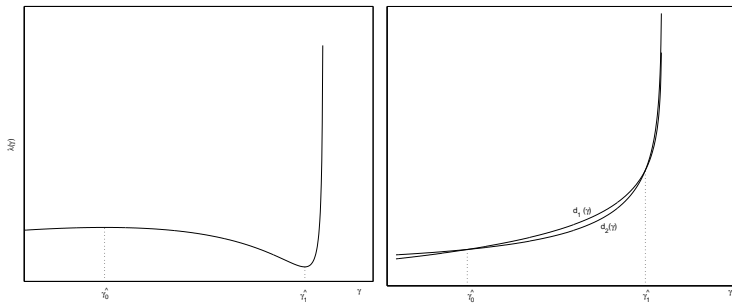
Byly navrženy i další, složitější postupy pro stanovení odhadů LMLE, např. Hirose [4] nebo Komori a Hirose [6] použili transformované logaritmicko-normální rozdělení. Popis těchto metod však není cílem tohoto příspěvku. Tím je popsat novou metodu, která umožňuje při libovolné volbě startovacího bodu $\gamma_0 < x_{\min}$ spolehlivě zmapovat tvar funkce λ ve smyslu nalezení stacionárních bodů, a tedy najít odhady LMLE.

3 Nalezení odhadů LMLE

Označme $f_1(\gamma) = \hat{\sigma}_\gamma^2$, $f_2(\gamma) = \hat{\mu}_\gamma$. Potom pro funkci $\lambda(\gamma)$ (příklad na obr. 1) a její derivaci $\lambda'(\gamma)$ platí

$$\begin{aligned}\lambda(\gamma) &= -\frac{n}{2}[\ln 2\pi + 1 + \ln f_1(\gamma) + 2f_2(\gamma)], \\ \lambda'(\gamma) &= -\frac{n}{2} \left[\frac{f_1'(\gamma)}{f_1(\gamma)} + 2f_2'(\gamma) \right].\end{aligned}$$

Naším cílem je najít lokální maximum funkce λ . Hledáme tedy $\hat{\gamma}$ tak, aby platilo



Obrázek 1: Funkce $\lambda(\gamma)$, $d_1(\gamma)$ $d_2(\gamma)$ se stacionárními body $\hat{\gamma}_0$ (hledané lokální maximum) a $\hat{\gamma}_1$.

$$\lambda'(\hat{\gamma}) = 0 \Leftrightarrow \frac{f_1'(\hat{\gamma})}{f_1(\hat{\gamma})} = -2f_2'(\hat{\gamma}), \quad (6)$$

$$\lambda'(\hat{\gamma} - \epsilon) > 0 \quad \text{a} \quad \lambda'(\hat{\gamma} + \epsilon) < 0$$

pro libovolně malé $\epsilon > 0$. Protože $f_1(\gamma) > 0$ (viz důkaz následujícího tvrzení), obě strany rovnic (6) lze vynásobit $f_1(\hat{\gamma})$, čímž získáme rovnici

$$f_1'(\hat{\gamma}) = 2f_1(\hat{\gamma})[-f_2'(\hat{\gamma})].$$

Pro větší přehlednost zavedeme funkce $d_1(\gamma) = f_1'(\gamma)$ a $d_2(\gamma) = 2f_1(\gamma)[-f_2'(\gamma)]$, budeme tedy hledat řešení rovnice $d_1(\hat{\gamma}) = d_2(\hat{\gamma})$.

Tvrzení: Pro libovolně x_1, \dots, x_n , pro které $\exists i, j : x_i \neq x_j$, jsou funkce $d_1(\gamma)$, $d_2(\gamma)$ ostře rostoucí a konvexní pro $\gamma \in (-\infty, x_{\min})$.

Důkaz. Definičním oborem všech funkcí proměnné γ v tomto důkazu je $D_\gamma = \{\gamma \in \mathbb{R} : \gamma < x_{\min}\}$. Bez újmy na obecnosti budeme předpokládat uspořádanost hodnot $x_1 < x_2 < \dots < x_n$.

Nejprve prozkoumáme průběh funkce d_1 . Funkce $f_1(\gamma) = \hat{\sigma}_\gamma^2$ je rozptylem hodnot $\ln(x_i - \gamma)$, $i = 1, \dots, n$, jedná se tedy o kladnou funkci. Dále je

$$\begin{aligned} f_1(\gamma) &= \frac{1}{n} \sum_{i=1}^n [\ln(x_i - \gamma) - \hat{\mu}_\gamma]^2 = \frac{1}{n} \sum_{i=1}^n [\ln(x_i - \gamma) - f_2(\gamma)]^2 \\ &= \frac{1}{n} \sum_{i=1}^n [\ln(x_i - \gamma)]^2 - f_2^2(\gamma), \\ f_1'(\gamma) &= 2 \left[\frac{1}{n} \sum_{i=1}^n \ln(x_i - \gamma) \frac{-1}{x_i - \gamma} - f_2(\gamma) f_2'(\gamma) \right], \\ f_1''(\gamma) &= 2 \left[\frac{1}{n} \sum_{i=1}^n \ln(x_i - \gamma) \frac{-1}{(x_i - \gamma)^2} - f_2(\gamma) f_2''(\gamma) \right] + \\ &\quad + 2 \left[\frac{1}{n} \sum_{i=1}^n \frac{1}{(x_i - \gamma)^2} - [f_2'(\gamma)]^2 \right]. \end{aligned}$$

Funkce $f_2(\gamma)$ je průměrem hodnot $\ln(x_i - \gamma)$, funkce $f_2'(\gamma)$ je průměrem hodnot $\frac{-1}{x_i - \gamma}$, atd. Hranatá závorka ve funkci f_1' tedy představuje výpočetní tvar výběrové kovariance vektorů $[\ln(x_i - \gamma), \frac{-1}{x_i - \gamma}]$, $i = 1, \dots, n$, ve funkci f_1'' najdeme v hranatých závorkách výběrovou kovarianci a rozptyl (v dalším textu budeme pod pojmem kovariance vždy mít výběrovou kovarianci).

Označme $g(x, \gamma) = \ln(x - \gamma)$ a $c_{kl}(\gamma)$ kovarianci vektorů $\left[\frac{\partial^k g(x_i, \gamma)}{\partial \gamma^k}, \frac{\partial^l g(x_i, \gamma)}{\partial \gamma^l} \right]$, $i = 1, \dots, n$, $k, l \in \mathbb{N}_0$. Indexy k, l tedy představují stupeň derivace funkce g podle γ . Snadno lze ukázat, že derivace $\frac{dc_{k,l}(\gamma)}{d\gamma} = c_{k+1,l}(\gamma) + c_{k,l+1}(\gamma)$. Ze symetrie kovariance plyne $c_{k,l}(\gamma) = c_{l,k}(\gamma)$. Je tedy

$$\begin{aligned}
f_1(\gamma) &= c_{0,0}(\gamma), \\
f_1'(\gamma) &= 2c_{0,1}(\gamma), \\
f_1''(\gamma) &= 2[c_{0,2}(\gamma) + c_{1,1}(\gamma)], \\
f_1'''(\gamma) &= 2[c_{0,3}(\gamma) + c_{1,2}(\gamma) + c_{1,2}(\gamma) + c_{2,1}(\gamma)] = 2[c_{0,3}(\gamma) + 3c_{1,2}(\gamma)], \text{ atd.}
\end{aligned}$$

Dále ukážeme, že kovariance $c_{k,l}(\gamma)$ jsou kladné pro všechna $k, l \in \mathbb{N}_0$ a pro libovolné $\gamma \in D_\gamma$. Označme $\alpha_i = \frac{\partial^k g(x_i, \gamma)}{\partial \gamma^k}$ a $\beta_i = \frac{\partial^l g(x_i, \gamma)}{\partial \gamma^l}$, $i = 1, \dots, n$ pro daná k, l . Protože funkce $\frac{\partial^k g(x, \gamma)}{\partial \gamma^k}$ je ostře rostoucí v proměnné x pro všechna $k \in \mathbb{N}_0$, je $\alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_n$, $\beta_1 \leq \beta_2 \leq \dots \leq \beta_n$. Podle Čebyševovy nerovnosti (viz např. Kufner [7], str. 31) platí

$$n(\alpha_1\beta_1 + \alpha_2\beta_2 + \dots + \alpha_n\beta_n) \geq (\alpha_1 + \alpha_2 + \dots + \alpha_n)(\beta_1 + \beta_2 + \dots + \beta_n), \quad (7)$$

přičemž rovnost nastává pro $\alpha_1 = \alpha_2 = \dots = \alpha_n$ nebo $\beta_1 = \beta_2 = \dots = \beta_n$. Protože stále předpokládáme, že existuje i, j , pro které $x_i \neq x_j$, a tedy $\alpha_i \neq \alpha_j$ a $\beta_i \neq \beta_j$, je v (7) vždy ostrá nerovnost.

Pro kovarianci $c_{k,l}$ potom platí

$$c_{k,l} = \frac{1}{n} \sum_{i=1}^n \alpha_i \beta_i - \left[\frac{1}{n} \sum_{i=1}^n \alpha_i \right] \left[\frac{1}{n} \sum_{i=1}^n \beta_i \right] > 0$$

a pro funkci d_1 a její derivace je

$$\begin{aligned}
d_1(\gamma) &= f_1'(\gamma) = 2c_{0,1} > 0, \\
d_1'(\gamma) &= f_1''(\gamma) = 2[c_{0,2}(\gamma) + c_{1,1}(\gamma)] > 0, \\
d_1''(\gamma) &= f_1'''(\gamma) = 2[c_{0,3}(\gamma) + 3c_{1,2}(\gamma)] > 0.
\end{aligned}$$

Funkce d_1 je tedy rostoucí a konvexní na daném definičním oboru.

Nyní dokážeme tvrzení pro funkci d_2 . Označme $f_3(\gamma) = [-f_2'(\gamma)]$, je

$$\begin{aligned}
f_3(\gamma) &= -f_2'(\gamma) = \frac{1}{n} \sum_{i=1}^n \frac{1}{x_i - \gamma} > 0, \\
f_3'(\gamma) &= -f_2''(\gamma) = \frac{1}{n} \sum_{i=1}^n \frac{1}{(x_i - \gamma)^2} > 0, \\
f_3''(\gamma) &= -f_2'''(\gamma) = \frac{1}{n} \sum_{i=1}^n \frac{1}{(x_i - \gamma)^3} > 0.
\end{aligned}$$

Pro funkci d_2 a její derivace pak platí

$$\begin{aligned}
d_2(\gamma) &= 2f_1(\gamma)[-f_2'(\gamma)] = 2f_1(\gamma)f_3(\gamma) > 0, \\
d_2'(\gamma) &= 2[f_1'(\gamma)f_3(\gamma) + f_1(\gamma)f_3'(\gamma)] > 0, \\
d_2''(\gamma) &= 2[f_1''(\gamma)f_3(\gamma) + 2f_1'(\gamma)f_3'(\gamma) + f_1(\gamma)f_3''(\gamma)] > 0.
\end{aligned}$$

Z posledních dvou nerovností vyplývá platnost tvrzení pro funkci d_2 . \square

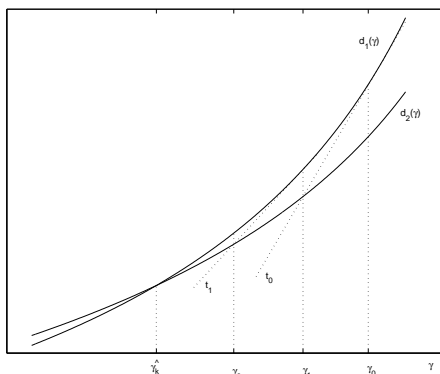
Postup pro získání odhadů LMLE novou metodou ukážeme na vzorku 0,73; 0,75; 0,76; 0,80; 0,96; 1,07; 1,11; 1,44; 1,73; 1,81. Tento vzorek rozsahu $n = 10$ byl získán generováním náhodné veličiny $X \sim LN(0; 0; 1)$. Na obrázku 1 jsou znázorněny funkce λ , d_1 a d_2 tohoto vzorku s průsečíky $\hat{\gamma}_0$ a $\hat{\gamma}_1$. Hledaným lokálním maximem funkce λ je $\hat{\gamma} = \hat{\gamma}_0$, $\hat{\gamma}_1$ je jejím minimem.

3.1 Postup nalezení stacionárního bodu

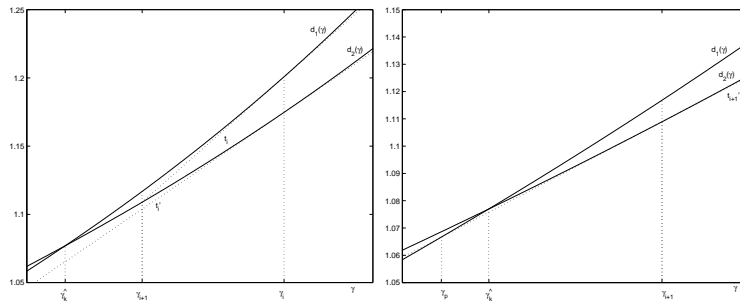
Mějme výchozí bod $\gamma_0 < x_{\min}$ a hledejme $\hat{\gamma}_k$, kde $k = 0$ nebo $k = 1$. V bodě γ_0 sestrojíme tečnu t_0 k funkci, která má v γ_0 vyšší funkční hodnotu, v našem příkladu (viz obr. 2) tedy k funkci d_1 . Tečna t_0 má předpis $\tau_0(\gamma) = d_1(\gamma_0) + d_1'(\gamma_0)(\gamma - \gamma_0)$, obecně je $t_i : \tau_i(\gamma) = d_1(\gamma_i) + d_1'(\gamma_i)(\gamma - \gamma_i)$. Najdeme bod $\gamma_1 < \gamma_0$, pro který platí $\tau_0(\gamma_1) = d_2(\gamma_1)$, což je průsečík tečny t_0 a funkce d_2 vlevo od bodu γ_0 . Postup opakujeme, z bodu $[\gamma_i, d_1(\gamma_i)]$ sestrojíme tečnu t_i a najdeme další bod γ_{i+1} . Tímto postupem dostaneme posloupnost $\{\gamma_0, \gamma_1, \dots\}$, pro kterou platí $\gamma_0 > \gamma_1 > \gamma_2 > \dots > \hat{\gamma}_k$, navíc pro libovolné $\epsilon > 0$ existuje $i_0 > 0$ takové, že pro všechna $i \geq i_0$ je $|\gamma_i - \hat{\gamma}_k| \leq \epsilon$. Toto tvrzení platí pouze v případě, že se vlevo od γ_0 nachází stacionární bod; jeho důkaz je jednoduchý a nebudeme ho proto uvádět. Posloupnost $\{\gamma_0, \gamma_1, \dots\}$ v našem příkladu konverguje ke stacionárnímu bodu $\hat{\gamma}_0$.

Chtějme nyní stanovit interval I_k , ve kterém leží pouze hledané $\hat{\gamma}_k$, tedy $\hat{\gamma}_k \in I_k$ a $\hat{\gamma}_{k'} \notin I_k$, $k' = 1 - k$. Libovolné γ_i , $i > 0$, získané popsáním postupem lze vzít za pravou mez intervalu I_k . Pro nalezení levé meze v každém kroku předchozího postupu stanovíme navíc tečnu „dolní“ funkce, $t'_i : \tau'_i(\gamma) = d_2(\gamma_i) + d_2'(\gamma_i)(\gamma - \gamma_i)$ (viz obr. 3). Jestliže existuje bod $\gamma_p < \gamma_i$, pro který je $\tau'_i(\gamma_p) = d_1(\gamma_p)$, je dolní mez nalezena. Interval $I_k = (\gamma_p, \gamma_{i_0})$, kde $i_0 > 0$ lze zvolit libovolně, obsahuje hledané $\hat{\gamma}_k$. Hodnotu $\hat{\gamma}_k$ lze potom stanovit např. metodou půlení intervalu pro funkci $\lambda'(\gamma)$.

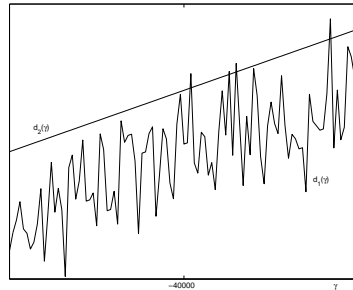
Při hledání intervalu I_k omezíme počet provedených iterací (tj. počet γ_i), při dosažení tohoto počtu hledání stacionárního bodu ukončíme. V 99% pro-



Obrázek 2: Z výchozího bodu γ_0 najdeme γ_1 jako průsečík tečny t_0 a funkce d_2 vlevo od γ_0 , poté pokračujeme do γ_2 atd.



Obrázek 3: Tečna t'_i nemá průsečík s funkcí d_1 , tečna t'_{i+1} ano.



Obrázek 4: Zobrazení funkce d_1 není rostoucí vlivem zaokrouhlovacích chyb.

vedených simulací byl interval I_k nalezen po nejvýše 30 iteracích, počet iterací však patrně ovlivňuje volba startovací hodnoty a hodnota parametru σ (při nižších hodnotách σ bylo potřeba více iterací). Námí používaná startovací hodnota byla $\gamma_0 = x_{\min} - R/(n-1)$, kde R je rozpětí souboru x_1, \dots, x_n .

Pokud jsme našli průsečík funkcí d_1 a d_2 , je vhodné zkontrolovat, zda k tomu nedošlo důsledkem zaokrouhlovacích chyb, které vznikají při použití výpočetní techniky. Na obr. 4 uvádíme funkce d_1 a d_2 nakreslené programem Matlab 6 pro velmi malá γ . Funkce d_1 není zobrazena jako rostoucí, což je způsobeno právě zaokrouhlovacími chybami. Považovat průsečíky z obr. 4 za stacionární body by zřejmě bylo chybou. V případě, že v blízkém okolí nalezeného průsečíku netvoří funkční hodnoty d_1 a d_2 rostoucí konvexní posloupnosti, ukončíme postup s tím, že řešení rovnice $d_1(\hat{\gamma}) = d_2(\hat{\gamma})$ nebylo nalezeno.

3.2 Nalezení lokálního maxima

Ukázali jsme, jak najít stacionární bod vlevo od zvolené výchozí hodnoty γ_0 . Tento postup použijeme, pokud $\lambda'(\gamma_0) < 0$ a naším předpokladem tedy je $\gamma_0 \in (\hat{\gamma}_0, \hat{\gamma}_1)$. Nalezené $\hat{\gamma} = \hat{\gamma}_0$ je pak hledaným odhadem parametru γ .

Pokud není $\hat{\gamma}$ nalezeno, potom lokální maximum funkce λ (a též L) neexistuje, popř. nebylo dosaženo kvůli vyčerpání maximálního počtu iterací. Není též vyloučeno, že se nachází v oblasti, kterou nejsme schopni prozkoumat kvůli relativně velkým zaokrouhlovacím chybám.

Jestliže je $\lambda'(\gamma_0) > 0$, uvažujeme tři možnosti: $\gamma_0 \in (-\infty, \hat{\gamma}_0)$ nebo $\gamma_0 \in (\hat{\gamma}_1, x_{\min})$, popř. funkce λ nemá ani jeden stacionární bod. Je třeba zvolit, který směr prohlédneme jako první. Na základě našich simulací doporučujeme nejprve prohledat oblast směrem vpravo od γ_0 (možnost $\gamma_0 \in (-\infty, \hat{\gamma}_0)$). Použijeme **postup nalezení stacionárního bodu** ze strany 466, nyní je však třeba hledat posloupnost $\gamma_0 < \gamma_1 < \gamma_2 < \dots < \gamma_p$.

Při našich simulacích (s volbou $\gamma_0 = x_{\min} - R/(n - 1)$) nenastala pro $\lambda'(\gamma_0) > 0$ nikdy možnost $\gamma_0 \in (\hat{\gamma}_1, x_{\min})$, přesto ji nelze vyloučit. Jestliže tedy není maximum nalezeno vpravo od γ_0 , hledáme **postupem nalezení stacionárního bodu** interval $I_1 = (\gamma_p, \gamma_{i_0})$, který obsahuje sedlový bod $\hat{\gamma}_1$, vlevo od γ_0 . Protože v případě nalezení tohoto intervalu platí nerovnosti $\hat{\gamma}_0 < \gamma_p < \hat{\gamma}_1$, je $\lambda'(\gamma_p) < 0$ a γ_p použijeme jako novou startovací hodnotu $\gamma_0 := \gamma_p$. Poté najdeme $\hat{\gamma} = \hat{\gamma}_0$ **postupem nalezení stacionárního bodu**.

Literatura

- [1] Calitz F. (1973). *Maximum likelihood estimation of the parameters of the three-parameter lognormal distribution - a reconsideration*. Austral. J. Statist. **15**, 185–190.
- [2] Harter H.L., Moore A.L. (1966). *Local-maximum-likelihood estimation of the parameters of three-parameter lognormal populations from complete and censored samples*. J. Amer. Statist. Assoc. **61**, 842–851.
- [3] Hill B.M. (1963). *The three-parameter lognormal distribution and Bayesian analysis of a point-source epidemic*. J. Amer. Statist. Assoc. **58**, 72–84.
- [4] Hirose H. (1997). *Maximum likelihood parameter estimation in the three-parameter log-normal distribution using the continuation method*. Computational Statistics & Data Analysis **24**, 139–152.
- [5] Kane V.E. (1982). *Standard and goodness-of-fit parameter estimation methods for the three-parameter lognormal distribution*. Commun. Statist. — Theor. Meth. **11**, 1935–1957.
- [6] Komori Y., Hirose H. (2004). *Yeasy estimation by a new parametrization for the three-parameter lognormal distribution*. Journal of Statistical Computation & Simulation **74**, 1, 63–74.
- [7] Kufner A. (1975). *Nerovnosti a odhady*. ÚV Matematické olympiády v nakladatelství Mladá fronta, Praha.

Poděkování: Práce byla podpořena výzkumným záměrem MSM 4977751301.
Adresa: V. Václavík, Katedra matematiky, FAV ZČU, Univerzitní 22, 306 14 Plzeň

E-mail: vladimir.vaclavik@email.cz

ON EFFICIENCIES OF DECISIONS ABOUT STATISTICAL MODELS BASED ON f -DIVERGENCES OF EMPIRICAL DISTRIBUTIONS

Igor Vajda

Keywords: Statistical models, goodness of fit tests, divergence statistics, limit theorems, Pitman efficiency, Bahadur efficiency.

Abstract: Statistical model of stochastic reality is a hypothetical probability distribution on a given measurable space. True distributions of measurements (signals, images) are generally unknown. Usually are available digitalized (appropriately quantized or classified) descriptions of measurements which can be represented by empirical probability distributions. We consider statistical decisions about the models consisting in testing hypotheses about their distributions. The tests are based on divergence statistics (f -divergences of hypothetical and empirical distributions). The first part of this paper deals with limit laws under hypotheses and local alternatives needed for the testing. The second part of the paper deals with efficiencies of the tests. Both Pitman and Bahadur approaches to the relative efficiency of statistics are studied. This leads to new results concerning optimality of the likelihood ratio test.

Abstrakt: V práci jsou uvažovány testy hypotézy o typu rozdělení založené na f -divergenci hypotetické a empirické distribuce. První část článku pojednává o limitním rozdělení za hypotézy a lokální alternativy. Je uvažován jak Pitmanův tak Bahadurův přístup k relativní eficienci. Získané výsledky vedou k novým výsledkům týkajícím se optimality testů založených na poměru věrohodnosti.

1 Statistical model

We consider a statistical model (Ω, \mathcal{S}, P) with known measurable observation space (Ω, \mathcal{S}) and unknown probability distribution P producing i.i.d. realizations Y_1, \dots, Y_n . Available are only the digitalized (appropriately quantized) data

$$X_j = \sum_{i=1}^n 1_{\{A_j\}}(Y_i), \quad 1 \leq j \leq k$$

for a given partition

$$\mathcal{A} = \{A_1, \dots, A_k\} \subset \mathcal{S}$$

of Ω . We admit that the partition \mathcal{A} , the partition sets A_j and the partition sizes k depend on the sample size n , i.e.

$$\mathcal{A} = \mathcal{A}_n, \quad A_j = A_{j,n}, \quad k = k_n. \quad (1)$$

In this paper we study testing the hypothesis \mathbf{H} that the stochastic outputs Y_i of the model are generated by a given distribution P^0 against the alternative \mathbf{A} represented by the true distribution of these outputs. The testing is assumed to be carried out by means of the available data

$$\mathbf{X} = (X_1, \dots, X_k). \quad (2)$$

This means that, in fact, we study the problem of testing

$$\mathbf{H} \sim p = p^0 \quad \text{against} \quad \mathbf{A} \sim p \quad (\text{true})$$

where

$$p^0 = (p_j^0 \equiv P^0(A_j) : 1 \leq j \leq k) \quad (3)$$

is a *discrete hypothetical distribution* and

$$p = (p_j \equiv P(A_j) : 1 \leq j \leq k) \quad (4)$$

a *discrete true distribution*, and that the testing is carried out by means of the data (2) uniquely represented by the *discrete empirical distribution*

$$\hat{p} = (\hat{p}_1 \equiv X_1/n, \dots, \hat{p}_k \equiv X_k/n). \quad (5)$$

In view of (1) this means that

$$X_j = X_{j,n}, \quad p_j^0 = p_{j,n}^0, \quad p_j = p_{j,n} \quad \text{and} \quad \hat{p}_j = \hat{p}_{j,n} \quad (6)$$

in (2)-(5).

We study various methods of the testing and preferences between them in the situation where the sample size n increases above all bounds. In this context we respect throughout this paper the following conventions and assumptions.

Conventions: (i) The subscripts n considered in (1) and (6) are suppressed and (ii) unless otherwise explicitly stated, all convergences and asymptotic expressions like \lim , \longrightarrow , \xrightarrow{p} , \xrightarrow{d} , $o(1)$ or $O(1)$ are considered for $n \rightarrow \infty$.

Assumptions: It holds $k \rightarrow \infty$, and for some $\beta \geq 1$ also

$$k^{\beta+1}/n = o(1) \quad \text{and} \quad \min_n k^\beta p_{\min}^0 \geq \text{const} > 0 \quad (7)$$

where $p_{\min}^0 = \min\{p_j^0 : 1 \leq j \leq k\}$.

2 Divergence statistics

Let us denote by \mathcal{F} the class of all functions $f(t)$ twice differentiable with $f''(t) > 0$ in the domain $t \in (0, \infty)$ which are Lipschitz around $t = 1$ and standardized in the sense $f(1) = 0$. The assumption of positive second derivative means that $f(t)$ is strictly convex on $(0, \infty)$. By $f(0) \in (-\infty, \infty]$ we denote the extension for $t \downarrow 0$. This paper studies the following class of statistics.

Definition 1. The *divergence statistics* are defined by the formula

$$\mathcal{D}_{f,n} = \frac{2n D_f(\hat{p}, p^0)}{f''(1)}, \quad f \in \mathcal{F} \tag{8}$$

where

$$D_f(\hat{p}, p^0) = \sum_{j=1}^k p_j^0 f\left(\frac{\hat{p}_j}{p_j^0}\right) \tag{9}$$

is the f -divergence of distributions \hat{p}, p^0 .

Notice that by (7) it holds $p_j^0 > 0$ in (9). For the properties of the f -divergence (9) see e.g. [7]. Next follow some well known examples of the divergence statistics (8).

Example 1 (*classical Pearson statistic*). The quadratic function $f(t) = (t - 1)^2$ leads to the Pearson divergence $\chi^2(\hat{p}, p^0)$ and the classical Pearson statistic

$$\chi_n^2 = n\chi^2(\hat{p}, p^0) = n \sum_{j=1}^k \frac{(\hat{p}_j - p_j^0)^2}{p_j^0} = \sum_{j=1}^k \frac{(X_j - np_j^0)^2}{np_j^0}.$$

Example 2 (*likelihood ratio statistic*). The logarithmic function $f(t) = t \ln t$ leads to the information divergence $I(\hat{p}, p^0)$ and the likelihood ratio statistic

$$\mathcal{I}_n = 2nI(\hat{p}, p^0) = 2n \sum_{j=1}^k \hat{p}_j \ln \frac{\hat{p}_j}{p_j^0} = 2 \sum_{j=1}^k X_j \ln \frac{X_j}{np_j^0}. \tag{10}$$

Example 3 (*power divergence statistics*). The class of power functions

$$f_\alpha(t) = \frac{t^\alpha - \alpha(t - 1) - 1}{\alpha(\alpha - 1)} \quad \text{where } \alpha \in \mathbb{R}, \quad \alpha(\alpha - 1) \neq 0$$

with the limits

$$f_0(t) = -\ln t + t - 1 \quad \text{and} \quad f_1(t) = t \ln t - t + 1$$

define power divergences $D_\alpha(\hat{p}, p^0) \equiv D_{f_\alpha}(\hat{p}, p^0)$ and the corresponding power divergence statistics $\mathcal{D}_{\alpha,n} \equiv \mathcal{D}_{f_{\alpha,n}}$ for various power parameters $\alpha \in \mathbb{R}$. It is easy to see that $\chi^2(\hat{p}, p^0) \equiv 2D_2(\hat{p}, p^0)$ and $I(\hat{p}, p^0) \equiv D_1(\hat{p}, p^0)$ are special cases of the power divergences and $\chi_n^2 \equiv \mathcal{D}_{2,n}$ and $\mathcal{I}_n \equiv \mathcal{D}_{1,n}$ are special cases of the power divergence statistics.

3 Limit laws

Let throughout this section the conditions and assumptions introduced in Section 1 hold. Then for all $f \in \mathcal{F}$

$$\frac{\mathcal{D}_{f,n} - k}{\sqrt{2k}} \xrightarrow{d} N(0, 1) \quad \text{under } \mathbf{H} \quad (11)$$

according to [3]. This extends the classical limit law

$$\frac{\chi_n^2 - k}{\sqrt{2k}} \xrightarrow{d} N(0, 1) \quad \text{under } \mathbf{H} \quad (12)$$

of [8] valid with the present assumption $k^{1+\beta}/n = o(1)$ replaced by the weaker $k/n = o(1)$. It is natural to ask whether a universal asymptotically normal law similar to (11) remains valid also when the hypothetical equality $\mathbf{H} : p = p^0$ is replaced by the alternative $\mathbf{A} \sim p$ local in the sense that p is close to p^0 . The answer is yes provided that p tends sufficiently fast to p^0 in terms of their mutual Pearson divergence $\chi^2(\hat{p}, p^0)$. Before going into details note that a partial variant of this answer for the simple but important uniform hypotheses

$$p^0 = (p_j^0 \equiv 1/k : 1 \leq j \leq k). \quad (13)$$

was obtained previously in [10]. Here the hypotheses are restricted only by the second condition in (7) which is obviously weaker than (13).

Definition 2. The alternative $\mathbf{A} \sim p$ is said to be *local* or *weakly local* if for some $\Delta \geq 0$

$$\frac{n\chi^2(p, p^0)}{\sqrt{k}} \longrightarrow \Delta \quad \text{or} \quad \frac{n\chi^2(p, p^0)}{\sqrt{k}} = O(1)$$

respectively.

Example 4. The classical statistical local alternative $\mathbf{A} \sim p$ is of the form

$$p = \left(1 - \frac{1}{\sqrt{n}}\right)p^0 + \frac{1}{\sqrt{n}}q$$

for some $q = (q_j \equiv Q(A_j) : 1 \leq j \leq k)$ (cf. (4)). Here $\chi^2(p, p^0) = \chi^2(q, p^0)/n$ so that this alternative is local in the present sense if $\chi^2(q, p^0)/\sqrt{k}$ is convergent and weakly local if $\chi^2(q, p^0)/\sqrt{k}$ is bounded.

Under the assumptions considered in this paper (11) can be extended into the following *Universal Asymptotic Normality Theorem*.

Theorem 1. All f -divergence statistics $D_{f,n}$ satisfy the limit law

$$\frac{\mathcal{D}_{f,n} - k - \sqrt{k}\Delta}{\sqrt{2k}} \xrightarrow{d} N(0, 1) \quad \text{under local } \mathbf{A}. \quad (14)$$

Proof of this theorem is based on the following *Extension Lemma*.

Lemma 1. If for some $\mu_n \in \mathbb{R}$ and $\sigma_n > 0$

$$\frac{\chi_n^2 - \mu_n}{\sigma_n} \xrightarrow{d} N(0, 1) \quad \text{under weakly local } \mathbf{A}$$

then for all divergence statistics $\mathcal{D}_{f,n}$

$$\frac{\mathcal{D}_{f,n} - \mu_n}{\sigma_n} \xrightarrow{d} N(0, 1) \quad \text{under weakly local } \mathbf{A}.$$

Proof. Let \mathbf{A} be weakly local. It suffices to prove for all sufficiently small $\varepsilon > 0$

$$\mathbb{P} \left(\left| \frac{\mathcal{D}_{f,n} - \chi_n^2}{\sqrt{k}} \right| > \varepsilon \mid \mathbf{A} \right) = o(1). \quad (15)$$

By the inequalities in [3, p. 65], for all sufficiently small $\varepsilon > 0$ there exist constants $c(\varepsilon) > 0$ such that for all n

$$\mathbb{P} \left(\left| \frac{\mathcal{D}_{f,n} - \chi_n^2}{\sqrt{k}} \right| > \varepsilon \mid \mathbf{A} \right) \leq c(\varepsilon) \left(\frac{nA_n}{\sqrt{k}} + B_n \right)$$

where

$$A_n = \sum_{j=1}^k \frac{\mathbb{E} |\hat{p}_j - p_j^0|^3}{(p_j)^2}$$

and

$$B_n = \sum_{j=1}^k \frac{\mathbb{E} (\hat{p}_j - p_j^0)^2}{(p_j)^2}.$$

Using the fact that under \mathbf{A}

$$n\hat{p} \equiv \mathbf{X} \sim M_n(p, k) \quad (\text{see (2)})$$

and employing the assumptions, one obtains from here $A_n = o(\sqrt{k}/n)$, $B_n = o(1)$ and thus the desired asymptotics (15). The detailed steps are similar to those undertaken in the more special situation in [10].

Proof of Theorem 1. After some effort it is possible to verify that under the assumptions of this paper Theorem 5.1 of [8] implies

$$\frac{\chi_n^2 - k - \sqrt{k}\Delta}{\sqrt{2k}} \xrightarrow{d} N(0, 1) \quad \text{under local } \mathbf{A}.$$

The desired result follows by applying this in the Lemma 1.

Example 5 (*likelihood ratio statistic*). By our UAN theorem

$$\frac{2\mathcal{I}_n - k - \sqrt{k}\Delta}{\sqrt{2k}} \xrightarrow{d} N(0, 1).$$

This particular limit law was proved directly in Theorem 5.2 of [8]) under weaker but

- (i) less intuitive and much more complicated assumptions
- (ii) and incomparably more complicated proof.

Our theorem is not only simpler than the mentioned Theorem 5.2, but also universal, e.g. applicable to all statistics $\mathcal{D}_{\alpha,n}$. Among the well known examples different from $\mathcal{D}_{1,n} = \mathcal{I}_n$ and $\mathcal{D}_{2,n} = \chi_n^2$ one can mention the *Freeman-Tukey statistic*

$$\mathcal{D}_{1/2,n} = nH^2(\hat{p}, p^0) = 4n \sum_{j=1}^k \left(\sqrt{\hat{p}_j} - \sqrt{p_j^0} \right)^2$$

or the *Neyman statistic* $\mathcal{D}_{-1,n}$ and the *reversed likelihood ratio statistic* $\mathcal{D}_{0,n}$.

4 Asymptotic efficiencies

In this section we study the hypotheses \mathbf{H} and alternatives \mathbf{A} introduced in Section 1 and the tests of these hypotheses based on the divergence statistics $\mathcal{D}_{f,n}$ introduced in Section 2 for various functions $f \in \mathcal{F}$. As usually, these tests reject \mathbf{H} when the corresponding statistics $\mathcal{D}_{f,n}$ exceed certain critical values $c_{f,n}$. We consider an arbitrary fixed pair f_1, f_2 of mutually different functions from \mathcal{F} , and for $f \in \{f_1, f_2\}$ we compare asymptotic efficiencies of the test statistics $\mathcal{D}_{f,n}$. These efficiencies refer to the powers $\pi_{f,n}(s) = \mathbb{P}(\mathcal{D}_{f,n} < c_{f,n} | \mathbf{A})$ of the mentioned tests achieved under the critical values $c_{f,n}$ satisfying the asymptotic size condition $s = \lim \mathbb{P}(\mathcal{D}_{f,n} > c_{f,n} | \mathbf{H})$ for a given $s > 0$, and to the sizes $s_{f,n}(\pi) = \mathbb{P}(\mathcal{D}_{f,n} > \tilde{c}_{f,n} | \mathbf{H})$ of these tests under the critical values $\tilde{c}_{f,n}$ satisfying the asymptotic power condition $\pi = \lim \mathbb{P}(\mathcal{D}_{f,n} < \tilde{c}_{f,n} | \mathbf{A})$ for a given $\pi > 0$. Similarly as before, we respect in this section the conditions and assumptions introduced in Section 1. Next follow two classical approaches to the definition of the asymptotic efficiency of the statistic $\mathcal{D}_{f_1,n}$ relative $\mathcal{D}_{f_2,n}$ depending on the parameters s and π (see [9]).

Definition 3. The Pitman asymptotic relative efficiency $PE_s(\mathcal{D}_{f_1,n}, \mathcal{D}_{f_2,n})$ is the limit (if it exists) of the ratio $\pi_{f_1,n}(s)/\pi_{f_2,n}(s)$ of powers of the corresponding divergence tests of equal asymptotic size s . The Bahadur asymptotic relative efficiency $BE_\pi(\mathcal{D}_{f_1,n}, \mathcal{D}_{f_2,n})$ is the limit (if it exists) of the ratio $s_{f_1,n}(\pi)/s_{f_2,n}(\pi)$ of the sizes of the corresponding divergence tests of equal asymptotic power π .

Theorem 2. If the alternative \mathbf{A} is local in the sense of Definition 2 then for all f_1, f_2 and s, π under consideration $PE_s(\mathcal{D}_{f_1,n}, \mathcal{D}_{f_2,n}) = BE_\pi(\mathcal{D}_{f_1,n}, \mathcal{D}_{f_2,n}) = 1$.

Proof. Let Φ be the distribution function of the normal random variable $N(0, 1)$, Φ^{-1} the corresponding quantile function and put for every and $\Delta > 0$

$$c_{f,n}(s) = k + \sqrt{2k}\Phi^{-1}(1 - s), \quad \tilde{c}_{f,n}(\pi) = k + \sqrt{k}\Delta + \sqrt{2k}\Phi^{-1}(\pi)$$

By the limit laws (11) and (14), the critical values $c_{f,n} = c_{f,n}(s)$ and $\tilde{c}_{f,n} = \tilde{c}_{f,n}(\pi)$ satisfy for all $0 < s, \pi < 1$ and all $f \in \mathcal{F}$ the above considered asymptotic size and power conditions and also the limit laws

$$s_{f,n}(\pi) = P\left(\frac{\mathcal{D}_{f,n} - k}{\sqrt{2k}} > \frac{\Delta}{\sqrt{2}} + \Phi^{-1}(\pi) \mid \mathbf{H}\right) \longrightarrow \Phi\left(\frac{\Delta}{\sqrt{2}} + \Phi^{-1}(\pi)\right)$$

and

$$P\left(\frac{\mathcal{D}_{f,n} - k - \sqrt{k}\Delta}{\sqrt{2k}} < \Phi^{-1}(1 - s) - \frac{\Delta}{\sqrt{2}} \mid \mathbf{A}\right) \longrightarrow \Phi\left(\Phi^{-1}(\pi) - \frac{\Delta}{\sqrt{2}}\right)$$

where the last probability represents $\pi_{f,n}(s)$. The rest is clear from here.

Obvious reason why Definition 3 fails is the too small (asymptotically vanishing) deviation $\chi^2(p, p^0)$ of the local alternative $\mathbf{A} \sim p$ from $\mathbf{H} \sim p^0$ required in Definition 2, leading to the same asymptotically vanishing deviation $D_f(p, p^0)$ in terms of all f -divergences as it is visible from (14). Thus, following [1]), in the rest of this section we consider the alternatives $\mathbf{A} \sim p$ satisfying the *large deviation condition*

$$D_f(p, p^0) \longrightarrow \Delta_f > 0 \quad \text{for } f \in \{f_1, f_2\} \subset \mathcal{F}. \tag{16}$$

In accordance with [9] we suppose that the statistics $\mathcal{D}_{f,n}$ are for every $f \in \{f_1, f_2\}$ consistent in the sense

$$\frac{\mathcal{D}_{f,n}}{n} \xrightarrow{p} \begin{cases} 0 & \text{under } \mathbf{H} \\ \Delta_f & \text{under } \mathbf{A} \end{cases}. \tag{17}$$

This means that the asymptotic power condition $\pi = \lim \Pr(\mathcal{D}_{f,n} < \tilde{c}_{f,n} \mid \mathbf{A})$ holds for the critical values of the form $\tilde{c}_{f,n} = n\Delta_f + o(n)$ so that the test

sizes $s_{f,n}(\pi)$ considered in the definition of the Bahadur efficiency above are of the form $s_{f,n}(\pi) = \Pr(\mathcal{D}_{f,n} > n\Delta_f + o(n) | \mathbf{H}) \approx \Pr(\mathcal{D}_{f,n} > n\Delta_f | \mathbf{H})$. Thus the new concept of relative efficiency in the next definition follows the above stated Bahadur approach, just the "small deviation condition" $D_f(p, p^0) \rightarrow 0$ imposed on the alternative is replaced by the large deviation condition (16).

Definition 3. Let for every $f \in \{f_1, f_2\}$ the test statistic $\mathcal{D}_{f,n}$ be consistent in the sense of (17) and let there exist a sequence $a_{f,n} \rightarrow \infty$ and a continuous function $g_f : (0, \infty) \rightarrow \mathbb{R}$ such that for all $\Delta > 0$

$$P(\mathcal{D}_{f,n} > n\Delta | \mathbf{H}) = \exp\{-a_{f,n}[g_f(\Delta) + o(1)]\} \approx \exp\{-a_{f,n}g_f(\Delta)\}. \tag{18}$$

Then we say that the limit

$$\mathcal{BE}(\mathcal{D}_{f_1,n}, \mathcal{D}_{f_2,n}) = \lim \frac{a_{f_1,n} g_{f_1}(\Delta_{f_1})}{a_{f_2,n} g_{f_2}(\Delta_{f_2})} \tag{19}$$

(if it exists) is the *Bahadur asymptotic relative efficiency* of $\mathcal{D}_{f_1,n}$ with respect to $\mathcal{D}_{f_2,n}$.

Throughout the past decades this concept of efficiency was applied to the tests based on various power divergence statistics $\mathcal{D}_{f_{\alpha},n} \equiv \mathcal{D}_{\alpha,n}$, $\alpha \in \mathbb{R}$. The first known result of this kind is $\mathcal{BE}(\mathcal{I}_n, \chi_n^2) \equiv \mathcal{BE}(\mathcal{D}_{1,n}, \mathcal{D}_{2,n}) = \infty$ obtained by [9]. Results concerning the Bahadur functions $g_{f_{\alpha}}(\Delta)$ for some power divergence statistics $\mathcal{D}_{f_{\alpha},n}$ can be found in [4], [2], and [5]. Recently [6] proved the following result.

Theorem 3. If $k^{1+\beta} \ln n/n \rightarrow 0$ holds instead of $k^{1+\beta}/n \rightarrow 0$ assumed in Section 1 then the Bahadur efficiency $\mathcal{BE}(\mathcal{D}_{\alpha_1,n}, \mathcal{D}_{\alpha_2,n})$ exists for all $0 < \alpha_1 < \alpha_2$ and is given by the formula

$$\mathcal{BE}(\mathcal{D}_{\alpha_1,n}, \mathcal{D}_{\alpha_2,n}) = \begin{cases} g_{\alpha_1}(\Delta_{\alpha_1})/g_{\alpha_2}(\Delta_{\alpha_2}) & \text{if } 0 < \alpha_2 \leq 1 \\ \infty & \text{if } \alpha_2 > 1 \end{cases}$$

where

$$g_{\alpha}(\Delta) = \begin{cases} \ln(1 + \alpha(\alpha - 1)\Delta)/(\alpha - 1) & \text{if } 0 < \alpha < 1 \\ \lim_{\alpha \uparrow 1} g_{\alpha}(\Delta) = \Delta & \text{if } \alpha = 1 \end{cases}$$

are the functions corresponding in the sense of (18) to the statistics $\mathcal{D}_{\alpha,n}$.

This result extends what was previously achieved in this area. It says that the assumptions (16)-(18) can be satisfied by all power divergence statistics

$\mathcal{D}_{\alpha,n}$, $\alpha > 0$ including the particular likelihood ratio statistic $\mathcal{D}_{1,n} = \mathcal{I}_n$, and that the Bahadur efficiency of these statistics decreases in the power parameter domain $\alpha \in [1, \infty)$. Moreover, for small $\Delta_{\alpha_1} = \Delta_{\alpha_2}$ (i.e. for "not too large" large deviations) it holds $g_{\alpha_i}(\Delta_{\alpha_i}) \approx \alpha_i \Delta_{\alpha_i}$. Therefore

$$\mathcal{BE}(\mathcal{D}_{\alpha_1,n}, \mathcal{D}_{\alpha_2,n}) = g_{\alpha_1}(\Delta_{\alpha_1})/g_{\alpha_2}(\Delta_{\alpha_2}) \approx \alpha_1/\alpha_2 < 1 \quad \text{if } 0 < \alpha_1 < \alpha_2,$$

i.e. the Bahadur efficiency increases in the power parameter domain $\alpha \in (0, 1]$. The main message of this research is thus the global maximum achieved at $\alpha = 1$, i.e. the Bahadur supremacy of the likelihood ratio statistic $\mathcal{I}_n = \mathcal{D}_{1,n}$ over the remaining power divergence divergence statistics $\mathcal{D}_{\alpha,n}$ with positive power parameters $\alpha > 0$.

References

- [1] Bahadur R.R. (1981). *Some Limit Theorems in Statistics*. SIAM, Philadelphia, PA.
- [2] Beirlant J., Devroye L., Györfi L., Vajda I. (2001). *Large deviations of divergence measures on partitions*. Journal of Statistical Planning and Inference, **39**, 1–16.
- [3] Györfi L., Vajda I. (2002). *Asymptotic distributions for goodness-of-fit statistics in a sequence of multinomial models*. Statistics and Probability Letters, **56**, 57–67.
- [4] Györfi L., Morvai G., Vajda I. (2000). *Information-theoretic methods in testing the goodness-of-fit*. Proceedings of International Symposium of IEEE on Information Theory, p. 28, Sorrento, Italy, June 25–30, 2000.
- [5] Harremoës P., Vajda I. (2008a). *On the Bahadur-efficient testing of uniformity by means of the entropy*. IEEE Transactions on Information Theory, **54**, 321–331.
- [6] Harremoës P., Vajda I. (2008b). *Efficiency of entropy testing*. Proceedings of International Symposium of IEEE on Information Theory, Toronto, Canada, June 2008.
- [7] Liese F., Vajda I. (2006). *On divergences and informations in statistics and information theory*. IEEE Transactions on Information Theory, **52**, 4394–4412.
- [8] Morris C. (1975). *Central limit theorems for multinomial sums*. Annals of Statistics, **3**, 165–188.
- [9] Quine M.P., Robinson J. (1985). *Efficiencies of chi-square and likelihood ratio tests*. Annals of Statistics, **13**, 727–742.
- [10] Vajda I. (2003). *Asymptotic laws for stochastic disparity statistics*. Tatra Mountains Mathematical Publications, **26**, 1–12.

Acknowledgement: This paper was supported by the GAČR grant 102/07/1131.

Address: I. Vajda, ÚTIA AV ČR, Pod Vodárenskou věží 4, 18208 Praha 8

E-mail: vajda@utia.cas.cz

RATIO STATISTICS

František Vávra, Tomáš Pavelka, Blanka Šedivá,
Kateřina Vokáčová, Patrice Marek, Martina Neumanová

Keywords: Ratio statistics, tolerance bounds, point estimates, stochastic ordering.

Abstract: The results of independent experiments used for testing of certain method or certain phenomenon are often represented by samples $(r_1, s_1), \dots, (r_n, s_n)$, where $r_i, s_i \geq 0; \forall i = 1, \dots, n, \sum_{i=1}^n r_i > 0$. The main goal of our contribution is to analyse the random variable $\xi = \frac{R}{R+S}$. The analysis is concerned with tolerance bounds of statistic $\xi_n = \frac{\sum_{i=1}^n r_i}{\sum_{i=1}^n r_i + \sum_{i=1}^n s_i}$ and introduces its asymptotic model as well. Furthermore, two particular cases of n are examined. In the first case, n is assumed to be non-random but changeable (e.g. comparison of automatic speech recognition methods). In the other case, n is assumed to be fixed (e.g. calculation of state unemployment rate from all district unemployment rates).

Abstrakt: Předložená práce se zabývá studiem tolerančních mezí statistiky $\xi_n = \frac{\sum_{i=1}^n r_i}{\sum_{i=1}^n r_i + \sum_{i=1}^n s_i}$, kde $(r_1, s_1), \dots, (r_n, s_n)$ jsou pozorování, které popisují výsledky nějakých nezávislých experimentů testujících nějakou metodu nebo jev, $r_i, s_i \geq 0; \forall i = 1, \dots, n, \sum_{i=1}^n r_i > 0$ a $\sum_{i=1}^n s_i > 0$. O statistice je předpokládáno, že měří některou náhodnou proměnnou ve tvaru $\xi_n = \frac{R}{R+S}$. V práci je prezentován asymptotický model této statistiky. Jsou analyzovány dva dílčí případy. V prvním je n nenáhodné ale proměnné (např. testy úspěšnosti metod pro rozpoznávání mluvených slov) V druhém případě je n pevně dané a neměnné (míra nezaměstnanosti státu počítaná z nezaměstnanosti všech okresů).

1 Motivation and model

The observations of the results of independent experiments measuring phenomenon or testing a method are often represented as $(r_1, s_1), \dots, (r_n, s_n)$, where $r_i, s_i \geq 0; \forall i = 1, \dots, n, \sum_{i=1}^n r_i > 0$. The mean value, the variance and the covariance of each variable are known. $\forall i = 1, \dots, n, E\{r_i\} = e_{ir}, E\{s_i\} = e_{is}, \sigma^2(r_i) = \sigma_{ir}^2, \sigma^2(s_i) = \sigma_{is}^2$ and $\forall i = 1, \dots, n; j = 1, \dots, n; E\{(s_i - e_{is})(r_j - e_{jr})\}$. The quality of the tested method or the rate of certain phenomenon is usually measured by the following criterion

$$\xi_n = \frac{\sum_{i=1}^n r_i}{\sum_{i=1}^n r_i + \sum_{i=1}^n s_i}.$$

This criterion is used in many tasks. One example of such application, which is demonstrated in our contribution, is the testing of automatic speech recognition methods, where:

r_i represents the number of recognised words in i -th sentence of the length N_i ,

s_i represents the number of incorrectly recognised words in i -th sentence of the length N_i ,

N_i represents the length of the sentence in number of words,

n represents the number of sentences on which the method is tested.

The number n is changeable, i.e. can be different for different tests. Another example of application is the measurement of the unemployment rate of the whole country, which uses the numbers of the unemployed in all particular districts for its evaluation. In this case,

r_i represents the number of the registered unemployed in i -th district,

s_i represents the number of the employed in i -th district,

n represents the number of districts which is fixed, i.e. it is the same for all the tests performed.

2 Probability description

The criterion

$$\xi_n = \frac{\sum_{i=1}^n r_i}{\sum_{i=1}^n r_i + \sum_{i=1}^n s_i}$$

is the random variable, which is computed from the results of the experiments. The criterion could be transformed to

$$\xi_n = \frac{1}{1 + \frac{\sum_{i=1}^n s_i}{\sum_{i=1}^n r_i}} = \frac{1}{1 + \eta_n}.$$

Let us assume the distribution function of the variable $\eta_n = \frac{\sum_{i=1}^n s_i}{\sum_{i=1}^n r_i}$, $F_{\eta_n}(x)$ is known. Then the distribution function

$$F_{\xi_n}(x) = P\{\xi_n < x\} = P\left\{\frac{1}{1 + \eta_n} < x\right\}$$

could be expressed as follows

$$F_{\xi_n}(x) = 1 - F_{\eta_n}\left(\frac{1}{x} - 1\right) \Leftrightarrow 0 < x < 1; \quad F_{\xi_n} = 0 \Leftrightarrow x \leq 0; \quad F_{\xi_n} = 1 \Leftrightarrow x \geq 1.$$

Distribution function of variable η_n

$$F_{\eta_n}(x) = P\{\eta_n < x\} = P\left\{\frac{\sum_{i=1}^n s_i}{\sum_{i=1}^n r_i} < x\right\} = P\left\{\sum_{i=1}^n s_i - x \sum_{i=1}^n r_i < 0\right\} = (A)$$

Further on, the following notation will be used $e_r = \frac{1}{n} \sum_{i=1}^n e_{ir}$ and

$e_s = \frac{1}{n} \sum_{i=1}^n e_{is}$. The random variable $\sum_{i=1}^n s_i - x \sum_{i=1}^n r_i$ has the mean value of

$$E\left\{\sum_{i=1}^n s_i - x \sum_{i=1}^n r_i\right\} = n(e_s - xe_r).$$

The variance of this variable has the following form

$$\sigma^2\left\{\sum_{i=1}^n s_i - x \sum_{i=1}^n r_i\right\} = n[\sigma_s^2 + x^2\sigma_r^2 - 2x\rho_{s,r}\sigma_s\sigma_r],$$

where

$$\sigma_r^2 = \frac{1}{n} \sum_{i=1}^n E\{(r_i - e_r)^2\},$$

$$\sigma_s^2 = \frac{1}{n} \sum_{i=1}^n E\{(s_i - e_s)^2\},$$

$$\rho_{s,r}\sigma_s\sigma_r = \frac{1}{n} \sum_{i=1}^n E\{(s_i - e_s)(r_i - e_r)\}$$

and x is a fixed known value for which the distribution function is computed.

We can continue with the derivation of the distribution function $F_{\eta_n}(x)$:

$$\begin{aligned} (A) &= P\left\{\sum_{i=1}^n s_i - x \sum_{i=1}^n r_i < 0\right\} = \\ &= P\left\{\sum_{i=1}^n s_i - x \sum_{i=1}^n r_i - n(e_s - xe_r) < -n(e_s - xe_r)\right\} = \\ &= P\left\{\frac{\sum_{i=1}^n s_i - x \sum_{i=1}^n r_i - n(e_s - xe_r)}{\sqrt{n(\sigma_s^2 + x^2\sigma_r^2 - 2x\rho_{s,r}\sigma_s\sigma_r)}} < \frac{-n(e_s - xe_r)}{\sqrt{n(\sigma_s^2 + x^2\sigma_r^2 - 2x\rho_{s,r}\sigma_s\sigma_r)}}\right\} = \\ &= (B) \end{aligned}$$

The left side of the inequality $\frac{\sum_{i=1}^n s_i - x \sum_{i=1}^n r_i - n(e_s - xe_r)}{\sqrt{n(\sigma_s^2 + x^2\sigma_r^2 - 2x\rho_{s,r}\sigma_s\sigma_r)}}$ is a centralized and standardized random variable whose distribution function for increasing n converges to the distribution function of the standard normal distribution

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{z^2}{2}} dz.$$

(The proofs with different variants of additional assumptions are described in e.g. [1] p. 374 - 382.) Therefore, for a large enough n the following is valid:

$$(B) = \Phi \left(-\sqrt{n} \frac{e_s - xe_r}{\sqrt{\sigma_s^2 + x^2\sigma_r^2 - 2x\rho_{s,r}\sigma_s\sigma_r}} \right).$$

$$\text{Summing up: } F_{\eta_n}(x) = P\{\eta_n < x\} = \Phi \left(-\sqrt{n} \frac{e_s - xe_r}{\sqrt{\sigma_s^2 + x^2\sigma_r^2 - 2x\rho_{s,r}\sigma_s\sigma_r}} \right).$$

Distribution function of variable ξ_n

The distribution function of the criterion ξ_n can be described as follows:

$$F_{\xi_n}(x) = 1 - \Phi \left(-\frac{\sqrt{n} \cdot (e_s - (\frac{1}{x} - 1)e_r)}{\sqrt{(\sigma_s^2 + (\frac{1}{x} - 1)^2\sigma_r^2 - 2(\frac{1}{x} - 1)\rho_{s,r}\sigma_s\sigma_r)}} \right) \Leftrightarrow 0 < x < 1.$$

Such expression is quite complicated, nevertheless easy to evaluate.

3 Point estimate of the measure of location

The most representative measure of location for such type of distribution is the median, which is the solution of the equation $F_{\xi_n}(x_{med}) = \frac{1}{2}$. Thus

$$e_s = \left(\frac{1}{x_{med}} - 1 \right) \cdot e_r \Rightarrow x_{med} = \frac{e_r}{e_r + e_s}.$$

4 Tolerance interval

The goal of this part is to find bounds $x_L < x_U$ such that

$$P(x_L \leq \xi_n \leq x_U) = 1 - \alpha,$$

where α is the confidence level. In other words, we want to find the percentile range that represents a specified proportion of the population. The relation $P(x_L \leq \xi_n \leq x_U) = 1 - \alpha$ does not uniquely determine the bounds. Therefore, α_1, α_2 such that

$$\alpha = \alpha_1 + \alpha_2; \quad 0 < \alpha, \alpha_1, \alpha_2 < 1$$

are chosen to satisfy $P(x_U < \xi_n) = \alpha_1 \Leftrightarrow P(\xi_n < x_U) = 1 - \alpha_1$ and $P(\xi_n < x_L) = \alpha_2$. The particular bound determination is realized by solving the equations for the distribution function $F_{\xi_n}(x_U) = 1 - \alpha_1, F_{\xi_n}(x_L) = \alpha_2$.

Determination of the lower bound x_L

x_L is the solution of the equation

$$1 - \Phi \left(-\sqrt{n} \frac{e_s - \left(\frac{1}{x} - 1\right) e_r}{\sqrt{\left(\sigma_s^2 + \left(\frac{1}{x} - 1\right)^2 \sigma_r^2 - 2\left(\frac{1}{x} - 1\right) \rho_{s,r} \sigma_s \sigma_r\right)}} \right) = \alpha_2,$$

thus

$$\Phi^{-1}(1 - \alpha_2) = \left(-\sqrt{n} \frac{e_s - \left(\frac{1}{x_L} - 1\right) e_r}{\sqrt{\left(\sigma_s^2 + \left(\frac{1}{x_L} - 1\right)^2 \sigma_r^2 - 2\left(\frac{1}{x_L} - 1\right) \rho_{s,r} \sigma_s \sigma_r\right)}} \right).$$

If the notation $z = \frac{1}{x_L} - 1$ is used, the equation has the form of:

$$\begin{aligned} \Phi^{-1}(1 - \alpha_2) &= \left(-\sqrt{n} \frac{e_s - z e_r}{\sqrt{\sigma_s^2 + z^2 \sigma_r^2 - 2z \rho_{s,r} \sigma_s \sigma_r}} \right) \Rightarrow \\ \Phi^{-1}(1 - \alpha_2) \sqrt{\sigma_s^2 + z^2 \sigma_r^2 - 2z \rho_{s,r} \sigma_s \sigma_r} &= -\sqrt{n}(e_s - z e_r) \end{aligned}$$

The next equation modification is squaring. By this step, the solution is represented by one of the roots of the equation

$$[\Phi^{-1}(1 - \alpha_2)]^2 (\sigma_s^2 + z^2 \sigma_r^2 - 2z \rho_{s,r} \sigma_s \sigma_r) = n (e_s - z e_r)^2,$$

which will be identified by backward testing of the validity of these roots on the equation before the squaring. Finally, this letter notation was used to express the equation in a more transparent form:

$$Az^2 + 2zB + C = 0,$$

where $A = (a\sigma_r^2 - e_r^2),$
 $B = (e_s e_r - \rho_{s,r} a \sigma_r \sigma_s),$
 $C = a\sigma_s^2 - e_s^2$ and
 $a = \frac{[\Phi^{-1}(1 - \alpha_2)]^2}{n}.$

Two roots $z_{1,2} = \frac{-B \pm \sqrt{B^2 - AC}}{A}$ are evaluated by solving this quadratic equation. As mentioned before, the identification of the valid solution is arranged by checking whether the particular root satisfies the following equation

$$\frac{\Phi^{-1}(1 - \alpha_2)}{\sqrt{n}} \sqrt{\sigma_s^2 + z^2 \sigma_r^2 - 2z \rho_{s,r} \sigma_s \sigma_r} + (e_s - z e_r) = 0.$$

Only one root is valid. Let us denote the solution z_* . Using the transformation $z_* = \frac{1}{x_L} - 1$, the solution $x_L = \frac{1}{1+z_*}$ is obtained.

Determination of the upper bound x_U

The determination of the x_U results from solving the equation

$$\Phi \left(-\sqrt{n} \frac{e_s - \left(\frac{1}{x_U} - 1\right) e_r}{\sqrt{\left(\sigma_s^2 + \left(\frac{1}{x_U} - 1\right)^2 \sigma_r^2 - 2 \left(\frac{1}{x_U} - 1\right) \rho_{s,r} \sigma_s \sigma_r\right)}} \right) = \alpha_1.$$

The calculation is almost identical to the previous calculation of x_L . The only differences are

$$a = \frac{[\Phi^{-1}(\alpha_1)]^2}{n}, z_{1,2} = \frac{-B \pm \sqrt{B^2 - AC}}{A}$$

and the equation for testing the validity of the roots is

$$\frac{\Phi^{-1}(\alpha_1)}{\sqrt{n}} \sqrt{\sigma_s^2 + z^2 \sigma_r^2 - 2z \rho_{s,r} \sigma_s \sigma_r} + (e_s - z e_r) = 0.$$

5 Acceptability of asymptotic representation

The formula of the asymptotic version of the distribution function $F_{\xi_n}(x)$ does not necessarily need to have the characteristics of the distribution function, and it usually has not. Furthermore, ξ_n is the random variable defined on the interval $(0, 1)$. The formula of $F_{\xi_n}(x)$ exceeds this interval by not having the functional value equal to zero or one in the boundary elements of the interval. These circumstances could be used for testing the acceptability of such asymptotic representation. For these purposes, following assumptions are made:

1. The function $F_{\xi_n}(x)$ should be a non-decreasing function in the interval $(0, 1)$.
2. The value of $F_{\xi_n}(x)$ for x in the "zero value" should be non-negative and smaller than a known small positive number β_0 . The value of $F_{\xi_n}(x)$ for x in the point $x = 1$ should not exceed one and should be greater than $1 - \beta_1$, where β_1 is positive and sufficiently small.

The first assumption

The first assumption is fulfilled if $\frac{d}{dx} F_{\xi_n}(x) = f_{\xi_n}(x)$ is non-negative in the interval $(0, 1)$. The artificial variable $z = \frac{1}{x} - 1$ is used again for the sake of transparency of the expressions mentioned below.

$$F_{\xi_n}(z) = 1 - \Phi \left(-\sqrt{n} \frac{e_s - z e_r}{\sqrt{\sigma_s^2 + z^2 \sigma_r^2 - 2z \rho_{s,r} \sigma_s \sigma_r}} \right); \quad 0 < z < +\infty \quad \text{and} \quad \frac{dz}{dx} = -\frac{1}{x^2}.$$

Then

$$f_{\xi_n}(x) = \frac{d}{dx} F_{\xi_n}(x) = \frac{d}{dz} F_{\xi_n}(z) \frac{dz}{dx} = \frac{d}{dx} \left[1 - \Phi \left(-\sqrt{n} \frac{e_s - ze_r}{\sqrt{\sigma_s^2 + z^2 \sigma_r^2 - 2z \rho_{s,r} \sigma_s \sigma_r}} \right) \right] \frac{dz}{dx}.$$

Taking into account the negative sign of $\frac{dz}{dx}$, the sign of $f_{\xi_n}(x)$ is decisive.

$$\begin{aligned} \frac{d}{dz} \left[1 - \Phi \left(-\sqrt{n} \frac{e_s - ze_r}{\sqrt{\sigma_s^2 + z^2 \sigma_r^2 - 2z \rho_{s,r} \sigma_s \sigma_r}} \right) \right] &= \\ = -\sqrt{n} \varphi \left(-\sqrt{n} \frac{e_s - ze_r}{\sqrt{\sigma_s^2 + z^2 \sigma_r^2 - 2z \rho_{s,r} \sigma_s \sigma_r}} \right) &\left(\frac{e_r (\sigma_s^2 - z \rho_{s,r} \sigma_s \sigma_r) + e_s (z \sigma_r^2 - \rho_{s,r} \sigma_s \sigma_r)}{(\sigma_s^2 + z^2 \sigma_r^2 - 2z \rho_{s,r} \sigma_s \sigma_r)^{\frac{3}{2}}} \right) \end{aligned}$$

The condition $f_{\xi_n}(x) \geq 0$ is fulfilled for such $z = \frac{1}{x} - 1$, where

$$e_r (\sigma_s^2 - z \rho_{s,r} \sigma_s \sigma_r) + e_s (z \sigma_r^2 - \rho_{s,r} \sigma_s \sigma_r) \geq 0.$$

The first step of the test is to find the intercept point of the monotonic function, i.e. the solution of the equation

$$e_r (\sigma_s^2 - z \rho_{s,r} \sigma_s \sigma_r) + e_s (z \sigma_r^2 - \rho_{s,r} \sigma_s \sigma_r) = 0.$$

Thus

$$z = \frac{\sigma_s \rho_{s,r} e_s \sigma_r - e_r \sigma_s}{\sigma_r e_s \sigma_r - \rho_{s,r} e_r \sigma_s}.$$

The other step is to check if the intercept point x is out of the interval $(0, 1)$. This is assured by $\frac{\rho_{s,r} e_s \sigma_r - e_r \sigma_s}{e_s \sigma_r - \rho_{s,r} e_r \sigma_s} \leq 0$, because $x = \frac{1}{1+z}$. The next step is the verification of the inequality in the median point.

$$x_{med} = \frac{e_r}{e_r + e_s} \Rightarrow z_{med} = \frac{e_s}{e_r},$$

hence

$$e_r (\sigma_s^2 - z_{med} \rho_{s,r} \sigma_s \sigma_r) + e_s (z_{med} \sigma_r^2 - \rho_{s,r} \sigma_s \sigma_r) = \frac{(e_r \sigma_s - e_s \sigma_r)^2}{e_r} \geq 0.$$

For this reason, the condition $\frac{\rho_{s,r} e_s \sigma_r - e_r \sigma_s}{e_s \sigma_r - \rho_{s,r} e_r \sigma_s} \leq 0$ represents the necessary and sufficient condition of $f_{\xi_n}(x) \geq 0$; $0 < x \leq 1$. If this condition is satisfied, the asymptotic approximation will be considered correct. The correctness is important for the “small n”.

Note: The mentioned correctness condition is not dependent on the number of observations.

The second assumption

When the numbers $0 < \beta_0, \beta_1, \beta = \beta_0 + \beta_1 < 1$ are set, the asymptotic approximation could be considered as β -acceptable, if the inequalities

$$1 \geq F_{\xi_n}(1) \geq 1 - \beta_1 \text{ and } 0 \leq \lim_{x \rightarrow 0} F_{\xi_n}(0) \leq \beta_0$$

are fulfilled. It means that $F_{\xi_n}(1) = 1 - \Phi\left(-\sqrt{n}\frac{e_s}{\sigma_s}\right)$, so

$$1 - \Phi\left(-\sqrt{n}\frac{e_s}{\sigma_s}\right) \geq 1 - \beta_1 \Rightarrow n \geq \frac{\sigma_s^2}{e_s^2} (\Phi^{-1}(\beta_1))^2.$$

The other $\lim_{x \rightarrow 0^+} F_{\xi_n} = 1 - \Phi\left(-\sqrt{n}\frac{e_r}{\sigma_r}\right)$, and so

$$1 - \Phi\left(-\sqrt{n}\frac{e_r}{\sigma_r}\right) \leq \beta_0 \Rightarrow n \geq \frac{\sigma_r^2}{e_r^2} (\Phi^{-1}(1 - \beta_0))^2.$$

Summing up: The asymptotic approximation could be considered as $\beta = \beta_0 + \beta_1$ acceptable if and only if $n \geq \max\left\{\frac{\sigma_s^2}{e_s^2} (\Phi^{-1}(\beta_1))^2; \frac{\sigma_r^2}{e_r^2} (\Phi^{-1}(1 - \beta_0))^2\right\}$

Note: The condition of β -acceptability is dependent on the observation number n .

6 Conclusion

The mentioned methods can be applied in the cases where the measurements are in the form of a ratio criterion. The ratio criterion takes the values from the interval $(0; 1)$ for the purposes of our contribution. The extension of this interval by "one" could be arranged just by a technical procedure. Nevertheless, it is necessary to solve a few problems with the correctness and acceptability of the asymptotic approximation. Further research can focus on the stochastic comparison.

References

- [1] Rényi A. (1972). *Teorie pravděpodobnosti*. ACADEMIA, Paris.
- [2] Jílek M. (1988). *Statistické toleranční meze*. SNTL, Praha.

Acknowledgement: The research has been supported by MSM4977751301 research plan - Continuous and Discrete Mathematical Structures and Development of Corresponding Methods of their Study; MPO 2A - 2TP1/051 grant - Improving Reliability and Safety of Electrical Networks; NPV II: 2C06009 Cot-Sewing grant - Complex knowledge base tools for natural language communication with the semantic web.

Address: F. Vávra, T. Pavelka, B. Šedivá, K. Vokáčová, P. Marek, M. Neumanová, ZČU FAV, KMA, Univerzitní 22, 306 14 Plzeň

E-mail: vavra@kma.zcu.cz

WEIGHTED DATA DEPTH AND ITS PROPERTIES

Ondřej Vencálek

Keywords: Data depth, weighted halfspace depth, support of probability measure.

Abstract: The paper deals with nonparametric methods for analysis of multivariate data. Generalization of the halfspace depth, so called weighted halfspace depth, is introduced. We consider its properties with respect to the weight function in its definition. One particular property of our interest: points out of the support of probability measure should have depth equal to zero.

Abstrakt: Tento příspěvek z oblasti neparametrických metod analýzy mnohorozměrných dat představuje možnost zobecnění poloprostorové hloubky bodu, tzv. váženou poloprostorovou hloubku, a zabývá se jejími vlastnostmi v závislosti na volbě váhy v její definici. Speciálně nás zajímá oblast nulové hloubky a její vztah k nosiči rozdělení.

1 Introduction

Many statistical problems deal with multivariate random vectors. The classical theory of multivariate data analysis is based on the assumption of multivariate normal distribution of the data. This restriction is often disputable so another, nonparametric, approach should be considered.

One of the most important approaches to the multivariate data is based on so called data depth. The concept of data depth is based on ordering of points in \mathbb{R}^p . The principle of ordering is very simple. We can add one real value number to a certain point $\mathbf{x} \in \mathbb{R}^p$ that describes its centrality with respect to some probability distribution P . This number is called *depth of point \mathbf{x} with respect to P* . We denote it $D_P(\mathbf{x})$. The way we add these real numbers to points in \mathbb{R}^p is described by so called *depth function*. We denote it $D_P : \mathbb{R}^p \rightarrow \mathbb{R}^+$. We order points in \mathbb{R}^p according to their depth.

The J.Tukey's definition of the halfspace depth function (1975, see [4]) brought quite simple but very natural way of central-outward ordering of points with respect to the probability distribution. The rigorous definition is recalled in Section 2. The approach based on the data depth has been developed since then. A new wave of interest of this approach came in 1990s and has continued so far. Review of this methodology can be found e.g. in [1]. New results are introduced in [2].

Nonparametric statistical inference for multivariate random vectors has been developed based on the data depth. But the definition of the depth function alone is still not unified. Many various depth functions have been

introduced. The most well known are simplicial depth, Oja depth, L_1 -depth and convex hull peeling depth. General definition of the depth function was proposed by Zuo and Serfling (see [3]) in 2000. This general definition says that the depth function is an arbitrary function which has some desirable properties (we introduce them here without detailed comment: a depth function is any nonnegative bounded function which is affine invariant, maximal at the center, monotone relatively to the deepest point and which goes to zero at infinity).

We propose generalization of the halfspace depth. We call it weighted halfspace depth. Its formal definition can be found in Section 2. We investigate its properties. They depend on a weight function that we use for weighted halfspace depth. Not all weight functions are proper. We found some very mild conditions under that the weight function is proper in the sense that all points out of the support of probability have weighted halfspace depth equal to zero. These conditions are presented in Section 3 for a probability distribution with convex support. If we consider nonconvex support of probability measure the conditions are very similar. They are presented in Section 4.

2 Halfspace depth and weighted halfspace depth

In this section we present some basic definitions. Namely the definition of the halfspace depth and the weighted halfspace depth. We also recall the definition of the closed convex support of probability measure.

Definition 1: Let P be a probability measure on \mathbb{R}^p . The halfspace depth of the point $\mathbf{x} \in \mathbb{R}^p$ with respect to P is defined as

$$D_P(\mathbf{x}) = \inf_{\mathbf{u}: \|\mathbf{u}\|=1} P\{\mathbf{y} : \mathbf{u}^T(\mathbf{y} - \mathbf{x}) \geq 0\}$$

If we consider empirical probability measure \widehat{P}_n instead of P we get a sample version of the halfspace depth. Zuo and Serfling showed (in [3]) that the halfspace depth has all desirable properties of the depth function. One particular property of the halfspace depth is convexity of the areas of the deepest points i.e. for arbitrary P and $d \in (0, 1)$ the set $\{\mathbf{x} : D_P(\mathbf{x}) \geq d\}$ is convex. This property could be considered as undesirable (e.g. for the uniform distribution on a nonconvex support). Therefore we introduce a generalization of the halfspace depth which has not this property in general. Defining the weighted halfspace depth we follow several consecutive steps. The definition of the halfspace depth can be expressed in the following form:

$$D_P(\mathbf{x}) = \inf P\{H : \text{a closed halfspace: } \mathbf{x} \in H\} = \inf_H \int_H dP,$$

where the infimum is over all closed halfspaces H that contain \mathbf{x} . Our modification consists in integration of some weight function (w.r.t. P). So we have

$$p_H = \int_H w(X)dP.$$

Now we are ready to define weighted halfspace depth of a certain point \mathbf{x} :

Definition 2: Consider p -dimensional space \mathbb{R}^p , a point $\mathbf{x} = (x_1, \dots, x_p)$, $\mathbf{x} \in \mathbb{R}^p$ and a probability measure P on \mathbb{R}^p . Let $w_+ : \mathbb{R}^p \rightarrow [0, \infty)$ be a bounded measurable weight function such that $w_+(\mathbf{x}) = 0$ for $x_p < 0$ and $w_-(\mathbf{x}) = w_-(x_1, \dots, x_{p-1}, x_p) = w_+(x_1, \dots, x_{p-1}, -x_p)$.

We define the weighted halfspace depth of the point \mathbf{x} with respect to P as

$$D_{\mathbf{X}}(\mathbf{x}) = D_P(\mathbf{x}) = \inf_{\mathbf{A} \in O_p} \frac{\mathbb{E}_P w_+(\mathbf{A}(\mathbf{X} - \mathbf{x}))}{\mathbb{E}_P w_-(\mathbf{A}(\mathbf{X} - \mathbf{x}))}, \quad (1)$$

where O_p denotes the space of all orthogonal $p \times p$ matrices.

We define the value of $0/0 = 1$.

Let's start with a lemma about two general properties of this depth that hold for any weight function.

Lemma 1: The depth function defined by the formula (1) is

- translation invariant, i.e. $D_{\mathbf{X}+\mathbf{a}}(\mathbf{x} + \mathbf{a}) = D_{\mathbf{X}}(\mathbf{x})$
- rotation invariant, i.e. $D_{\mathbf{A}\mathbf{X}}(\mathbf{A}\mathbf{x}) = D_{\mathbf{X}}(\mathbf{x})$,

where \mathbf{A} is the matrix of the rotation.

Proof: Proof follows directly from the definition. Considering rotation invariance we just note that $\{\mathbf{B}\mathbf{A} : \mathbf{B} \in O_p\} = O_p$. \square

We want to discuss the choice of the weight function in Definition 2. One possible criterion is the property that all points out of the closed convex support of probability measure have depth equal to zero. It means:

$$\mathbf{x} \notin \text{csp}(P) \Rightarrow H_P(\mathbf{x}) = 0. \quad (2)$$

Recall that the support of a probability measure P (we denote it $\text{sp}(P)$) is the smallest closed set with the probability equal to 1, i.e.

$$\text{sp}(P) = \bigcap \{F \in \mathcal{F} : P(F) = 1\},$$

where \mathcal{F} is the class of all closed subsets of \mathbb{R}^p . The closed convex support of probability measure P is defined as closed convex hull of the support $\text{sp}(P)$. It is denoted $\text{csp}(P)$.

3 Convex support of probability measure

Consider a probability distribution with convex support, e.g. the multivariate exponential distribution or the uniform distribution on a convex support. Let's consider only two-dimensional space ($p = 2$) for simplicity. The following theorem introduces the sufficient condition for property (2):

Theorem 1: Let $\text{sp}(P)$ be convex. We denote $W = \{\mathbf{y} : w_+(\mathbf{y}) > 0\}$. Suppose that

$$\forall \mathbf{x} : x_1 = 0, x_2 > 0 \exists U \text{ a neighbourhood of the point } \mathbf{x} : U \subset W, \quad (3)$$

holds for the weight function. Then $H_P(\mathbf{x}) = 0$ holds for all $\mathbf{x} \notin \text{sp}(P)$.

Proof: Suppose $\mathbf{x}_0 \notin \text{sp}(P)$. We want to prove that its weighted depth will be equal to zero when condition (3) holds.

We denote $\mathbf{x}_m = \arg \min \{|\mathbf{x} - \mathbf{x}_0| : \mathbf{x} \in \text{sp}(P)\}$, i.e. \mathbf{x}_m is the point of the support of P with the smallest distance from \mathbf{x}_0 . Existence and uniqueness of this point arise from the convexity of $\text{sp}(P)$. We can suppose (without loss of generality) $\mathbf{x}_0 = (0, 0)$ because of the translation invariance of the weighted depth (see Lemma 1) and we can suppose (without loss of generality) $\mathbf{x}_m = (0, v)$, where $v = |\mathbf{x} - \mathbf{x}_0|$ because of the rotation invariance of the weighted depth.

For such a rotation $\text{sp}(P) \subset H_v \subset H_0$ holds, where $H_v = \{\mathbf{x} : x_2 \geq v\}$ and $H_0 = \{\mathbf{x} : x_2 \geq 0\}$. We can prove it by an absurdum proof. Suppose that there exist $\mathbf{y} \in \text{sp}(P)$ such that $y_2 < v$. Then (from the convexity of $\text{sp}(P)$) all points on the abscissa \mathbf{x}_m, \mathbf{y} are in $\text{sp}(P)$, i.e.

$$\mathbf{x}_m + \alpha(\mathbf{y} - \mathbf{x}_m) \in \text{sp}(P) \quad \forall \alpha \in [0, 1].$$

The distance of these points from the origin (\mathbf{x}_0) can be expressed as $[(\alpha y_1)^2 + (v + \alpha(y_2 - v))^2]^{1/2}$, what is, for α small enough, smaller than v . But this is in conflict with the assumption that $\mathbf{x}_m = (0, v)$ is the point of $\text{sp}(P)$ with the smallest distance from $\mathbf{x}_0 = (0, 0)$.

From (3) we have that there exist $U_{\mathbf{x}_m}$ a neighbourhood of the point \mathbf{x}_m such that $U_{\mathbf{x}_m} \subset W$. So we have $U_{\mathbf{x}_m} \cap W \cap \text{sp}(P) \neq \emptyset$, hence $E_P w_+(\mathbf{A}(\mathbf{X} - \mathbf{x}_0)) > 0$. It follows from $\text{sp}(P) \subset H_v \subset H_0$ that $E_P w_-(\mathbf{A}(\mathbf{X} - \mathbf{x}_0)) = 0$, so we have $H_P(\mathbf{x}_0) = 0$. \square

The class of functions that satisfy the condition (3) from Theorem 1 is still quite broad. Here are some examples of weight functions that satisfies condition (3):

- $w_+(\mathbf{x}) > 0$ iff $x_2 \geq 0$,
- $w_+(\mathbf{x}) > 0$ iff $x_2 \geq 0$ and $|x_1| \leq h$,

- $w_+(\mathbf{x}) > 0$ iff $x_2 \geq 0$ and $|x_1| \leq hx_2$,
- $w_+(\mathbf{x}) > 0$ iff $x_2 \geq 0$ and $|x_1| \leq h/x_2$,

where h is a positive constant. We can see that the condition is really very mild and not too restrictive. However, this condition is not necessary. E.g., consider the uniform distribution on the convex (bounded) support and weight function

$$w_+(\mathbf{x}) > 0 \text{ if and only if } x_p \in [0, h],$$

where h is a positive constant. The condition (3) is not satisfied, but (2) holds.

The form of the condition (3) for general multivariate case ($p \in \mathbb{N}$) can be written as follows: For all $\mathbf{x} : x_1 = 0, \dots, x_{p-1} = 0, x_p > 0$ there exist U , a neighbourhood of the point \mathbf{x} , such that $U \subset W$.

4 Nonconvex support of probability measure

Now we will consider a probability measure P with nonconvex support. Because of nonconvexity we do not request depth equal to zero for all points out of the support itself, but only for all points out of the closed convex hull of the support ($\text{csp}(P)$). E.g., if we consider the uniform distribution on an annulus, the center of this annulus is the center of symmetry of the distribution, but it is still out of the support. So we do not demand its depth equal to zero.

For a nonconvex support of P the condition (3) is no more tenable. We can either consider $\text{sp}(P)$ connected or strenghten the condition on weight function. One possible way how to do it for such a case is introduced in Theorem 2.

Theorem 2: Suppose there exists $n \in \mathbb{N}$ such that $\text{sp}(P) = \bigcup_{i=1}^n K_i$, where K_i ($i=1, \dots, n$) is connected subset of \mathbb{R}^2 , and $\text{sp}(P)$ has no singular point.

Denote $m_{ij} = \min \{|\mathbf{x} - \mathbf{y}| : \mathbf{x} \in K_i, \mathbf{y} \in K_j\}$, $i, j = 1, \dots, n$.

Consider $m = \max_{1 \leq i, j \leq n} m_{ij}$.

Let a weight function w_+ have the following property:

$$\forall \mathbf{x} : |x_1| \leq m/2 \exists U_{\mathbf{x}} \text{ a neighbourhood of } \mathbf{x} \text{ such that } U_{\mathbf{x}} \subset W. \quad (4)$$

Then $\mathbf{x} \notin \text{csp}(P) \Rightarrow H_P(\mathbf{x}) = 0$.

Proof: The proof is very similar to the previous one. We denote the point of $\text{csp}(P)$ with the smallest distance from $\mathbf{x}_0 = (0, 0)$ by $\mathbf{x}_m = (0, v)$.

We will prove via absurdum proof that there exist a point $\mathbf{x} = (x_1, x_2)$ such that $|x_1| \leq m/2$, which is in W ($\mathbf{x} \in W$). Suppose that there is no such a point \mathbf{x} . Then there must be points $\mathbf{y} = (y_1, y_2) \in \text{sp}(P)$ and $\mathbf{z} = (z_1, z_2) \in \text{sp}(P)$ such that $y_1 < -m/2$ and $z_1 > m/2$. Hence $|\mathbf{y} - \mathbf{z}| > m$. These points are from the different components of connectedness. We take two points from

these components with the smallest distance between each other: \mathbf{y}_m and \mathbf{z}_m . For all points \mathbf{y} of the one component $y_1 < -m/2$ holds and for all points \mathbf{z} of the other component $z_1 > m/2$ holds, we get $|\mathbf{y}_m - \mathbf{z}_m| > m$, but this is in conflict with the definition of m .

From the assumption (4) follows that there exists $U_{\mathbf{x}}$ a neighbourhood of the point \mathbf{x} such that $U_{\mathbf{x}} \cap W \cap \text{sp}(P) \neq \emptyset$. Hence (similary as in proof of the Theorem 1) $H_P(\mathbf{x}_0) = 0$. \square

Note that in a special case $n=1$ (i.e. for nonconvex connected support) we have depth equal to zero for all points out of the closed convex hull of the support when (3) holds.

We have been discussing the depth of the points out of the closed convex hull of the support so far. Now we will discuss properties of the depth of points that are in the convex hull of the support, but out of the support itself. It is easy to show that all points from the closed convex hull of the support have a positive halfspace depth. An advantage of the weighted halfspace depth is that points in the closed convex hull of the support but out of the support itself might have the depth equal to zero.

We explain the advantage on the following example. Consider the uniform distribution on some sector which originates from the circle with the radius r (see Figure 1, part a). All points from the closed convex hull of the support have the halfspace depth greater then zero (Figure 1, part b). Now we consider the weighted halfspace depth with the following weight function

$$\begin{aligned} w_+(\mathbf{x}) &= 1 && \text{if } |x_1| < h, x_2 \geq 0 \\ &= 0 && \text{otherwise,} \end{aligned} \quad (5)$$

where h is some positive constant smaller than r . Then the area of points that have the weighted halfspace depth greater then zero is the union of the support and the circle with the same center as the big one, but with radius

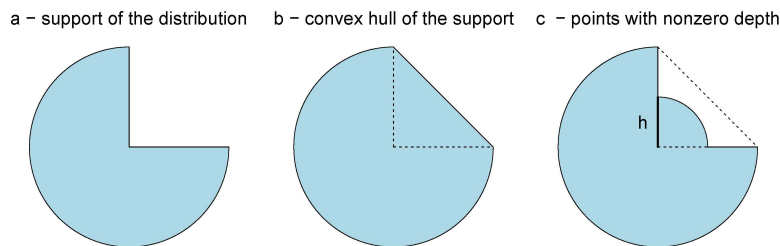


Figure 1: sector-shaped support of the uniform distribution (a); the convex hull of the support i.e. points with nonzero halfspace depth (b); points with nonzero weighted halfspace depth with the weight function (5), where $h = r/2$ (c).

equal to h (Figure 1, part c). Comparing shapes of the areas with nonzero halfspace depth and nonzero weighted halfspace depth we see that the second one is more similar to the shape of the support of probability measure.

References

- [1] Liu R.Y., Parelius J.M. a Singh K. (1999) *Multivariate analysis by data depth: descriptive statistics, graphics and inference*. The Annals of Statistics, Vol. 27, No. 3, 783–858.
- [2] Liu R.Y., Serfling R. a Souvaine D.L.(Eds.) (2006) *DIMACS; Data Depth: Robust Multivariate Analysis*, Computational Geometry and Applications. American Mathematical Society.
- [3] Zuo Y. a Serfling R. (2000) *General notion of statistical depth function*. Annals of Statistics, Vol. 28, 461–482.
- [4] Tukey J. (1975) *Mathematics and picturing data*. In Proceedings of the 1975 International Congress of Mathematics, Vol. 2, 523–531.

Acknowledgement: The research was supported by the grant GA ČR 201/08/0486.

Address: O. Vencálek, MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8 – Karlín

E-mail: `vencalek@karlin.mff.cuni.cz`

CONSISTENCY OF WEIGHTED GENERALIZED METHOD OF MOMENTS

Jan Ámos Víšek

Keywords: Generalized method of moments (GMM), weighted GMM, consistency, robustness.

Abstract: Generalized method of moment was proposed by Lars Peter Hansen [3] to solve the problem of overidentification of model. Paper offers a robustified version of the method and proves its consistency.

Abstrakt: L.P. Larsen navrhl v [3] zobecněnou metodu momentů pro řešení problémů spojených s přeúčněním modelu. Tento příspěvek se zabývá robustifikací uvedené metody a konzistencí navržené modifikace.

1 Introduction

Let \mathcal{N} and R denote the set of integers and the real line, respectively. In social sciences the underlying parametrized model (for a random “response” variable Y we are interested in) is frequently (over)identified by equations

$$\mathbb{E}f(V, \beta^0) = 0 \quad (1)$$

with random variables (r.v.'s) $V: \Omega \rightarrow R^s$ and parameters $\beta^0 \in \mathcal{S}^o \subset R^p$ to be estimated considering an initial part of length n of a sequence of random variables $\{V_i\}_{i=1}^\infty$ (Y is implicitly present in V). $\mathcal{S} \subset R^p$ is parameter space and \mathcal{S}^o its interior. In the rest of paper we shall assume that all r.v.'s are defined on a probability space (Ω, \mathcal{A}, P) and that the *true* value of parameters is vector β^0 . Without loss of generality we may assume that $\beta^0 = 0$ – otherwise we would write everywhere $\beta - \beta^0$ instead of β . We frequently have $V = (Y', X')'$ with $Y: \Omega \rightarrow R^k$ (*response variable*) and $X: \Omega \rightarrow R^p$ (*explanatory variables*), but the structure of V can be generally a bit more complicated, including e.g. also some lagged values of variables in question, instruments, etc.

(1) can be usually interpreted as a collection of orthogonality conditions since $f(V, \beta^0)$ has typically form

$$f(V, \beta^0) = [U_1(V, \beta^0) \cdot Z(V, \beta^0), U_2(V, \beta^0) \cdot Z(V, \beta^0), \dots, U_k(V, \beta^0) \cdot Z(V, \beta^0)] \quad (2)$$

where

$$U = U(V, \beta^0) : \Omega \rightarrow R^k \quad \text{and} \quad Z = Z(V, \beta^0) : \Omega \rightarrow R^q \quad (3)$$

are k -dimensional vector of *error terms (disturbances)* and p -dimensional vector of the (*explanatory variables*) (or, if you want, of the *instruments*), respectively. This framework (introduced - in formally a bit modified way

employing Kronecker product - in the pioneering paper by Hansen [3]) allows generally for possibility of $k > 1$, when we consider e. g. *seemingly unrelated regressions* or *simultaneous equations* (when we consider several regression equations simultaneously - seemingly unrelated equations, or simultaneous equations, see e. g. [4]). The dependence of Z on its second argument is usually trivial, see also [3]. To keep complexity of further text and of considerations in reasonable limits we restrict ourselves on the case $k = 1$ as well as on $Z_i = Z(V_i)$ for all $i \in \mathcal{N}$.

If $p = q$, employing *method of moments*, we can define $\hat{\beta}^{(MM,n)}$ as any solution of the empirical counterpart of (1), i. e. any solution of

$$g_n(\mathbf{V}, \beta) = \frac{1}{n} \sum_{i=1}^n f(V_i, \beta) = 0 \quad (4)$$

where $\mathbf{V} = (V_1, V_2, \dots, V_n)'$. For the case when $p < q$, i. e. when number of orthogonal conditions is larger than dimension of parameter β , (1) need not have any solution. Nevertheless, the orthogonal conditions can keep a sense, although the system of equations (1) is overidentified, see e. g. [10]. Hence Lars Peter Hansen [3] proposed¹ to consider an estimator by *generalized method of moments* (GMM)

$$\hat{\beta}^{(GMM,n)} = \arg \min_{\beta \in R^p} [d_n(\mathbf{V}, \beta)]' \cdot d_n(\mathbf{V}, \beta) \quad (5)$$

where $d_n(\mathbf{V}, \beta) = a_n \cdot g_n(\mathbf{V}, \beta)$ and $\{a_n\}_{n=1}^{\infty}$ is a sequence of matrices of type $(p \times q)$ - generally with possibility of $a_n = a_n(\mathbf{V})$ - which he assumed to converge *almost surely* to a nonrandom matrix a of the full rank. Assuming the matrices $a_n = a_n(V)$ (i. e. assuming a_n to be function of V and n) allows to consider a_n as an estimate of an optimal, nevertheless unknown matrix a (see again [10]). On the other hand, as we assume *almost sure convergence*, the considerations on consistency or asymptotic normality of $\hat{\beta}^{(GMM,n)}$ are nearly the same as if we consider a fix matrix a . So to avoid routine technicalities, let us assume fix matrix a of the full rank².

2 Weighted GMM

Now, notice please that (4) are normal equations for (1). Together with (2) it immediately indicates that atypical values among *residuals* $U(V_1, \beta)$, $U(V_2, \beta), \dots, U(V_n, \beta)$ as well as among *instruments* $Z(V_1), Z(V_2), \dots, Z(V_n)$ can cause the same problems as in the case of regression model (in what follows let us write simply $U_i(\beta)$ instead of $U(V_i, \beta)$). Then, following [2] let

¹We are going to keep the notations as well as the sense of his pioneering paper as much as possible.

²To meet the ends with limited space.

us denote for any $\beta \in R^p$ by $U_{(i)}^2(\beta)$ and by $\pi(\beta, i)$ the i -th order statistic among squared residuals $U_1^2(\beta), U_2^2(\beta), \dots, U_n^2(\beta)$ and the rank of i -th squared residual among the squared residuals, respectively, i. e.

$$U_{(1)}^2(\beta) \leq U_{(2)}^2(\beta) \leq \dots \leq U_{(n)}^2(\beta) \tag{6}$$

and

$$\pi(\beta, i) = j \in \{1, 2, \dots, n\} \Leftrightarrow U_i^2(\beta) = U_{(j)}^2(\beta). \tag{7}$$

Finally, for a nonincreasing continuous function $w : [0, 1] \rightarrow [0, 1]$ let us consider

$$g_n^{(WGMM)}(\mathbf{V}, \beta) = \frac{1}{n} \sum_{i=1}^n w\left(\frac{\pi(\beta, i) - 1}{n}\right) f(V_i, \beta) \tag{8}$$

instead of $g_n(\mathbf{V}, \beta)$ (see (4) and (5)). Prior to defining a robust version of $\hat{\beta}_n^{(GMM)}$, let us put

$$F_\beta(u) = P(|U_1(\beta)| < u) \tag{9}$$

and denote for any $\beta \in R^p$ the empirical distribution of the absolute values of residuals by $F_\beta^{(n)}(r)$, i. e.

$$F_\beta^{(n)}(u) = \frac{1}{n} \sum_{j=1}^n I\{|U_j(\beta)| < u\}. \tag{10}$$

It is straightforward to show (see again [7] or [8]) that

$$F_\beta^{(n)}(|U_i(\beta)|) = \frac{\pi(\beta, i) - 1}{n}.$$

Finally put (see [5])

$$g_n^{(WGMM)}(\mathbf{V}, \beta) = \frac{1}{n} \sum_{i=1}^n w(F_\beta^{(n)}(|U_i(\beta)|)) f(V_i, \beta).$$

Then we can propose, for $A = a' \cdot a$, following definition (compare with Definition 2 in [7]):

Definition 2.1. *The estimator $\hat{\beta}^{(WGMM, n)}$ implied by the Weighted Generalized Method of Moments is given by*

$$\hat{\beta}^{(WGMM, n)} = \underset{\beta \in R^p}{\operatorname{arg\,min}} \left[g_n^{(WGMM)}(\mathbf{V}, \beta) \right]^t \cdot A \cdot g_n^{(WGMM)}(\mathbf{V}, \beta). \tag{11}$$

3 Consistency of $\hat{\beta}^{(WGMM,n)}$

In what follows we need some conditions on r. v.'s (we follow again [3] – for discussion please see Conclusions at the end of paper). The conditions could be given for r. v.'s V_i 's but as we work with the residuals $U_i(\beta)$ and instruments Z_i , let us put them directly on them (let's recall that we have already assumed that the instruments Z_i 's don't depend on β). That is why we refer hereafter exclusively to $U_i(\beta)$'s and Z_i 's so that we'll write e. g. $f(U_i(\beta), Z_i)$ instead of $f(V_i, \beta)$,

$$g_n^{(WGMM)}(\mathbf{U}(\beta), \mathbf{Z}) = \frac{1}{n} \sum_{i=1}^n w(F_\beta^{(n)}(|U_i(\beta)|)) f(U_i(\beta), Z_i) \quad (12)$$

instead of $g_n^{(WGMM)}(\mathbf{V}, \beta)$, etc.. Further, denote for any $i \in \mathcal{N}$

$$h(U_i(\beta), Z_i) = w(F_\beta(|U_i(\beta)|)) f(U_i(\beta), Z_i). \quad (13)$$

Finally, for any $\beta \in R^p$ and any $\delta > 0$ put $\mathcal{B}(\beta, \delta) = \{\alpha \in R^p : \|\alpha - \beta\| < \delta\}$. Now, following [3] let us recall:

Definition 3.1. For any $\beta \in R^p$ and any $\delta > 0$ put

$$\epsilon(\beta, \delta) = \sup \left\{ \left\| h(U_1(\alpha), Z_1) - h(U_1(\beta), Z_1) \right\| : \alpha \in \mathcal{B}(\beta, \delta) \cap \mathcal{S} \right\}. \quad (14)$$

The function $h(U_1(\beta), Z_1)$ (see (13)) is said to be ℓ -moment-continuous ($\ell \in \mathcal{N}$) at β if

$$\lim_{\delta \rightarrow 0_+} \mathbb{E} \epsilon^\ell(\beta, \delta) = 0.$$

We'll need the following assertion:

Assertion 3.2. The first-moment-continuity of $h(U_1(\beta), Z_1)$ implies continuity of $\mathbb{E}h(U_1(\beta), Z_1)$ in β for all $\beta \in \mathcal{S}$.

Proof: Fix $\varepsilon > 0$ and $\beta \in \mathcal{S}$. Then find $\delta > 0$ so that for all $\alpha \in \mathcal{S}$, $\|\alpha - \beta\| < \delta$ we have $\mathbb{E} \epsilon(\beta, \delta) < \varepsilon$. Then for all $\alpha \in \mathcal{S}$, $\|\alpha - \beta\| < \delta$

$$\|\mathbb{E}h(U_1(\alpha), Z_1) - \mathbb{E}h(U_1(\beta), Z_1)\| \leq \mathbb{E}\epsilon(\beta, \delta) < \varepsilon.$$

C1 Parameter space \mathcal{S} is compact with inner point $\beta^0 = 0$. For any $\beta \in \mathcal{S}$:

- the sequence $\{(Z'_i, U_i(\beta))'\}_{i=1}^\infty$ is sequence of independent and identically distributed $(q+1)$ -dimensional random vectors³ (i. i. d. r. v. 's) with

$$\mathbb{E} \sup_{\beta \in \mathcal{S}} \|f(U_1(\beta), Z_1)\|^2 < \infty, \quad (15)$$

³Remember, we have assumed $k = 1$.

- function $h(U_1(\beta), Z_1)$ is first-moment-continuous,
- the matrix a is of the full row rank.

Remark 3.3. The assumption $\mathbb{E} \sup_{\beta \in \mathcal{S}} \|f(U_1(\beta), Z_1)\| < \infty$ seems to be rather strong. Recalling (2), let's consider estimating coefficients of linear regression by means of "instrumental variables", i. e. by solution of normal equations

$$Z_i (Y_i - X_i' \beta) = 0$$

where $\{(X_i', e_i)'\}_{i=1}^{\infty}$ and $\{Z_i\}_{i=1}^{\infty}$ are sequences of explanatory variables, error terms and of instruments, respectively. Then we have (remember $\beta^0 = 0$)

$$f(U_1(\beta), Z_1) = (Y_1 - X_1' \beta^0) \cdot Z_1 + X_1' (\beta^0 - \beta) \cdot Z_1 = Z_1 (e_1 - X_1' \beta),$$

i. e. $\|f(U_1(\beta), Z_1)\|^2 = \|Z_1 (e_1 - X_1' \beta)\|^2$. Then

$$\mathbb{E} \sup_{\beta \in \mathcal{S}} \|Z_i (e_i - X_i' \beta)\|^2 \leq \mathbb{E} \|Z_i e_i\|^2 + \mathbb{E} \left\{ \|Z_i X_i'\|^2 \right\} \cdot \sup_{\beta \in \mathcal{S}} \|\beta\|^2$$

and under usual assumptions on explanatory variables, error terms and instruments, compactness of \mathcal{S} then implies (15).

Remark 3.4. As a is assumed to be of the full rank, A is positive definite and symmetric and hence it has all eigenvalues real and positive. Denote by λ_A the largest one.

We'll need also the following assertion:

Assertion 3.5. Under **Conditions C1** we have

$$\sup_{u \in \mathbb{R}^+, \beta \in \mathbb{R}^p} \sqrt{n} \left| F_{\beta}^{(n)}(u) - F_{\beta}(u) \right| = \mathcal{O}_p(1). \quad (16)$$

Proof runs along the same lines as the proof of Corollary 14 in [6]. Although the proof in [6] was seemingly established (only) for regression scheme, it utilized only the fact that the sequence of residuals $\{U_i(\beta)\}_{i=1}^{\infty}$ is a sequence of i.i.d. r.v.'s. Q.E.D.

To be able to prove consistency of $\hat{\beta}^{(WGMM, n)}$ we need an identification condition (see also [3], assumptions of Theorem 2.1 and compare it with **C3** in [7, 8]). Let us denote

$$\kappa_0 = \inf_{\beta \in \mathcal{S}} \mathbb{E}'[h(U_1(\beta), Z_1)] \cdot A \cdot \mathbb{E}[h(U_1(\beta), Z_1)].$$

Remark 3.6. Due to the compactness of \mathcal{S} the infimum is attained at least one point of \mathcal{S} .

C2 There is the only $\beta \in R^p$, namely β^0 , for which

$$\mathbb{E}'[h(U_1(\beta^0), Z_1)] \cdot A \cdot \mathbb{E}[h(U_1(\beta^0), Z_1)] = \kappa_0.$$

Assertion 3.7. Let $\gamma_0 = \sup\{r : \mathcal{B}(0, r) \subset \mathcal{S}\}$. Let $0 < \gamma < \gamma_0$ and consider compact set $\mathcal{S}(\gamma) = \mathcal{S} \setminus \mathcal{B}(0, \gamma)$. Then, under **C1** and **C2**, there are $\kappa_\gamma > \kappa_0$ and $\beta_\gamma \in \mathcal{S}(\gamma)$ such that

$$\begin{aligned} \inf_{\beta \in \mathcal{S}(\gamma)} \mathbb{E}'[h(U_1(\beta), Z_1)] \cdot A \cdot \mathbb{E}[h(U_1(\beta), Z_1)] \\ = \mathbb{E}'[h(U_1(\beta_\gamma), Z_1)] \cdot A \cdot \mathbb{E}[h(U_1(\beta_\gamma), Z_1)] = \kappa_\gamma. \end{aligned} \quad (17)$$

Proof: Put $\psi(U_1(\beta), Z_1) = \{\mathbb{E}[h(U_1(\beta), Z_1)]\}' \cdot A \cdot \mathbb{E}[h(U_1(\beta), Z_1)]$. Then definition of infimum implies that there is a sequence $\{\beta_n\}_{n=1}^\infty$ so that

$$\lim_{n \rightarrow \infty} \psi(U_1(\beta_n), Z_1) = \inf_{\beta \in \mathcal{S}(\gamma)} \psi(U_1(\beta), Z_1)$$

and due to compactness of $\mathcal{S}(\gamma)$ we can select a convergent subsequence $\{\beta_{n_k}\}_{k=1}^\infty$ so that $\lim_{k \rightarrow \infty} \beta_{n_k} = \beta_\gamma \in \mathcal{S}(\gamma)$ and from Assertion 3.2

$$\lim_{k \rightarrow \infty} \psi(U_1(\beta_{n_k}), Z_1) = \psi(U_1(\beta_\gamma), Z_1) = \kappa_\gamma.$$

Then from **C2** immediately follows that $\kappa_\gamma > \kappa_0$ otherwise we would have two solutions of (5). Q.E.D.

Recalling (13), denote $h_n(U_i(\beta), Z_i) = w(F_\beta^{(n)}(|U_i(\beta)|))f(U_i(\beta), Z_i)$ and

$$\psi_n(\mathbf{U}(\beta), \mathbf{Z}) = \left[\frac{1}{n} \sum_{i=1}^n h_n(U_i(\beta), Z_i) \right]' \cdot A \cdot \left[\frac{1}{n} \sum_{i=1}^n h_n(U_i(\beta), Z_i) \right].$$

C3 Weight function $w : [0, 1] \rightarrow [0, 1]$ is absolutely continuous and nonincreasing, with the derivative $w'(\alpha)$ bounded in absolute value by L , $w(0) = 1$.

Lemma 3.8. Under **C1**, **C2** and **C3** for any $\varepsilon > 0$ and any $\gamma \in (0, \gamma_0)$ there are $\Delta_\gamma > 0$ and $n_{\varepsilon, \gamma} \in \mathcal{N}$ such that for any $n > n_{\varepsilon, \gamma}$

$$P \left(\left\{ \omega \in \Omega : \inf_{\beta \in \mathcal{S}(\gamma)} \psi_n(\mathbf{U}(\beta), \mathbf{Z}) > \kappa_0 + \Delta_\gamma \right\} \right) > 1 - \varepsilon. \quad (18)$$

Proof : Fix $\varepsilon > 0$, $\gamma \in (0, \gamma_0)$ and put $\Delta_\gamma = \frac{\kappa_\gamma}{6}$ (for κ_γ see (17)). Denote $\mathbb{E} \sup_{\beta \in \mathcal{S}} \|f(U_1(\beta), Z_1)\|^2 + 1 = \xi < \infty$ and put $\eta_\gamma = \frac{\varepsilon \cdot \Delta_\gamma}{12 \cdot \lambda_A \cdot L \cdot \xi}$. Employing Assertion 3.5 find $n^{(1)} \in \mathcal{N}$ so that for any $n > n^{(1)}$ for

$$B_n = \left\{ \omega \in \Omega : \sup_{u \in \mathbf{R}^+, \beta \in \mathbf{R}^p} \left| F_\beta^{(n)}(u) - F_\beta(u) \right| > \eta_\gamma \right\}$$

we have $P(B_n) < \frac{\varepsilon}{6}$. Denote

$$C_n = \left\{ \omega \in \Omega : \sup_{\beta \in \mathcal{S}(\gamma)} \left| \psi_n(\mathbf{U}(\beta), \mathbf{Z}) - \psi(\mathbf{U}(\beta), \mathbf{Z}) \right| > \Delta_\gamma \right\} \quad (19)$$

and recall that

$$P(C_n) = P(B_n \cap C_n) + P(B_n^c \cap C_n) \leq P(B_n) + P(B_n^c \cap C_n).$$

Then utilizing Chebysev's inequality for positive r. v. (write U_i instead of $U_i(\beta)$ for a while)

$$\begin{aligned} P(B_n^c \cap C_n) &\leq \frac{1}{\Delta_\gamma} \int_{\Omega} \sup_{\beta \in \mathcal{S}(\gamma)} \left| \frac{1}{n} \sum_{i=1}^n [(h_n(U_i, Z_i) - h(U_i, Z_i)) \times \right. \\ &\quad \left. \times A \cdot h_n(U_i, Z_i) + h(U_i, Z_i) \cdot A \cdot (h_n(U_i, Z_i) - h(U_i, Z_i))] \right| dP(\omega) \leq \\ &\leq \frac{2 \cdot L \cdot \eta_\gamma \cdot \lambda_A}{\Delta_\gamma} \int_{\Omega} \sup_{\beta \in \mathcal{S}(\gamma)} \|f(U_i, Z_i)\|^2 dP(\omega) < \frac{2 \cdot L \cdot \eta_\gamma \cdot \xi \cdot \lambda_A}{\Delta_\gamma} < \frac{\varepsilon}{6} \end{aligned} \quad (20)$$

and hence $P(C_n) < \frac{\varepsilon}{3}$. Utilizing moment-continuity, for any $\beta \in \mathcal{S}(\gamma)$ find $\delta_\beta > 0$ so that

$$\forall (\alpha \in \mathcal{S}(\gamma), \|\alpha - \beta\| < \delta_\beta) \quad \mathbb{E} \epsilon(\beta, \delta_\beta) < \frac{1}{2} \sqrt{\Delta_\gamma} \quad (21)$$

(for $\epsilon(\beta, \delta)$ see (14)). Then $\mathcal{S}(\gamma) \subset \bigcup_{\beta \in \mathcal{S}(\gamma)} \mathcal{B}(\beta, \delta_\beta)$ and the compactness of $\mathcal{S}(\gamma)$ implies that there are $\beta_1, \beta_2, \dots, \beta_K$ so that $\mathcal{S}(\gamma) \subset \bigcup_{k=1}^K \mathcal{B}(\beta_k, \delta_{\beta_k})$. Now for any $k \in \{1, 2, \dots, K\}$ define sequence of i.i.d. r.v.'s by

$$\begin{aligned} &\{\epsilon_i(\beta_k, \delta_{\beta_k})\}_{i=1}^\infty \\ &= \left\{ \sup \left\{ \left\| h(U_i(\alpha), Z_i) - h(U_i(\beta_k), Z_i) \right\| : \alpha \in \mathcal{B}(\beta_k, \delta_{\beta_k}) \cap \mathcal{S} \right\} \right\}_{i=1}^\infty \end{aligned}$$

and for all $k = 1, 2, \dots, K$ find $n_k \in \mathcal{N}$ so that for any $n > n_k$ we have

$$P \left(\left\{ \omega \in \Omega : \left\| \frac{1}{n} \sum_{j=1}^n \epsilon_j(\beta_k, \delta_{\beta_k}) - \mathbb{E} \epsilon_1(\beta_k, \delta_{\beta_k}) \right\| > \frac{1}{2} \sqrt{\Delta_\gamma} \right\} \right) < \frac{\varepsilon}{3K}.$$

Due to (21), it immediately implies that for any $n > n_k$

$$P \left(\left\{ \omega \in \Omega : \left\| \frac{1}{n} \sum_{j=1}^n \epsilon_j(\beta_k, \delta_{\beta_k}) \right\| < \Delta_\gamma \right\} \right) > 1 - \frac{\varepsilon}{3K}.$$

Putting $n^{(2)} = \max_{k=1,2,\dots,K} n_k$ and

$$D_n = \bigcap_{k=1}^K \left\{ \omega \in \Omega : \left\| \frac{1}{n} \sum_{j=1}^n \epsilon_j(\beta_k, \delta_{\beta_k}) \right\| < \sqrt{\Delta_\gamma} \right\}, \tag{22}$$

we have for any $n > n^{(2)}$ $P(D_n) > 1 - \frac{\varepsilon}{3}$. Further, for all $k = 1, 2, \dots, K$ find $m_k \in \mathcal{N}$ so that for any $n > m_k$ we have

$$P \left(\left\{ \omega \in \Omega : \left\| \frac{1}{n} \sum_{i=1}^n h(U_i(\beta_k), Z_i) - \mathbb{E}h(U_1(\beta_k), Z_1) \right\| > \sqrt{\Delta_\gamma} \right\} \right) < \frac{\varepsilon}{3K}.$$

Putting $n^{(3)} = \max_{k=1,2,\dots,K} m_k$ and

$$E_n = \bigcap_{k=1}^K \left\{ \omega \in \Omega : \left\| \frac{1}{n} \sum_{i=1}^n h(U_i(\beta_k), Z_i) - \mathbb{E}h(U_1(\beta_k), Z_1) \right\| < \sqrt{\Delta_\gamma} \right\}, \tag{23}$$

we have for any $n > n^{(3)}$ $P(E_n) > 1 - \frac{\varepsilon}{3}$. It's straightforward that $P(C_n^c \cap D_n \cap E_n) > 1 - \varepsilon$. Put $n_{\varepsilon,\gamma} = \max\{n^{(1)}, n^{(2)}, n^{(3)}\}$ and consider any $\alpha \in \mathcal{S}(\gamma)$, $n > n_{\varepsilon,\gamma}$ and $\omega \in C_n^c \cap D_n \cap E_n$. As $\alpha \in \mathcal{S}(\gamma)$, there is $k_0 \in \{1, 2, \dots, K\}$ so that $\alpha \in \mathcal{B}(\beta_{k_0}, \delta_{\beta_{k_0}})$. It means that (see subsequently (19), (22) and (23))

$$\left\| \frac{1}{n} \sum_{i=1}^n [h_n(U_i(\alpha), Z_i) - h(U_i(\alpha), Z_i)] \right\| < \sqrt{\Delta_\gamma}, \tag{24}$$

$$\left\| \frac{1}{n} \sum_{i=1}^n [h(U_i(\alpha), Z_i) - h(U_i(\beta_{k_0}), Z_i)] \right\| < \sqrt{\Delta_\gamma} \tag{25}$$

and

$$\left\| \frac{1}{n} \sum_{i=1}^n h(U_i(\beta_{k_0}), Z_i) - \mathbb{E}h(U_i(\beta_{k_0}), Z_i) \right\| < \sqrt{\Delta_\gamma}. \tag{26}$$

As $\Delta = \frac{\kappa_\gamma}{6}$ and taking into account Assertion 3.7, we have for any $\alpha \in \mathcal{S}(\gamma)$, $n > n_{\varepsilon,\gamma}$ and $\omega \in C_n^c \cap D_n \cap E_n$

$$\frac{1}{n} \sum_{i=1}^n h'_n(U_i(\alpha), Z_i) \cdot A \cdot h_n(U_i(\alpha), Z_i) > \frac{\kappa_\gamma}{2} + \kappa_0.$$

Q.E.D.

Theorem 3.9. *Under C1, C2 and C3 $\hat{\beta}^{(WGMM,n)}$ is (weakly) consistent.*

4 Conclusions

Paper proposes possible robustification of GMM under the simplest standard conditions. The next step is to be to consider heteroscedasticity of disturbances (however as the difference between the proof of Kolmogorov-Smirnoff result for regression scheme under homoscedasticity, on one hand, and under the heteroscedasticity on the other, for the case of heteroscedasticity we need also the parameter space to be separable metric space and the technicalities of the proof are complicated, see [9]) or to assume autocorrelation in the sequence of random variables $\{V_i\}_{i=1}^{\infty}$.

References

- [1] Cragg J.G. (1983). *More efficient estimation in the presence of heteroscedasticity of unknown form*. *Econometrica*, **51**, 751–763.
- [2] Hájek J., Šidák Z. (1967). *Theory of Rank Test*. Academic Press, New York.
- [3] Hansen L.P. (1982). *Large sample properties of generalized method of moments estimators*. *Econometrica*, **50**, 4, 1029–1054.
- [4] Judge G.G., Griffiths W.E., Hill R.C., Lütkepohl H., Lee T.C. (1982). *Introduction to the Theory and Practice of Econometrics*. New York: J.Wiley & Sons.
- [5] Víšek, J.Á. (2000). *Regression with high breakdown point*. In: J. Antoch & G. Dohnal (Eds.): *Robust 2000*. MatFyz Press, Prague, 324–356.
- [6] Víšek, J.Á. (2006). *Kolmogorov-Smirnov statistics in multiple regression*. In: J. Antoch & G. Dohnal (Eds.): *Robust 2006*. MatFyz Press, Prague, 367–374.
- [7] Víšek J.Á. (2008a). *Consistency of the instrumental weighted variables*. To appear in the *Annals of the Institute of Statistical Mathematics*.
- [8] Víšek J.Á. (2008b). *Consistency of the least weighted squares*. Submitted to the *Kybernetika*.
- [9] Víšek, J.Á. (2008c). *Empirical distribution function under heteroscedasticity*. Submitted to *Statistics*.
- [10] Wooldridge, J. M. (2001). *Applications of generalized method of moments estimation*. *Journal of Economic Perspective*, Vol. 15, no 4, 87 - 100.

Acknowledgement: Research was supported by grant of GA ČR number 402/06/0408. We are grateful to the anonymous referee for suggestions which improved the quality of paper.

Address: J.Á. Víšek, UK FSV & AV ČR ÚTIA, Smetanovo nábřeží 6, 110 01 Praha 1

E-mail: visek@mbox.fsv.cuni.cz

ALGORITMUS VÝPOČTU PŘIBLIŽNÝCH KONFIDENČNÝCH INTERVALOV PARAMETRA POLOHY Z DIGITALIZOVANÝCH MERANÍ

Gejza Wimmer, Jr.

Klíčové slová: Přibližný konfidenční interval, fiduciální distribúcia, parameter polohy, algoritmus.

Abstrakt: Príspevok sa zaoberá algoritmom pre výpočet približných konfidenčných intervalov parametra polohy z digitalizovaných meraní. Je prezentovaný všeobecný popis algoritmu pre generovanie náhodného výberu z fiduciálnej distribúcie, ako aj podrobnejšie znázornenie vytvorenia množiny možných hodnôt parametrov (μ, σ) tak, aby táto bola neprázdna. Nasledujúcim krokom je vytvorenie výberového mechanizmu, ktorý vygeneruje jedinou hodnotu (μ, σ) z polygónu všetkých možných hodnôt parametrov. Výsledky sú zhrnuté v simulačnej štúdii, ktorej cieľom bolo porovnať základné štatistické vlastnosti (pravdepodobnosť pokrytia a očakávanú dĺžku) približných konfidenčných intervalov pre parameter μ na základe malého výberu z digitalizovaných meraní.

Abstract: An algorithm for a construction of approximate confidence intervals for the location parameter based on digitized measurements is presented. An algorithm for generating random samples from a fiducial distribution is introduced and a more detail description of generating a non-empty set of all possible values of the parameters (μ, σ) is given. The next step is a creation of a sampling mechanism that generates one value of (μ, σ) from a polygon that contains all possible values of the parameters. The results are described in a simulation study where the main statistical characteristics (coverage probability and average length) of the approximate confidence intervals for μ are presented.

1 Úvod

Nech skutočné výsledky merania sú v digitálnom tvare v dôsledku známej a ohraničenej rozlíšiteľnosti meracieho prístroja δ . V [4] sú uvedené rôzne spôsoby zostrojenia konfidenčných intervalov na základe ML prístupu, Studentovho prístupu, Wilinkovho prístupu pomocou modifikovaného odhadu disperzie, ako aj (modifikovaný) fiduciálny prístup, motiváciu problému a ďalšie podrobnosti možno nájsť v [1, 2, 3]. V tomto príspevku uvidíme všeobecný popis algoritmu pre generovanie náhodného výberu z fiduciálnej

distribúcie, ktorého verzia bola implementovaná v prostredí MATLAB a následne v R (S-plus), jednotlivé kroky vedúce k zostrojeniu približných konfidenčných intervalov parametera polohy pomocou posledného zmieného prístupu, ako aj simulačné štúdiu, v ktorej sú porovnané základné štatistické vlastnosti vyššie uvedených štyroch konfidenčných intervalov pre parameter μ .

Algoritmus sa zakladá na združenej fiduciálnej distribučnej funkcii parametrov (μ, σ) (závisí od G, Q , a V), ktorá je daná ako

$$F_{(\mu, \sigma)}((\mu^*, \sigma^*)) = \Pr(V(Q(x, Z^*)) \leq (\mu^*, \sigma^*) | Z^* \in S(x)), \quad (1)$$

kde (μ^*, σ^*) je pevne zvolený vektor hodnôt a $Z^* = (Z_1^*, Z_2^*, \dots, Z_n^*)' \sim N(0, I_n)$. Tu $X = G(\xi, Z)$ je známy systém štruktúrálnych rovníc (dáta generujúci mechanizmus, spája náhodný vektor pozorovaní X s vektorom parametrov $\xi \in \Xi$), $Q(X, Z)$ reprezentuje 'inverznú funkciu' ku G (pre ľubovoľnú pevnú hodnotu x a z tvorí množinu všetkých možných parametrov) a $V(\cdot)$ je náhodný mechanizmus voľby parametra (μ, σ) z uzáveru množiny možných parametrov $Q(x, z)$, pričom $V(\cdot)$ má rovnomerné rozdelenie nad $Q(x, z)$. Navyše, ak $z \in S(x)$, kde

$$S(x) = \{(z_1, \dots, z_n) : \exists(\mu, \sigma), \text{ že} \\ x_i - \delta/2 \leq \mu + \sigma z_i < x_i + \delta/2, i = 1, 2, \dots, n\}, \quad (2)$$

prícom $x_i = \delta k_i = \delta[(\mu + \sigma Z_i)/\delta + 0.5]$, tak množina parametrov $Q(x, z)$ vytvára polygón.

2 Algoritmus na generovanie náhodného výberu z fiduciálnej distribúcie

Napriek tomu, že explicitný tvar rozdelenia (1) nie je známy, generovanie náhodného výberu parametrov (μ, σ) možno urobiť priamo podľa (1). Na to, aby sa počas generovania realizácii náhodného vektora Z^* nevyskytli prípady, keď $Q(x, Z^*) = \emptyset$ (t.j. aby nebolo nutné v každom kroku algoritmu overovať podmienku $Z^* \in S(x)$) s výhodou možno využiť metódy generovania založené na MCMC (Markov Chain Monte Carlo). Detailný popis algoritmu možno nájsť v práci Hannig *et al.* (2007), pričom je založený na nasledovnom postupe:

1. Ak je realizácia vektora meraní $x = (x_1, x_2, \dots, x_n)$ konzistentná s modelom (daným výberovým mechanizmom G), potom musí existovať najmenej jeden parameter, napr. $\xi^{(0)} = (\mu^{(0)}, \sigma^{(0)})$, že $x = G(\xi^{(0)}, z)$ pre nejakú hodnotu z , povedzme $z^{(0)}$, teda platí $z^{(0)} \in S(x)$.
2. Pre ľubovoľnú hodnotu vektora chýb $z \in S(x)$, povedzme $z^{(k-1)}$, algoritmus nájde polygón $Q(x, z^{(k-1)})$ s nekonečným počtom možných hodnôt parametrov (μ, σ) .

3. Pomocou výberového mechanizmu $V(\cdot)$ algoritmus vygeneruje jediná hodnotu parametra (μ, σ) z polygónu $Q(x, z^{(k-1)})$, napr. $(\mu^{(k)}, \sigma^{(k)})$, ktorý spĺňa podmienky $\alpha_i \leq \mu^{(k)} + \sigma^{(k)} z_i < \beta_i$ pre všetky z_i , také že $(\alpha_i - \mu^{(k)}) / \sigma^{(k)} \leq z_i < (\beta_i - \mu^{(k)}) / \sigma^{(k)}$, $i = 1, \dots, n$.
4. Z týchto možných hodnôt z_i algoritmus náhodne vyberie novú verziu vektora chýb z , napr. $z^{(k)} = (z_1^{(k)}, \dots, z_n^{(k)})$. Algoritmus potom pokračuje opakovaním krokov 2-4, pokiaľ sa nedosiahne požadovaný počet realizácií náhodného výberu parametrov (μ, σ) z fiduciálnej distribúcie.
5. Na vylepšenie rýchlosti konvergenzie k cieľovej fiduciálnej distribúcii je potrebné zaviesť nasledovný zmiešavací krok (mixing step): Pre dané $z = (z_1, \dots, z_n)$, $z \in S(x)$, algoritmus v každom kroku náhodne posunie z do $\tilde{z} = (\tilde{z}_1, \dots, \tilde{z}_n)$,

$$\tilde{z}_i = \frac{1}{n} U^* + \frac{\sqrt{W^*}}{\sum_{i=1}^n (z_i - \bar{z})^2} (z_i - \bar{z}), \quad i = 1, \dots, n, \quad (3)$$

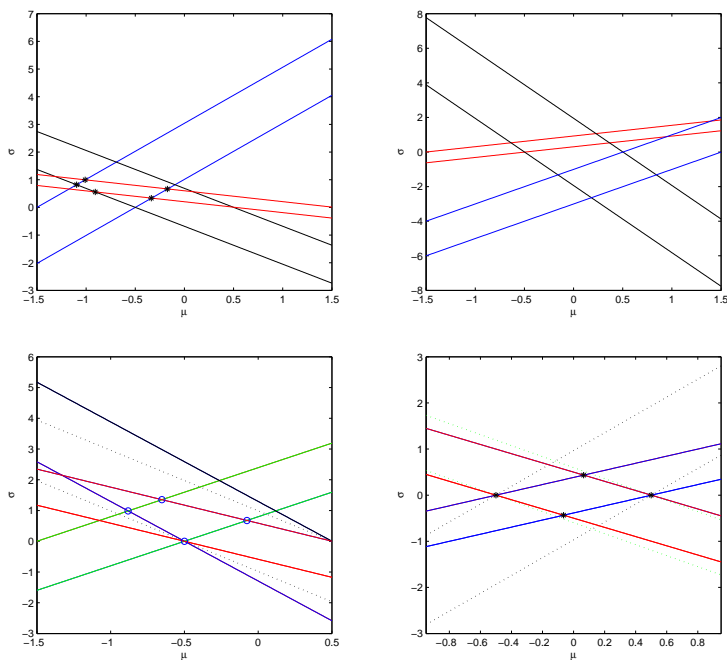
kde $\bar{z} = \frac{1}{n} \sum_{i=1}^n z_i$. $U^* \sim N(0, 1)$ a $W^* \sim \chi_{(n-1)}^2$ sú nezávislé náhodné premenné, spoločné pre všetky $i = 1, \dots, n$. Pri tomto postupe vždy platí, že $\tilde{z} \in S(x)$.

3 Jednotlivé časti algoritmu

Uvedený algoritmus bol zrealizovaný v prostredí (štatistických) programovacích balíkov MATLAB a R. Úvodným krokom je načítanie vstupných údajov, ako sú získané (resp. vygenerované) merania a rôzne konštanty (rozlišovacia schopnosť meracieho prístroja, rozsah náhodného výberu zo združenej fiduciálnej distribúcie, hladinu významnosti α_v a pod.). Ďalšiu časť tvorí „Inicializácia“, kde sa na základe meraní a rozlíšenia meracieho prístroja určí dolná a horná hranica $\alpha_i = x_i - \delta/2$ a $\beta_i = x_i + \delta/2$, $i = 1, 2, \dots, n$ z (2) a vygeneruje sa vektor chýb $z^* \in S(x)$ tak, aby existoval polygón $Q(x, z^*)$ a to nasledovne. Najprv sa vygenerujú „nové merania“ $\mathbf{Y}^* = (Y_1^*, \dots, Y_n^*)'$ z rovnomerného rozdelenia na intervale (α, β) . Následne sa spočíta $\sigma^* = S_{Y^*} \sqrt{(n-1)/W}$, kde $W \sim \chi_{(n-1)}^2$ a $\mu^* = \bar{Y}^* + (\sigma^*/\sqrt{n})Z$, kde $Z \sim N(0, 1)$. Tu \bar{Y}^* , resp. $S_{Y^*}^2$ je výberová stredná hodnota, resp. výberová disperzia, spočítaná z hodnôt (Y_1^*, \dots, Y_n^*) . Nakoniec sa vygeneruje vektor chýb $\mathbf{z}^* = (\mathbf{Y}^* - \mu^*) / \sigma^*$.

Na obrázku 1 sú znázornené príklady vygenerovaného vektora chýb z^* v prípade, že existuje polygón $Q(x, z^*)$, ako aj v opačnej situácii, keď tento polygón nevznikne. Potom nasleduje redukcia počtu nerovnic, ktorých prienik tvorí polygón $Q(x, z^*)$ na najmenší potrebný počet $n^* \leq n$ nerovnic. Táto časť je vhodná najmä v prípade, ak obdržíme opakované hodnoty meraní.

Následne sa vyhladá spoločný prienik zredukovaných nerovnic $\alpha_i \leq \mu + \sigma z_i^* < \beta_i$, $i = 1, \dots, n^*$, tzn. určí sa polygón $Q(x, z^*)$ z ktorého pomocou



Obrázek 1: Obrázok naľavo hore znázorňuje nerovnice $\alpha_i \leq \mu + \sigma z_i^* < \beta_i$, $i = 1, 2, 3$, s vektorom vhodne vygenerovaných chýb, podľa popisu v časti *Inicializácia*, [pre napozorované hodnoty $(y_1, y_2, y_3) = (-1, 0, 1)$ a vygenerované hodnoty chýb $(z_1, z_2, z_3) = (-0.49, 0.73, 2.52)$]. Obrázok napravo hore znázorňuje nerovnice $\alpha_i \leq \mu + \sigma z_i^* < \beta_i$, $i = 1, 2, 3$, s vektorom nevhodne vygenerovaných chýb z rozdelenia $N(0, 1)$, [pre napozorované hodnoty $(y_1, y_2, y_3) = (-1, 0, 1)$ a vygenerované hodnoty chýb $(z_1, z_2, z_3) = (-1.60, 0.26, -0.50)$]. Obrázky dolnej časti znázorňujú redukciu počtu potrebných nerovníc. Bodkovanou čiarou sú znázornené oblasti definované zredukovanými nerovnicami.

výberového mechanizmu $V(\cdot)$ vygenerujeme jeden bod z jeho vnútra, tzn. získame jednu hodnotu (μ_g, σ_g) . Záverečným krokom je vygenerovanie nového vektora chýb tak, aby $z_n \in S(x)$. Tu sa pri vstupných hodnotách (μ_g, σ_g) z predchádzajúceho kroku spočíta hodnota Φ distribučnej funkcie $N(0, 1)$ v bodoch $(\alpha - \mu_g)/\sigma_g$ a $(\beta - \mu_g)/\sigma_g$, tzn. hodnoty $r_1 = \Phi((\alpha - \mu_g)/\sigma_g)$ a $r_2 = \Phi((\beta - \mu_g)/\sigma_g)$. Pomocou nich potom vygenerujeme nové chyby z_n ako $z_n = \Phi^{-1}(r_1 + R(r_2 - r_1))$, kde R je vygenerované z rovnomerného $(0, 1)$ rozdelenia. Z týchto na záver pomocou zmiešavacieho kroku určíme hodnotu \tilde{z}_n . Celý postup sa potom opakuje podľa potreby, vzhľadom na to, koľko hodnôt z fiduciálnej distribúcie je požadovaných.

σ	μ	Fiducial	Willink	Student	ML
0.1	0.0	1.0000 (1.00)	1.0000 (1.60)	1.0000 (0.00)	1.0000 (0.00)
	0.1	1.0000 (1.00)	1.0000 (1.60)	0.0001 (0.00)	0.0000 (0.00)
	0.2	1.0000 (1.00)	1.0000 (1.60)	0.0064 (0.01)	0.0000 (0.00)
	0.3	1.0000 (1.01)	1.0000 (1.55)	0.1103 (0.12)	0.0000 (0.00)
	0.4	1.0000 (1.10)	1.0000 (1.36)	0.5834 (0.69)	0.5533 (0.10)
0.3	0.0	1.0000 (1.03)	1.0000 (1.62)	1.0000 (0.24)	0.8039 (0.01)
	0.1	1.0000 (1.04)	1.0000 (1.51)	0.2777 (0.32)	0.0090 (0.01)
	0.2	0.9986 (1.07)	0.9986 (1.41)	0.4592 (0.54)	0.0050 (0.01)
	0.3	0.9997 (1.13)	1.0000 (1.33)	0.6910 (0.84)	0.0018 (0.00)
	0.4	0.9955 (1.19)	1.0000 (1.30)	0.8750 (1.10)	0.6515 (0.13)
0.5	0.0	0.9999 (1.22)	1.0000 (1.30)	0.9332 (1.19)	0.6279 (0.15)
	0.0	0.9954 (1.39)	0.9970 (1.85)	0.9954 (1.25)	0.4656 (0.38)
	0.1	0.9857 (1.37)	0.9922 (1.69)	0.8446 (1.25)	0.2944 (0.36)
	0.2	0.9773 (1.37)	0.9805 (1.57)	0.8688 (1.28)	0.2697 (0.34)
	0.3	0.9880 (1.38)	0.9998 (1.52)	0.9088 (1.33)	0.2350 (0.30)
1.0	0.0	0.9718 (1.36)	0.9982 (1.47)	0.9235 (1.34)	0.6424 (0.35)
	0.0	0.9623 (2.37)	0.9697 (2.55)	0.9623 (2.37)	0.7475 (1.28)
	0.1	0.9438 (2.34)	0.9628 (2.50)	0.9356 (2.34)	0.7239 (1.26)
	0.2	0.9583 (2.36)	0.9771 (2.50)	0.9528 (2.37)	0.7419 (1.29)
	0.3	0.9561 (2.38)	0.9838 (2.50)	0.9509 (2.39)	0.7476 (1.30)
	0.4	0.9451 (2.34)	0.9823 (2.46)	0.9385 (2.35)	0.7965 (1.28)
	0.5	0.9519 (2.37)	0.9764 (2.45)	0.9447 (2.38)	0.8130 (1.29)

Tabulka 1: Pravdepodobnosť pokrytia a priemerná dĺžka približných 95% konfidenčných intervalov parametra μ , pre rozsah výberu $n = 5$.

4 Simulačná štúdia

Cieľom simulačnej štúdie bolo porovnať základné štatistické vlastnosti (pravdepodobnosť pokrytia a očakávanú dĺžku) približných konfidenčných intervalov pre parameter μ na základe malého výberu z digitalizovaných meraní. V simulačnej štúdií boli zahrnuté približné konfidenčné intervaly CI_{μ}^{ML} , CI_{μ}^{St} , CI_{μ}^W a CI_{μ}^{FD} pre parameter polohy μ (t.j. intervaly určené metódou maximálnej vierohodnosti, klasický Studentov konfidenčný interval, interval navrhnutý Willinkom a interval určený z fiduciálnej distribučnej funkcie), pozri [1, 2, 3].

Uvažovali sme malé rozsahy výberu n , $n \in \{5, 10, 30\}$, relatívne malé rozptyly chýb σ^2 (vzhľadom k rozlíšeniu prístroja $\delta > 0$), $\sigma \in \{0.1\delta, 0.3\delta\}$,

σ	μ	Fiducial	Willink	Student	ML
0.1	0.0	1.0000 (1.00)	1.0000 (1.31)	1.0000 (0.00)	1.0000 (0.00)
	0.1	1.0000 (1.00)	1.0000 (1.31)	0.0002 (0.00)	0.0000 (0.00)
	0.2	1.0000 (0.99)	1.0000 (1.30)	0.0144 (0.01)	0.0000 (0.00)
	0.3	1.0000 (0.93)	1.0000 (1.25)	0.2058 (0.10)	0.0000 (0.00)
	0.4	0.9999 (0.70)	0.9999 (0.95)	0.4865 (0.31)	0.8099 (0.09)
0.3	0.0	1.0000 (0.87)	1.0000 (1.28)	1.0000 (0.19)	0.6564 (0.02)
	0.1	0.9976 (0.83)	1.0000 (1.17)	0.4754 (0.24)	0.0304 (0.02)
	0.2	0.9907 (0.73)	0.9998 (1.03)	0.7069 (0.38)	0.0166 (0.01)
	0.3	0.9982 (0.67)	0.9989 (0.87)	0.9052 (0.55)	0.0062 (0.00)
	0.4	0.9948 (0.65)	0.9949 (0.77)	0.9008 (0.63)	0.7496 (0.10)
0.5	0.0	1.0000 (0.65)	0.9897 (0.72)	0.9776 (0.70)	0.8880 (0.15)
	0.0	0.9625 (0.80)	0.9817 (1.04)	0.9625 (0.76)	0.6902 (0.50)
	0.1	0.9353 (0.78)	0.9800 (0.94)	0.9478 (0.76)	0.6275 (0.47)
	0.2	0.9547 (0.78)	0.9816 (0.88)	0.9398 (0.77)	0.5543 (0.43)
	0.3	0.9713 (0.77)	0.9798 (0.84)	0.9648 (0.79)	0.4464 (0.35)
1.0	0.0	0.9670 (0.76)	0.9698 (0.82)	0.9472 (0.79)	0.6977 (0.35)
	0.0	0.9786 (0.77)	0.9852 (0.82)	0.9787 (0.81)	0.9057 (0.38)
	0.0	0.9497 (1.39)	0.9576 (1.41)	0.9524 (1.40)	0.8921 (1.09)
	0.1	0.9461 (1.39)	0.9578 (1.41)	0.9529 (1.40)	0.8895 (1.09)
	0.2	0.9454 (1.38)	0.9538 (1.40)	0.9497 (1.39)	0.8886 (1.08)
0.3	0.2	0.9524 (1.39)	0.9536 (1.40)	0.9517 (1.40)	0.8913 (1.09)
	0.4	0.9539 (1.39)	0.9565 (1.40)	0.9549 (1.40)	0.9025 (1.10)
	0.5	0.9527 (1.40)	0.9556 (1.41)	0.9447 (1.40)	0.9014 (1.09)

Tabulka 2: Pravdepodobnosť pokrytia a priemerná dĺžka približných 95% konfidenčných intervalov parametra μ , pre rozsah výberu $n = 10$.

$0.5\delta, 1.0\delta\}$, ako aj rôzne hodnoty parametra polohy μ v rozsahu jednej jednotky danej škály meracieho prístroja, $\mu \in \{0.0\delta, 0.1\delta, 0.2\delta, 0.3\delta, 0.4\delta, 0.5\delta\}$. Bez újmy na všeobecnosti sme uvažovali len s rozlíšením meracieho prístroja $\delta = 1$. Pre každú zvolenú kombináciu parametrov n, μ a σ sme vygenerovali realizácie digitalizovaných meraní (x_1, x_2, \dots, x_n) a na jej základe vypočítali hodnoty približných 95% konfidenčných intervalov pre parameter polohy μ : teda $CI_{\mu}^{ML}(1 - \alpha)$, $CI_{\mu}^{St}(1 - \alpha)$ a $CI_{\mu}^W(1 - \alpha)$ pre $\alpha = 0.05$. Na určenie konfidenčného intervalu $CI_{\mu}^{FD}(1 - \alpha)$ sme navyše generovali náhodný výber rozsahu $N = 10000$ z fiduciálnej distribúcie (1). Keďže rozdelenie náhodného vektora meraní (X_1, X_2, \dots, X_n) má diskretný charakter, realizácie meraní (x_1, x_2, \dots, x_n) viedli k relatívne malému počtu rôznych hodnôt, hlavne pre malé hodnoty σ . To umožnilo významne zredukovať výpočtové nároky na

σ	μ	Fiducial		Willink		Student		ML	
0.1	0.0	1.0000	(1.00)	1.0000	(1.18)	1.0000	(0.00)	0.9998	(0.00)
	0.1	1.0000	(1.00)	1.0000	(1.18)	0.0015	(0.00)	0.0000	(0.00)
	0.2	1.0000	(0.98)	1.0000	(1.18)	0.0000	(0.00)	0.0000	(0.00)
	0.3	1.0000	(0.78)	1.0000	(1.13)	0.0001	(0.00)	0.0000	(0.00)
	0.4	1.0000	(0.44)	1.0000	(0.81)	0.0913	(0.03)	0.7097	(0.04)
0.3	0.0	0.9708	(0.60)	1.0000	(0.95)	0.9950	(0.13)	0.4898	(0.08)
	0.1	0.9540	(0.54)	0.9306	(0.91)	0.7452	(0.14)	0.1711	(0.06)
	0.2	0.9881	(0.47)	0.9773	(0.86)	0.6912	(0.18)	0.0628	(0.02)
	0.3	0.9943	(0.39)	0.9912	(0.68)	0.7784	(0.25)	0.0138	(0.00)
	0.4	0.9885	(0.32)	0.9963	(0.45)	0.8559	(0.31)	0.3870	(0.03)
0.5	0.0	1.0000	(0.31)	0.9802	(0.37)	0.9591	(0.36)	0.9058	(0.09)
	0.1	0.9512	(0.41)	0.9646	(0.42)	0.9561	(0.40)	0.9264	(0.38)
	0.2	0.9431	(0.41)	0.9562	(0.41)	0.9534	(0.40)	0.9300	(0.38)
	0.3	0.9402	(0.41)	0.9516	(0.42)	0.9513	(0.41)	0.8864	(0.37)
	0.4	0.9421	(0.40)	0.9522	(0.42)	0.9512	(0.41)	0.8074	(0.34)
1.0	0.0	0.9454	(0.39)	0.9525	(0.42)	0.9479	(0.41)	0.7525	(0.31)
	0.1	0.9580	(0.39)	0.9528	(0.42)	0.9462	(0.41)	0.9290	(0.32)
	0.0	0.9508	(0.73)	0.9513	(0.74)	0.9513	(0.74)	0.9359	(0.68)
	0.1	0.9461	(0.73)	0.9473	(0.73)	0.9473	(0.73)	0.9324	(0.68)
	0.2	0.9470	(0.73)	0.9477	(0.73)	0.9477	(0.73)	0.9323	(0.68)
0.3	0.9470	(0.73)	0.9451	(0.73)	0.9451	(0.73)	0.9338	(0.68)	
0.4	0.9478	(0.73)	0.9472	(0.73)	0.9472	(0.73)	0.9325	(0.68)	
0.5	0.9522	(0.73)	0.9507	(0.73)	0.9507	(0.73)	0.9376	(0.68)	

Tabuľka 3: Pravdepodobnosť pokrytia a priemerná dĺžka približných 95% konfidenčných intervalov parametra μ , pre rozsah výberu $n = 30$.

určenie konfidenčného intervalu z fiduciálneho rozdelenia. Korektnosť uvažovaných konfidenčných intervalov sme posudzovali s ohľadom na kritérium $\Pr(\mu \in CI_\mu(1 - \alpha)) \geq 1 - \alpha$, kde μ je skutočná hodnota parametra.

Z výsledkov tejto simulačnej štúdie, ktoré sú uvedené v tabuľkách, sa dá usúdiť, že približný konfidenčný interval CI_μ^{ML} je pre uvažované situácie neadekvátny, hoci je teoreticky známe, že je asymptoticky správny. Ako sa ukázalo aj v iných situáciách pre malé hodnoty skutočného parametra σ , konfidenčný interval CI_μ^{ML} nemusí dosiahnuť stanovenú pravdepodobnosť pokrytia ani pre veľmi veľké rozsahy náhodného výberu z pozorovaní n . Približný konfidenčný interval CI_μ^{St} je korektný (správny) aj pre malé rozsahy výberov n , pokiaľ hodnota skutočného parametra σ je blízka danému rozlíšeniu prístroja δ (alebo väčšia). Približné konfidenčné intervaly CI_μ^W

a CI_{μ}^{FD} boli korektné takmer pre všetky uvažované situácie. Interval CI_{μ}^W dosahoval mierne väčšie hodnoty priemerných dĺžok intervalov ako konfidenčný interval založený na fiduciálnej distribúcii CI_{μ}^{FD} .

Literatúra

- [1] Hannig J., Iyer H.K., Wang C.M. (2007). *Fiducial approach to uncertainty assessment accounting for error due to instrument resolution*. *Metrologia* **44**, 476–483.
- [2] Wimmer G. (2008). *Približné konfidenčné intervaly pre meranú hodnotu v prípade digitalizovaných meraní*. 15. letná škola JČMF ROBUST 2008, 8.-12.9. 2008, Hotel Mier, Račková dolina, Pribylina.
- [3] Witkovský V. (2008). *Konfidenčné intervaly založené na digitalizovaných meraniach*. 15. letná škola JČMF ROBUST 2008, 8.-12.9. 2008, Hotel Mier, Račková dolina, Pribylina.
- [4] Witkovský V., Wimmer G. (2008). *Confidence intervals for the location parameter based on digitized measurements*. Submitted to *Mathematica Slovaca*.

Podakovanie: Práca bola podporená grantmi Vedeckej grantovej agentúry MŠ SR a SAV, VEGA 1/3016/06, a Agentúry pre podporu výskumu a vývoja, APVV RPEU-0008-06.

Adresa: G. Wimmer, Jr., Matematický ústav, SAV, Štefánikova 49, 814 73 Bratislava

E-mail: dzibo7@gmail.com

KONFIDENČNÉ INTERVALY ZALOŽENÉ NA DIGITALIZOVANÝCH MERANIACH

Viktor Witkovský, Gejza Wimmer

Kľúčové slová: Digitálne merania, ohraničená rozlíšiteľnosť, približný konfidenčný interval, zovšeobecnená fiduciálna inferencia.

Abstrakt: Nosnou témou príspevku je problematika konštrukcie konfidenčných oblastí pre parametre regresnej priamky, na základe digitalizovaných výsledkov merania. Tu sa však podrobnejšie zaoberáme len zjednodušeným špeciálnym prípadom: uvažujeme a porovnávame niekoľko metód na konštrukciu približných konfidenčných intervalov pre parameter polohy v prípade relatívne malého počtu digitalizovaných meraní. Jedným z cieľov je taktiež stručne popísať metódu založenú na zovšeobecnenom fiduciálnom prístupe, ktorú navrhli Hannig, Iyer a Wang (2007). V tomto prípade predpokladáme, že chyby meraní sú normálne rozdelené, nezávislé, s nulovou strednou hodnotou a rovnakým (neznámym) rozptylom. Výsledky merania sú však pozorovateľné len ako diskkrétne (digitalizované) hodnoty. Je to zapríčinené ohraničenou rozlíšiteľnosťou meracieho prístroja, ktorá je v procese merania považovaná za známu a nemennú. Keď je rozptyl chýb merania veľký vzhľadom k rozlíšiteľnosti prístroja, efekt digitalizácie zaniká a môžeme použiť štandardné štatistické postupy, ktoré predpokladajú, že realizácie pozorovaní pochádzajú zo spojitého rozdelenia. Preto sa zameriavame predovšetkým na analýzu situácie, keď je rozptyl chýb merania relatívne malý a efekt digitalizácie nemožno ignorovať.

Abstract: In this paper, the main topic of interest is the problem how to construct confidence regions for the parameters of regression line based on digitized measurements. However, here we focus only the simplified situation: we consider and compare several methods for making approximate confidence intervals for the location parameter based on relatively small number of digitized measurements. One of the goals is to present a brief description of the method based on the generalized fiducial inference, recently suggested by Hannig, Iyer and Wang (2007). In this case, it is assumed that the measurement errors come from normal distribution with zero mean and common unknown variance. However, the measurement results are observable on a discrete scale, only. This is due limited resolution of the used instrumentation, with fixed and known unit of resolution. If the error variance is large with respect to the resolution of the indicating measurement device, the effect of digitization diminishes and standard statistical inference is still valid. Hence, in this paper we consider mainly situations where the error variance is relatively small, and the effect of digitization could not be ignored.

1 Úvod

Efekt kvantizácie resp. diskretizácie spojitéch dát môže spôsobiť podstatnú stratu efektívnosti metód štatistickej inferencie, ktoré sú dosiahnuteľné v prípade platnosti pôvodného spojitého modelu. Tento problém je dávno známy v štatistickej literatúre, pozri napr. [16, 2] pričom bolo dosiahnutých mnoho výsledkov charakterizujúcich stratu efektívnosti rôznych metód založených na diskretizovaných pozorovaniach v porovnaní s metódami založenými na plnej informácii. Teda s využitím spojitéch (avšak nepozorovateľných) údajov, pričom boli navrhnuté aj alternatívne metódy s cieľom minimalizovať túto stratu. Mnoho významných výsledkov, motivovaných najmä problémami informačných a komunikačných technológií, je založených na asymptoticky platných tvrdeniach, keď možno uvažovať o zmene kvantizačnej jednotky pri rastúcom počte pozorovaní. Pre ilustráciu dávame do pozornosti výsledky a ďalšie citované práce uvedené v článku [15], kde bola uvažovaná komplikovanejšia situácia kvantizácie dát s konečným delením priestoru možných výsledkov, ktoré však môže narastať bez obmedzenia s rastúcim počtom pozorovaní.

Tento príspevok je motivovaný reálnymi problémami pri vyhodnocovaní nameraných údajov, špeciálne problémom konštrukcie približných konfidenčných oblastí pre parametre regresnej priamky na základe digitalizovaných výsledkov merania. Takéto výsledky merania sú zapríčinené ohraničenou rozlíšiteľnosťou meracieho prístroja alebo kvantizáciou vznikajúcou v procese konverzie analógového signálu na digitálny. Ide o problém motivovaný potrebami metrologickej praxe, keď je potrebné predovšetkým charakterizovať neistoty výsledkov merania. Pre ilustráciu uvádzame výber z dlhého zoznamu publikácií, ktoré tento problém študovali hlavne z pohľadu metrologie, [1, 3, 5, 10, 13, 14, 17, 18, 19].

Podrobnejšie sa však budeme zaoberať len zjednodušeným špeciálnym prípadom konštrukcie približných konfidenčných intervalov pre parameter polohy, za predpokladu, že nepozorovateľné chyby meraní pred kvantizáciou sú normálne rozdelené, nezávislé, s nulovou strednou hodnotou a rovnakým (neznámym) rozptylom. Výsledky merania sú však pozorovateľné len ako diskrétné (digitalizované) hodnoty. Na rozdiel od predchádzajúcej situácie je počet pozorovaní obyčajne veľmi malý a nie je možné meniť kvantizačnú jednotku (rozlíšiteľnosť meracieho prístroja), ktorá je v procese merania považovaná za známú a nemennú.

V príspevku uvidíme prehľad niektorých metód konštrukcie približných konfidenčných intervalov pre parameter polohy. Špeciálnu pozornosť venujeme metóde založenej na zovšeobecnenom fiduciálnom prístupe, ktorú navrhol J. Hannig a jeho kolegovia v prácach [6, 7, 8]. V poslednej časti spomenieme aj metódu konštrukcie približného konfidenčného intervalu, ktorý je založený na profilovej vierohodnostnej funkcii (profile likelihood), pozri [11, 12]. Výsledky porovnania vybraných konfidenčných intervalov pre parameter polohy, založené na simulačnej štúdii, sú prezentované v práci [20].

2 Efekt digitalizácie v prípade regresnej priamky

Nech $\beta = (\beta_0, \beta_1)'$ označuje skutočnú hodnotu parametrov regresnej priamky, ktorú budeme odhadovať na základe nezávislých digitalizovaných meraní, pozorovaných na diskretnej stupnici s jednotkou rozlíšenia $\delta \geq 0$. Je to najmenší rozdiel medzi dvomi hodnotami ktoré vieme (zmysluplne) rozlíšiť meracím prístrojom.

Najprv uvažujme klasický spojitý model, teda n nezávislých meraní, $n \geq 3$, ktoré odčítame na prístroji s ideálnou rozlíšiteľnosťou $\delta = 0$. Označme $y = (y_1, y_2, \dots, y_n)'$ realizáciu náhodného vektora $Y = (Y_1, Y_2, \dots, Y_n)'$, kde $Y_i = \beta_0 + \beta_1 f_i + \sigma Z_i$, pričom f_i reprezentujú hodnoty vysvetľujúcej premennej a $Z_i \sim N(0, 1)$ pre $i = 1, \dots, n$. Pre jednoduchosť predpokladajme, že matica plánu F má plnú hodnotnosť, pričom $F = [\mathbf{1}, f]$, kde $\mathbf{1} = (1, \dots, 1)'$ a $f = (f_1, \dots, f_n)'$. Optimálne odhady parametrov β a σ^2 sú

$$\hat{\beta} = (F'F)^{-1}F'Y \quad \text{a} \quad S^2 = \frac{1}{n-2} \sum_{i=1}^n (Y_i - \hat{\beta}_0 - \hat{\beta}_1 f_i)^2. \quad (1)$$

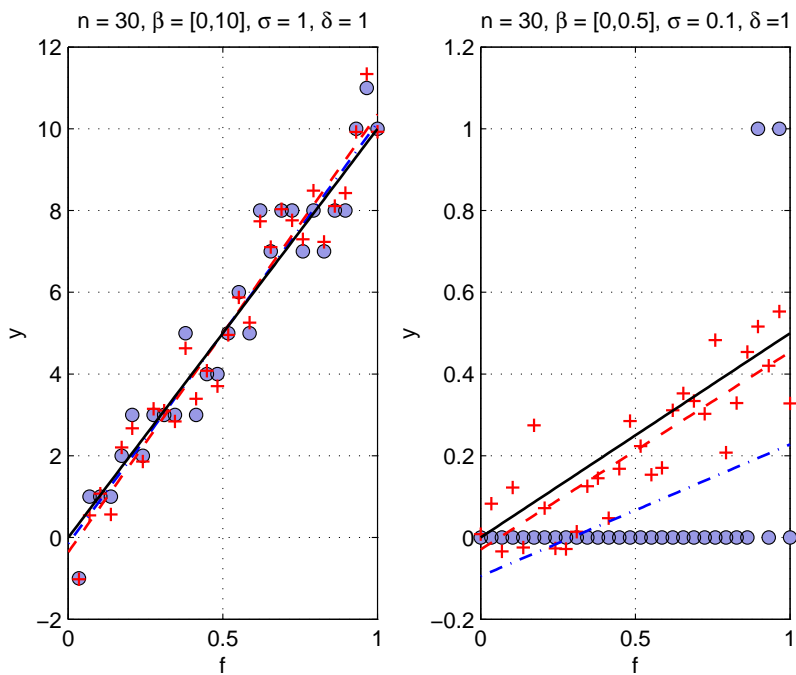
V prípade, že merací prístroj má len obmedzené rozlíšenie ($\delta > 0$), výsledkom merania sú hodnoty vektora $x = (x_1, x_2, \dots, x_n)'$ (údaje určené na stupnici s diskretnými hodnotami), pre ktoré platí $x_i = k_i \delta = \lfloor y_i / \delta + 0.5 \rfloor \delta = \lfloor (\beta_0 + \beta_1 f_i + \sigma z_i) / \delta + 0.5 \rfloor \delta$, $i = 1, 2, \dots, n$, pričom $k_i \in \{0, \pm 1, \pm 2, \dots\}$, $\lfloor a \rfloor$ označuje funkciu zaokrúhľovania na najväčšie celé číslo, ktoré nie je väčšie ako a . V tomto prípade teda predpokladáme, že ide o klasické zaokrúhľovanie y_i k najbližšej hodnote na danej digitálnej stupnici. Analogicky by sme mohli postupovať, ak prístroj namiesto zaokrúhľovania namerané hodnoty usekáva.

Na obr. 1 sú zobrazené dve konkrétne situácie so simulovanými hodnotami meraní, ktoré názorne ilustrujú možný efekt digitalizácie na kvalitu odhadu regresnej priamky. Kým na obrázku v ľavej časti je rozdiel medzi odhadnutými priamkami minimálny, ako vidieť na obrázku vpravo, tento efekt je výrazný v situácii keď parameter β_1 (sklon priamky) je malý a skutočná štandardná odchýlka chýb meraní σ je relatívne malá oproti danému rozlíšeniu meracieho prístroja δ .

Iný pohľad na efekt digitalizácie je ilustrovaný na obr. 2. Graficky sú tu zobrazené empirické pravdepodobnosti pokrytia skutočného parametra β , v prípade ak použijeme konfidenčnú oblasť CI , ktorá je optimálna v spojitom modeli, aj v prípade keď máme k dispozícii len digitalizované pozorovania. CI je, v prípade spojitého modelu, exaktná $(1 - \alpha) \times 100\%$ -konfidenčná oblasť pre parameter β

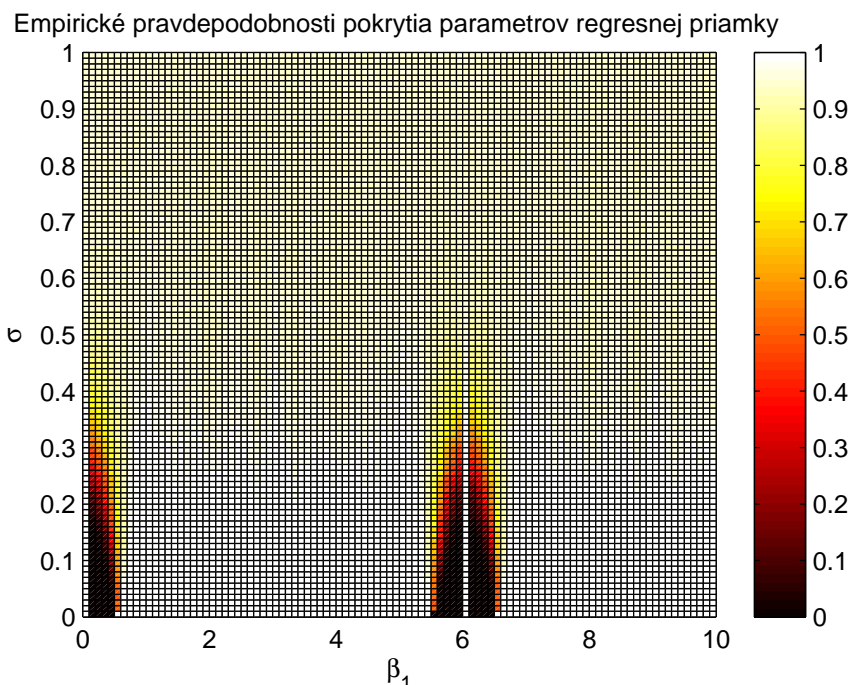
$$CI_{\beta}^F(1 - \alpha) := \left\{ \beta : (\hat{\beta} - \beta)'(F'F)(\hat{\beta} - \beta) \leq 2S^2 F_{2, n-2}(1 - \alpha) \right\}, \quad (2)$$

pričom $F_{2, n-2}(1 - \alpha)$ je $(1 - \alpha)$ -kvantil F -rozdelenia s 2 a $n - 2$ stupňami voľnosti.



Obrázok 1: Odhadnuté regresné priamky na základe spojitých a digitalizovaných meraní. Symbol + označuje hodnoty meraní registrované na prístroji s ideálnym rozlíšením $\delta = 0$, symbol \circ označuje hodnoty rovnakých meraní registrované na prístroji s obmedzeným rozlíšením $\delta = 1$. Plnou čiarou je zobrazená skutočná regresná priamka, prerušovaná čiara reprezentuje jej odhad založený na spojitých meraniach, bodko-čiarkovaná čiara reprezentuje odhad (metódou najmenších štvorcov) založený na digitalizovaných meraniach. Na obrázku vľavo sú tieto priamky nerozlíšiteľné, na rozdiel od obrázku vpravo, kde je rozdiel medzi odhadom založeným na digitalizovaných meraniach a odhadom založeným na spojitých meraniach výrazný.

Ako vidieť, výsledky simulačnej štúdie potvrdzujú známy fakt, že pokiaľ je skutočná hodnota parametra σ pomerne veľká vzhľadom k parametru rozlíšenia δ , efekt digitalizácie nie je významný, a z praktického pohľadu možno použiť štandardné metódy štatistickej inferencie (v tomto prípade konštrukciu konfidenčnej oblasti), ktoré sú vhodné pre pôvodný spojitý model. Na druhej strane vidieť dve kritické oblasti, kde sa empirická pravdepodobnosť veľmi vzdialila od očakávanej nominálnej hodnoty 0.95 a poklesla až na úroveň 0. Oblasť napravo ($\beta_1 \approx 6$) vznikla dôsledkom špecifickej situácie pre daný návrh experimentu s ekvidistantnými hodnotami vysvetľujúcej premennej f , keď $S^2 = 0$. Negatívny efekt digitalizácie, ktorý možno pozorovať pre všetky návrhy experimentu, je v situácii keď je skutočná hodnota para-



Obrázok 2: Empirické pravdepodobnosti pokrytia parametrov regresnej priamky $\beta = (\beta_0, \beta_1)'$ určené na základe $N = 1000$ simulovaných digitalizovaných meraní s rozlíšením $\delta = 1$. V simuláciách sme uvažovali regresný model pre rôzne kombinácie skutočných parametrov modelu: $n = 7$, $f = [0 : 1/6 : 1]$, $\beta = [0, \beta_1]$, pričom $\beta_1 \in \{0 : 0.1 : 10\}$ a $\sigma \in \{0 : 0.01 : 1\}$. Pre každú kombináciu parametrov sme zobrazili napozorovanú hodnotu pravdepodobnosti pokrytia $\Pr(\beta \in CI_{\beta}^F(1 - 0.05))$.

metra σ a skutočný sklon priamky β_1 nadobúdajú pomerne malé hodnoty vzhľadom k parametru rozlíšenia δ .

Z uvedených ilustračných príkladov vyplýva potreba hľadania nových vylepšených metód na konštrukciu konfidenčnej oblasti pre parametre regresnej priamky, na základe digitalizovaných údajov, obzvlášť v situáciách, ktoré sme identifikovali vyššie. V nasledujúcej časti príspevku popíšeme metódy, ktoré sú známe v zjednodušenej situácii, keď je cieľom určenie (približného) konfidenčného intervalu pre parameter polohy na základe digitalizovaných meraní.

3 Približné konfidenčné intervaly pre parameter polohy

Nech μ označuje skutočnú hodnotu meranej veličiny (parameter polohy), ktorú budeme odhadovať na základe nezávislých digitalizovaných meraní. V tejto časti stručne uvedieme základné metódy konštrukcie približného konfidenčného intervalu pre parameter μ , ktoré boli odporúčané pre metrologické aplikácie.

Podobne ako v predchádzajúcej časti, najprv uvažujme n nezávislých meraní, $n \geq 2$, ktoré odčítame na prístroji s ideálnou rozlíšiteľnosťou $\delta = 0$. Označme $Y_i = \mu + \sigma Z_i$, pričom $Z_i \sim N(0, 1)$. Optimálne odhady parametrov μ a σ^2 sú potom $\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$ a $S_Y^2 = \frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y})^2$. Exaktný dvojbstranný $(1 - \alpha) \times 100\%$ -konfidenčný interval pre μ je potom

$$\bar{Y} \mp \frac{S_Y}{\sqrt{n}} t_{n-1}(1 - \alpha/2), \quad (3)$$

kde $t_{n-1}(1 - \alpha/2)$ je $(1 - \alpha/2)$ -kvantil Studentovho t -rozdelenia s $n - 1$ stupňami voľnosti.

V prípade nezávislých digitalizovaných meraní sú výsledkami meraní hodnoty realizácie nezávislých náhodných veličín X_1, X_2, \dots, X_n , kde $X_i = K_i \delta$, a $K_i = \lfloor (\mu + \sigma Z_i) / \delta + 0.5 \rfloor$, $Z_i \sim N(0, 1)$, $i = 1, 2, \dots, n$. Rozdelenie pravdepodobnosti náhodnej premennej X_i je určené pravdepodobnostnou funkciou $\{P_{k\delta}\}$, $k = 0, \pm 1, \pm 2, \dots$, pričom $P_{k\delta} = \Pr(X_i = k\delta)$ je

$$P_{k\delta} = \int_{(k-0.5)\delta}^{(k+0.5)\delta} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(y-\mu)^2}{2\sigma^2}\right\} dy. \quad (4)$$

Na základe náhodného výberu X_1, X_2, \dots, X_n z rozdelenia (4) možno odvodiť približný konfidenčný interval pre μ na základe približného rozdelenia pre odhady parametrov metódou maximálnej vierohodnosti (ML). Pre dané hodnoty x_1, x_2, \dots, x_n , kde $x_i = k_i \delta$, je logaritmus vierohodnostnej funkcie pre parametre μ a σ

$$\ell(\mu, \sigma | x_1, \dots, x_n) = \sum_{i=1}^n \ln(P_{k_i \delta}) = \sum_{i=1}^m n_i \ln(P_{k_i \delta}), \quad (5)$$

pričom m označuje počet rôznych hodnôt $k_i \delta$, $i = 1, \dots, m$ s ich početnosťami n_i , $n = \sum_{i=1}^m n_i$. Realizácie ML odhadov

$$(\hat{\mu}_{ML}, \hat{\sigma}_{ML})' = \operatorname{argmax}_{(\mu, \sigma) \in \Theta} \ell(\mu, \sigma | X_1, \dots, X_n), \quad (6)$$

kde $\Theta = (-\infty, \infty) \times (0, \infty)$ možno spočítať numericky pre ľubovoľné hodnoty digitalizovaných pozorovaní x_1, x_2, \dots, x_n . ML odhad však neexistuje v prípade keď $X_{(n)} - X_{(1)} \leq \delta$, kde $X_{(1)}, \dots, X_{(n)}$ sú poriadkové štatistiky.

Pri dostatočne veľkom počte pozorovaní n sa na konštrukciu konfidenčného intervalu pre μ odporúča využiť približné rozdelenie ML odhadov (za predpokladu ich existencie)

$$\begin{pmatrix} \hat{\mu}_{ML} \\ \hat{\sigma}_{ML} \end{pmatrix} \overset{apr.}{\sim} N \left(\begin{pmatrix} \mu \\ \sigma \end{pmatrix}, I_F^{-1}(\mu, \sigma) \right), \quad (7)$$

kde $I_F(\mu, \sigma) = nJ(\mu, \sigma)$ má prvky

$$\begin{aligned} J_{1,1}(\mu, \sigma) &= \sum_{k=0, \pm 1, \dots} \left[\frac{\partial P_{k\delta}}{\partial \mu} \frac{1}{P_{k\delta}} \right]^2 P_{k\delta}, \\ J_{1,2}(\mu, \sigma) &= \sum_{k=0, \pm 1, \dots} \left[\frac{\partial P_{k\delta}}{\partial \mu} \frac{\partial P_{k\delta}}{\partial \sigma} \frac{1}{P_{k\delta}^2} \right] P_{k\delta} = J_{2,1}(\mu, \sigma), \\ J_{2,2}(\mu, \sigma) &= \sum_{k=0, \pm 1, \dots} \left[\frac{\partial P_{k\delta}}{\partial \sigma} \frac{1}{P_{k\delta}} \right]^2 P_{k\delta}, \end{aligned} \quad (8)$$

pričom

$$\begin{aligned} P_{k\delta} &= \Phi(u) - \Phi(l), \\ \frac{\partial P_{k\delta}}{\partial \mu} &= -\frac{1}{\sigma} [\varphi(u) - \varphi(l)], \\ \frac{\partial P_{k\delta}}{\partial \sigma} &= -\frac{1}{\sigma} [u\varphi(u) - l\varphi(l)], \end{aligned} \quad (9)$$

$u = [(k + 0.5)\delta - \mu]/\sigma$, $l = [(k - 0.5)\delta - \mu]/\sigma$, $\Phi(\cdot)$ je distribučná funkcia a $\varphi(\cdot)$ je hustota $N(0, 1)$ -rozdelenia.

Na základe (7) možno skonštruovať približný dvojstranný $(1 - \alpha) \times 100\%$ -konfidenčný interval pre μ

$$CI_{\mu}^{ML}(1 - \alpha) := \hat{\mu}_{ML} \mp \sqrt{\frac{1}{n} \frac{\hat{J}_{2,2}}{\hat{J}_{1,1}\hat{J}_{2,2} - \hat{J}_{1,2}^2}} u(1 - \alpha/2), \quad (10)$$

kde $\hat{J}_{i,j}$, $i, j = 1, 2$ sú prvky odhadnutej matice $\hat{J} = J(\hat{\mu}_{ML}, \hat{\sigma}_{ML})$ a $u(1 - \alpha/2)$ je $(1 - \alpha/2)$ -kvantil $N(0, 1)$ -rozdelenia.

Alternatívnou metódou konštrukcie konfidenčných intervalov pre μ je metóda založená na invertovaní testu pomerom vierohodností (LR) hypotézy $H_0 : \mu = \mu_0$. Definujme LR testovaciu štatistiku

$$LR(\mu_0) = -2 \left[\sup_{\sigma} \{ \ell(\mu_0, \sigma | X_1, \dots, X_n) \} - \sup_{(\mu, \sigma)} \{ \ell(\mu, \sigma | X_1, \dots, X_n) \} \right], \quad (11)$$

Za platnosti hypotézy $H_0 : \mu = \mu_0$ asymptoticky platí

$$LR(\mu_0) \stackrel{\mathcal{L}}{\sim} \chi_1^2. \quad (12)$$

Potom približný $(1 - \alpha) \times 100\%$ konfidenčný interval pre parameter μ , určený LR metódou (likelihood-ratio), definujeme ako

$$CI_{\mu}^{LR}(1 - \alpha) := \{ \mu : LR(\mu) < \chi_1^2(1 - \alpha) \}, \quad (13)$$

kde $\chi_1^2(1 - \alpha)$ označuje $(1 - \alpha)$ -kvantil χ^2 -rozdelenia s jedným stupňom voľnosti.

Ako sme uviedli, efekt digitalizácie je zanedbateľný ak hodnota parametra σ je relatívne veľká oproti parametru rozlíšiteľnosti δ . V tomto prípade možno použiť štandardný postup. Tento argument často slúži ako odôvodnenie pre použitie konfidenčného intervalu (3) aj v prípade digitalizovaných meraní. Preto ako *prírodzený* približný obojstranný $(1 - \alpha) \times 100\%$ -konfidenčný interval pre μ možno považovať práve klasický štandardný intervalový odhad

$$CI_{\mu}^{St}(1 - \alpha) := \bar{X} \mp \frac{S_X}{\sqrt{n}} t_{n-1}(1 - \alpha/2), \quad (14)$$

kde $S_X^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$. Konfidenčný interval (14) však degeneruje ak realizácie všetkých X_i , $i = 1, \dots, n$, sú rovnaké hodnoty, teda ak $X_{(1)} = X_{(n)}$.

R. Willink v [19] navrhol modifikovaný odhad disperzie výberového priemeru $\sigma_{\bar{X}}^2 = Var(\bar{X})$:

$$\hat{\sigma}_{\bar{X}}^2 = \begin{cases} \frac{c_n \delta^2}{12}, & \text{ak } X_{(1)} = X_{(n)}, \\ \max \left\{ \frac{S_X^2}{n}, \frac{[(X_{(1)} + X_{(n)})/2 - \bar{X}]^2}{3} \right\}, & \text{ak } X_{(n)} - X_{(1)} = \delta, \\ \frac{S_X^2}{n}, & \text{inokedy,} \end{cases} \quad (15)$$

pričom c_n je vhodná korekcia závislá na rozsahu výberu n ($c_2 = 6.4$, $c_3 = 1.3$ a $c_n = 1$ pre $n \geq 4$). Pomocou (15) získavame približný obojstranný $(1 - \alpha) \times 100\%$ -konfidenčný interval pre μ :

$$CI_{\mu}^W(1 - \alpha) := \bar{X} \mp \hat{\sigma}_{\bar{X}} t_{n-1}(1 - \alpha/2). \quad (16)$$

4 Približné konfidenčné intervaly pre parameter polohy založené na fiduciálnej distribúcii

4.1 Zovšeobecnená fiduciálna inferencia

Fiduciálny argument R.A. Fishera, pozri napr. [4], bol pokusom robiť rigorózne pravdepodobnostné tvrdenia o neznámych parametroch rozdelení, z ktorých máme k dispozícii realizácie náhodného výberu, bez predpokladu akejkoľvek apriórnej pravdepodobnostnej znalosti o ich možných hodnotách - čo Fisher vnímal ako slabinu štatistickej inferencie založenej na bayesovských metódach.

Hoci je tento postup intuitívne veľmi atraktívny, a v blízkom vzťahu k široko akceptovaným metódam štatistickej inferencie založeným na maximalizácii funkcie vierohodnoti, Fisherovi sa nepodarilo vybudovať zrozumiteľnú a jasnú teóriu fiduciálnej inferencie. Fisherovi oponenti poukazovali na nekonzistentnosť s klasickými (frekventistickými) metódami štatistickej inferencie a na častú nejednoznačnosť fiduciálnych metód. Navyše, fiduciálna

distribúcia parametrov je častokrát v zhode s bayesovskou posteriornou distribúciou, získanou pri neinformatívnom apriórnom rozdelení (vo všeobecnosti však nie je totožná s bayesovským riešením).

V posledných rokoch badať obnovený záujem o fiduciálny prístup k inferencii. J. Hannig v prácach [7, 6] preformuloval princípy fiduciálnej inferencie a zovšeobecnil ich pre prípad vektorových parametrov spojitéj ako aj diskretných rozdelení. Podarilo sa identifikovať zdroje nejednoznačnosti fiduciálneho rozdelenia a ukázať dobré asymptotické vlastnosti metód založených na ňom.

Fiduciálny argument je založený na výmene úlohy parametrov a dát. Uvádzame tu stručný popis konštrukcie fiduciálnej distribúcie podľa [7, 6]:
Nech

$$X = G(\xi, U) \quad (17)$$

je známy systém štruktúrnych rovníc (dáta generujúci mechanizmus), ktorý spája náhodný vektor pozorovaní X s vektorom parametrov $\xi \in \Xi$ a náhodným vektorom U (chybový - šumový proces), so známou distribučnou funkciou, ktorá nezávisí od parametrov.

Pre ľubovoľnú (pozorovanú) hodnotu x náhodného vektora X a pre realizáciu u vektora U definujeme množinovo-hodnotovú funkciu

$$Q(x, u) = \{\xi : x = G(\xi, u)\}. \quad (18)$$

Funkcia $Q(X, U)$ reprezentuje *inverznú funkciu* ku G . Pre ľubovoľnú pevnú hodnotu x a u je výsledkom $Q(x, u)$ buď prázdna množina, jednoprvková množina obsahujúca jednoznačný parameter ξ , alebo zložitejšia mnohoprvková množina možných parametrov ξ , ktorá závisia od vzťahov medzi x , u a G . Nech $V(\cdot)$ definuje akýkoľvek náhodný mechanizmus na merateľných množinách. Pre ľubovoľnú množinu S , $V(S)$ náhodne generuje prvok z uzavere množiny S . Potom, špeciálny výber (v závislosti od G , Q , a V) zovšeobecnenej fiduciálnej distribúcie parametra ξ , označme ju $F_\xi(\xi^*)$ pre pevné ξ^* , je definovaný ako akákoľvek verzia podmienenej distribúcie náhodnej premennej $V(Q(x, U^*))$ za podmienky $Q(x, U^*) \neq \emptyset$, teda

$$F_\xi(\xi^*) = \Pr(V(Q(x, U^*)) \leq \xi^* | Q(x, U^*) \neq \emptyset), \quad (19)$$

kde x je napozorovaná hodnota náhodného vektora X a U^* je nezávislá náhodná premenná s rovnakým rozdelením ako U . Symbolom ' \leq ' označujeme usporiadanie po jednotlivých zložkách vektorov. Úloha nájsť explicitný tvar fiduciálnej distribúcie môže byť veľmi zložitá. Avšak, uvedený postup dáva priamy návod na generovanie náhodného výberu parametrov z danej fiduciálnej distribúcie.

Pre ilustráciu, uvažujme náhodný výber meraní registrovaných na zariadení s dokonalým rozlíšením: Y_1, Y_2, \dots, Y_n , kde $Y_i = \mu + \sigma Z_i$, pričom $Z_i \sim N(0, 1)$. Pre postačujúce štatistiky \bar{Y} a S_Y^2 platí $\bar{Y} = \mu + (\sigma/\sqrt{n})Z$, kde $Z \sim N(0, 1)$ a , kde $S_Y^2 = \sigma^2 W/(n-1)$ $W \sim \chi_{(n-1)}^2$, pričom sú navzájom

nezávislé. Štruktúrálné rovnice definujeme v tvare $X = G(\xi, U)$, kde $\xi = (\mu, \sigma)$, $U = (Z, W)$, $X = (\bar{Y}, S_Y^2)$, teda

$$G(\xi, U) = (\mu + (\sigma/\sqrt{n})Z, \sigma^2 W/(n-1)).$$

Inverznú funkciu $Q(x, u)$ definujeme ako

$$Q(x, u) = Q((\bar{y}, s_Y^2), (z, w)) = \left\{ \left(\bar{y} - \sqrt{\frac{(n-1)s_Y^2}{nw}} z, \sqrt{\frac{(n-1)s_Y^2}{w}} \right) \right\},$$

kde \bar{y} je pozorovaná hodnota (realizácia) \bar{Y} a s_Y^2 je realizácia S_Y^2 . Tu $Q(x, u)$ vždy definuje neprázdnu množinu, ktorá obsahuje jedinú hodnotu z parametrického priestoru. Teda fiduciálna distribúcia nezávisí od žiadneho náhodného mechanizmu $V(\cdot)$:

$$F_{(\mu, \sigma)}((\mu^*, \sigma^*)) = \Pr \left(\bar{y} - \sqrt{\frac{(n-1)s_Y^2}{nW_{n-1}^*}} Z_1^* \leq \mu^*, \sqrt{\frac{(n-1)s_Y^2}{W_{n-1}^*}} \leq \sigma^* \right), \quad (20)$$

kde $W_{n-1}^* \sim \chi_{(n-1)}^2$ a $Z_1^* \sim N(0, 1)$ sú navzájom nezávislé náhodné premenné. Toto rozdelenie je totožné s fiduciálnym rozdelením, ktoré odvodil Fisher v [4]. Ľahko možno analogickým postupom odvodiť fiduciálne rozdelenie parametrov (β, σ) spojitého lineárneho regresného modelu (priamky):

$$F_{(\beta, \sigma)}((\beta^*, \sigma^*)) = \Pr \left(\hat{\beta}_{obs} - \sqrt{\frac{(n-l)s^2}{W_{n-l}^*}} (F'F)^{-1/2} Z_l^* \leq \beta^*, \sqrt{\frac{(n-l)s^2}{W_{n-l}^*}} \leq \sigma^* \right), \quad (21)$$

kde $\hat{\beta}_{obs}$ a s^2 sú napozorované hodnoty odhadov $\hat{\beta}$ a S^2 uvedených v (1), $W_{n-l}^* \sim \chi_{(n-l)}^2$ a $Z_l^* \sim N(0, I_l)$ sú navzájom stochasticky nezávislé, pričom pre regresnú priamku $l = 2$. Vidieť, že združené fiduciálne rozdelenie parametrov (μ, σ) (teda rozdelenie fiduciálnych náhodných premenných $\tilde{\mu} = \bar{y} - \sqrt{(n-1)s_Y^2/(nW_{n-1}^*)} Z_1^*$ a $\tilde{\sigma} = \sqrt{(n-1)s_Y^2/(W_{n-1}^*)}$) resp. rozdelenie parametrov (β, σ) (t.j. rozdelenie $\tilde{\beta} = \hat{\beta}_{obs} - \sqrt{(n-l)s^2/W_{n-l}^*} (F'F)^{-1/2} Z_l^*$ a $\tilde{\sigma} = \sqrt{(n-l)s^2/W_{n-l}^*}$) je spojité, pričom združenú fiduciálnu hustotu možno priamo odvodiť z uvedeného tvaru fiduciálnych náhodných premenných.

4.2 Fiduciálna distribúcia parametrov modelu digitalizovaných meraní

Hannig, Iyer a Wang v práci [8] navrhli metódu pre generovania náhodného výberu so združenej fiduciálnej distribúcie parametrov μ a σ , na základe napozorovaného náhodného výberu z digitalizovaných meraní ($\delta \geq 0$).

Nech $X = (X_1, X_2, \dots, X_n)$ je náhodný výber digitalizovaných meraní a $x = (x_1, x_2, \dots, x_n)$ označuje napozorované hodnoty. Štruktúralne rovnice $X = G(\xi, Z)$, kde $\xi = (\mu, \sigma)$ a $Z = (Z_1, \dots, Z_n)$, $Z_i \sim N(0, 1)$ sú nezávislé rovnako rozdelené náhodné premenné, sú dané vzťahmi

$$X_i = \delta K_i = \delta[(\mu + \sigma Z_i)/\delta + 0.5],$$

kde δ je rozlíšenie meracieho prístroja. Odtiaľ teda platia nerovnosti

$$X_i - \frac{\delta}{2} \leq \mu + \sigma Z_i < X_i + \frac{\delta}{2}, \quad i = 1, \dots, n. \quad (22)$$

Definujme funkciu $Q(x, z)$ ako

$$Q(x, z) = \{(\mu, \sigma) : x_i - \delta/2 \leq \mu + \sigma z_i \leq x_i + \delta/2, \quad i = 1, \dots, n\}. \quad (23)$$

Pre ľubovoľné pevné hodnoty x a z , výsledkom funkcie $Q(x, z)$ je buď prázdna množina, alebo množina hodnôt, ktorá vytvára polygón v 2-rozmernom parametrickom priestore (μ, σ) . Množina parametrov $Q(x, z)$ vytvára polygón ak $z \in S(x)$, kde

$$S(x) = \{(z_1, \dots, z_n) : \exists(\mu, \sigma), \text{ že} \\ x_i - \delta/2 \leq \mu + \sigma z_i < x_i + \delta/2, \quad i = 1, 2, \dots, n\}. \quad (24)$$

Ak $Q(x, z)$ je neprázdna množina s nekonečným množstvom možných hodnôt parametrov (μ, σ) , potom vyberieme náhodný mechanizmus $V(\cdot)$ pre výber parametra (μ, σ) . Hoci akákoľvek voľba $V(\cdot)$ je prípustná (s nosičom na množine danej uzáverom $Q(x, z)$), tu navrhujeme použiť $V(\cdot)$ s rovnomerným rozdelením na $Q(x, z)$.

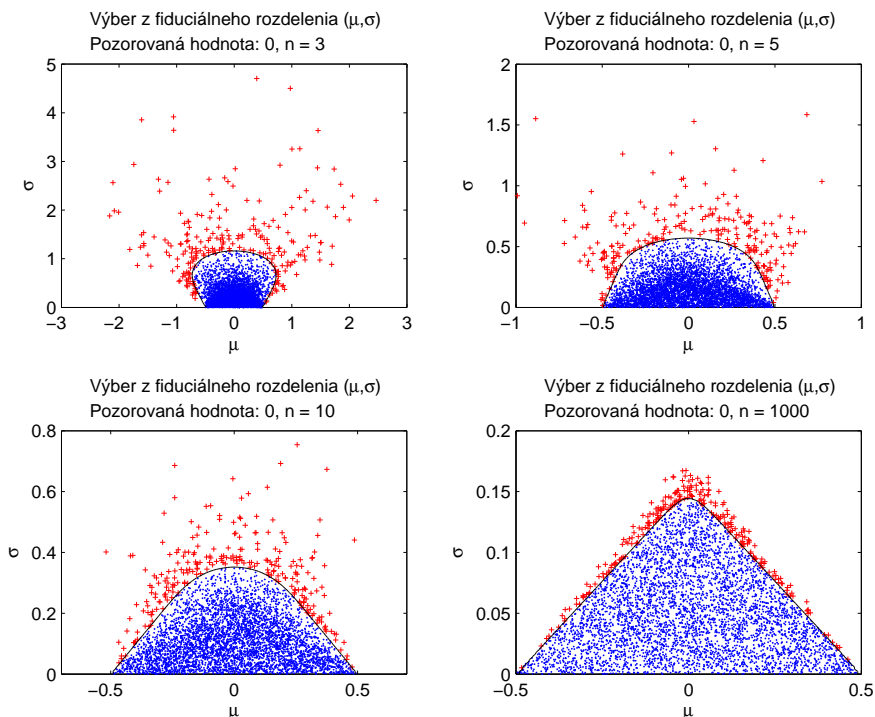
Združená fiduciálna distribučná funkcia parametrov (μ, σ) (závisí od G, Q , a V) je potom podľa (19) daná vzťahom

$$F_{(\mu, \sigma)}((\mu^*, \sigma^*)) = \Pr(V(Q(x, Z^*)) \leq (\mu^*, \sigma^*) | Z^* \in S(x)). \quad (25)$$

kde (μ^*, σ^*) je pevne zvolený vektor hodnôt a $Z^* = (Z_1^*, Z_2^*, \dots, Z_n^*)' \sim N(0, I_n)$. V takomto prípade (keď $V(\cdot)$ je náhodný mechanizmus s rovnomerným rozdelením na $Q(x, z)$) je fiduciálne rozdelenie parametrov spojité, avšak určenie explicitného tvaru združenej fiduciálnej hustoty je zložité.

Na obr. 3 a 4 sú pre ilustráciu zobrazené náhodné výbery zo združenej fiduciálnej distribúcie parametrov μ a σ pre niekoľko konkrétnych situácií, ktoré sú založené na nezávislých pozorovaniach digitalizovaných meraní na prístroji s rozlíšením $\delta = 1$.

Pre dané pozorované merania (x_1, \dots, x_n) je možné skonštruovať združenú (približnú) $(1 - \alpha) \times 100\%$ konfidenčnú oblasť pre parametre (μ, σ) s pomocou fiduciálnej distribúcie (25). Náhodný výber parametrov (μ, σ) z fiduciálnej distribúcie (25) možno usporiadať podľa hodnoty logaritmu vierohodnostnej funkcie (5). Empirická 95% konfidenčná oblasť je potom určená ako oblasť parametrov (μ, σ) , pre ktoré je hodnota funkcie $\ell(\mu, \sigma | x_1, \dots, x_n)$ väčšia ako empirický odhad α -kvantilu $\ell_{(\alpha)}(\mu, \sigma | x_1, \dots, x_n)$, pre zvolené $\alpha \in (0, 1)$, napr. $\alpha = 0.05$, pozri obr. 5.



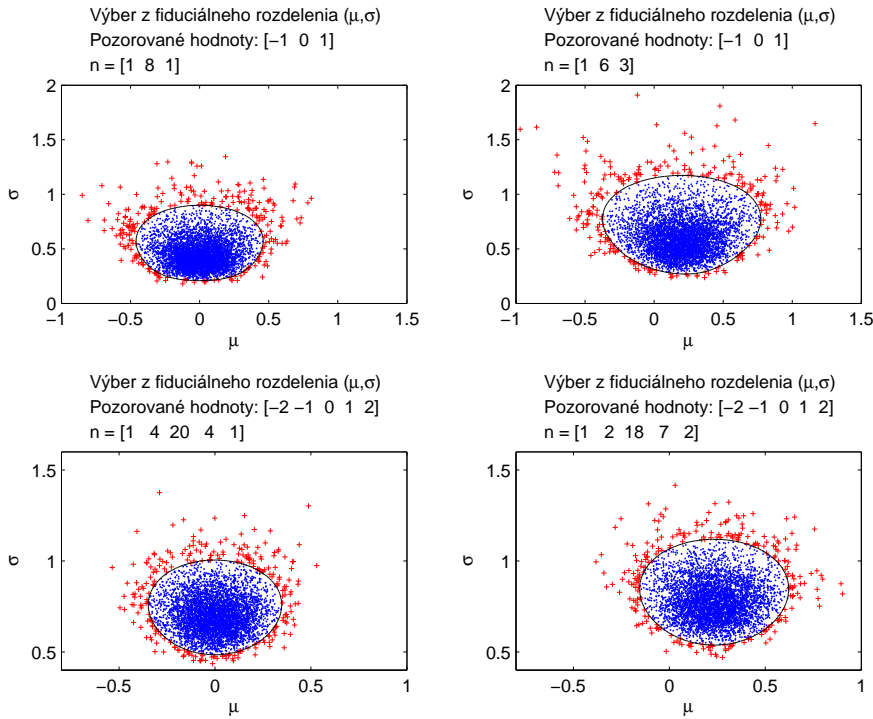
Obrázok 3: Náhodný výber rozsahu $N = 5000$ so združenej fiduciálnej distribúcie parametrov μ a σ založený na n nezávislých pozorovaniach nameranej hodnoty 0 na prístroji s rozlíšením $\delta = 1$ ($n = 3, 5, 10$, and 1000). Realizácie výberu z fiduciálneho rozdelenia parametrov (μ, σ) boli usporiadané podľa hodnoty funkcie vierohodnosti s vyznačením empirickej 95% oblasti.

4.3 Exaktná konfidenčná oblasť pre parametre modelu spojitých meraní

V prípade štandardných spojitých meraní (na prístroji s dokonalým rozlíšením $\delta = 0$) možno dokázať, že takto skonštruovaná oblasť je exaktná $(1 - \alpha) \times 100\%$ konfidenčná oblasť pre parametre (μ, σ) : Uvažujme náhodný výber meraní, Y_1, \dots, Y_n , ($Y_i = \mu + \sigma Z$, $Z \sim N(0, 1)$), na prístroji s dokonalým rozlíšením $\delta = 0$. Logaritmus funkcie vierohodnosti je

$$\ell(\mu, \sigma | y_1, \dots, y_n) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu)^2. \quad (26)$$

Nech y_1, \dots, y_n označuje pozorované hodnoty meraní Y_1, \dots, Y_n a $\tilde{\mu} = \bar{y} - \sqrt{(n-1)s_Y^2 / (nW_{n-1}^*)} Z_1^*$ a $\tilde{\sigma} = \sqrt{(n-1)s_Y^2 / W_{n-1}^*}$ sú fiduciálne náhodné premenné. Odvodíme rozdelenie náhodnej premennej $\ell(\tilde{\mu}, \tilde{\sigma} | y_1, \dots, y_n)$:

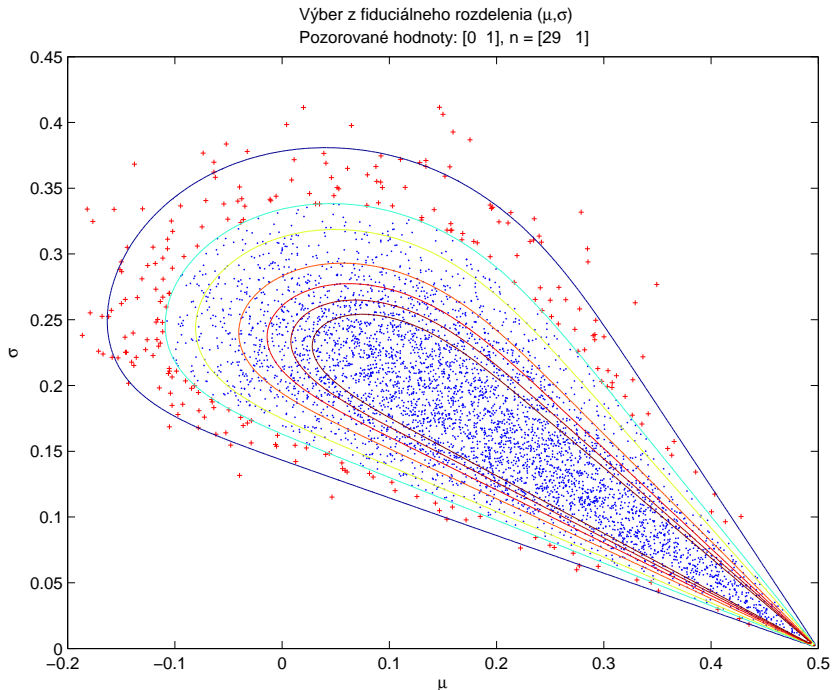


Obrázok 4: Náhodný výber rozsahu $N = 5000$ zo združenej fiduciálnej distribúcie parametrov μ a σ založený na n ($n = 10$ a 30) nezávislých pozorovaniach hodnôt $[-1, 0, 1]$ a $[-2, -1, 0, 1, 2]$ na prístroji s rozlíšením $\delta = 1$. Realizácie výberu z fiduciálneho rozdelenia parametrov (μ, σ) boli usporiadané podľa hodnoty funkcie vierohodnosti s vyznačením empirickej 95% oblasti.

$$\begin{aligned}
\ell(\tilde{\mu}, \tilde{\sigma} | y_1, \dots, y_n) &\sim -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log(\tilde{\sigma}^2) - \frac{1}{2\tilde{\sigma}^2} \sum_{i=1}^n (y_i - \tilde{\mu})^2 \\
&\sim -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log\left(\frac{(n-1)s_Y^2}{W_{n-1}^*}\right) \\
&\quad - \frac{1}{2} \frac{W_{n-1}^*}{(n-1)s_Y^2} \sum_{i=1}^n (y_i - \bar{y} + \bar{y} - \tilde{\mu})^2 \\
&\sim -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log\left(\frac{(n-1)s_Y^2}{W_{n-1}^*}\right) \\
&\quad - \frac{1}{2} \frac{W_{n-1}^*}{(n-1)s_Y^2} \left(\sum_{i=1}^n (y_i - \bar{y})^2 + \frac{(n-1)s_Y^2}{W_{n-1}^*} Z_1^{*2} \right) \\
&\sim -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log\left(\frac{(n-1)s_Y^2}{W_{n-1}^*}\right)
\end{aligned} \tag{27}$$

$$\begin{aligned}
& -\frac{1}{2} \frac{W_{n-1}^*}{(n-1)s_Y^2} \left((n-1)s_Y^2 + \frac{(n-1)s_Y^2}{W_{n-1}^*} W_1^* \right) \\
& \sim -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log((n-1)s_Y^2) + \frac{n}{2} \log(W_{n-1}^*) \\
& -\frac{1}{2}(W_{n-1}^* + W_1^*),
\end{aligned}$$

kde $W_{n-1}^* \sim \chi_{n-1}^2$ a $W_1^* \sim \chi_1^2$ sú nezávislé náhodné premenné s chi-kvadrát rozdelením s $n-1$ a 1 stupňom voľnosti. Nech $\ell_{(\alpha)}(\mu, \sigma | y_1, \dots, y_n)$ označuje α -kvanil rozdelenia náhodnej premennej $\ell(\tilde{\mu}, \tilde{\sigma} | y_1, \dots, y_n)$. Potom fiduciálna $(1-\alpha) \times 100\%$ konfidenčná oblasť pre parametre (μ, σ) je daná ako



Obrázok 5: Náhodný výber rozsahu $N = 5000$ zo združenej fiduciálnej distribúcie parametrov μ a σ založený na $n = [29, 1]$ nezávislých pozorovaniach hodnôt 0 a 1 na prístroji s rozlíšením $\delta = 1$. Realizácie výberu z fiduciálneho rozdelenia parametrov (μ, σ) boli usporiadané podľa hodnoty funkcie vierohodnosti s vyznačením empirickej $(1-\alpha) \times 100\%$ oblasti, pre rôzne hodnoty $\alpha \in (0, 1)$. Zobrazené krivky vznikli spojením bodov na hustej mriežke v parametrickom priestore (μ, σ) s rovnakou hodnotou funkcie vierohodnosti (na úrovni $\ell_{(\alpha)}(\mu, \sigma | x_1, \dots, x_n)$ pomocou hladkých kriviek (použitá bola funkcia `contour` v prostredí jazyka MATLAB, ktorá využíva kubickú interpoláciu, resp. vyhladzovanie pomocou splajnov).

$$CR_{(\mu, \sigma)}^{FD}(1 - \alpha) = \{(\mu, \sigma) : \ell(\mu, \sigma | y_1, \dots, y_n) > \ell_{(\alpha)}(\mu, \sigma | y_1, \dots, y_n)\}. \quad (28)$$

Táto konfidenčná oblasť je ekvivalentná s konfidenčnou oblasťou získanou invertovaním exaktného testu pomerom vierohodností (LRT - Likelihood Ratio Test), pozri [9, 21].

4.4 Približné konfidenčné intervaly pre parameter μ

Približný $(1 - \alpha) \times 100\%$ konfidenčný interval pre μ možno určiť priamo pomocou marginálnej fiduciálnej distribúcie parametra μ (v reálnych aplikáciách generujeme veľký počet, napr. $N = 10000$, nezávislých realizácií fiduciálnej náhodnej premennej $\tilde{\mu}$). Približný $(1 - \alpha) \times 100\%$ konfidenčný interval sa načastejšie určuje ako dvojstranný interval

$$CI_{\mu}^{FD}(1 - \alpha) := \langle \mu_{(\alpha/2)}, \mu_{(1-\alpha/2)} \rangle, \quad (29)$$

kde $\mu_{(\alpha/2)}$ a $\mu_{(1-\alpha/2)}$ označujú kvantily marginálnej fiduciálnej distribučnej funkcie pre parameter μ .

Avšak takýto prístup, t.j. dvojstranný $(1 - \alpha) \times 100\%$ konfidenčný interval, môže byť neadekvátny ak počet rôznych napozorovaných hodnôt m je $m = 1$, alebo $m = 2$: Napríklad, uvažujme situáciu keď $m = 2$ (napr. napozorovali sme hodnoty 0 a 1 pričom $n = [29, 1]$, pozri obr. 5). Potom pre ľubovoľne zvolenú hladinu významnosti α existuje parameter μ , blízko hranice 0.5 ($\mu < \approx 0.5$), a dostatočne malý parameter σ , že náhodný výber rozsahu n bude vo veľkej väčšine nadobúdať hodnoty rovné 0, a dvojstranný $(1 - \alpha) \times 100\%$ konfidenčný interval pre μ nepokryje túto skutočnú hodnotu parametra μ .

Na prekonanie tejto negatívnej vlastnosti navrhujeme modifikovaný približný $(1 - \alpha) \times 100\%$ konfidenčný interval pre parameter μ založený na marginálnej fiduciálnej distribúcii pre μ :

$$CI_{\mu}^{FDM}(1 - \alpha) := \begin{cases} \langle \mu_{(\alpha/2)}^*, \mu_{(1-\alpha/2)}^* \rangle, & \text{ak } x_{(n)} - x_{(1)} \leq \delta, \\ \langle \mu_{(\alpha/2)}, \mu_{(1-\alpha/2)} \rangle, & \text{inak,} \end{cases} \quad (30)$$

kde

$$\langle \mu_{(\alpha/2)}^*, \mu_{(1-\alpha/2)}^* \rangle = \langle \min \{ \mu_{(\alpha/2)}, x_{(n)} - \delta/2 \}, \max \{ \mu_{(1-\alpha/2)}, x_{(1)} + \delta/2 \} \rangle.$$

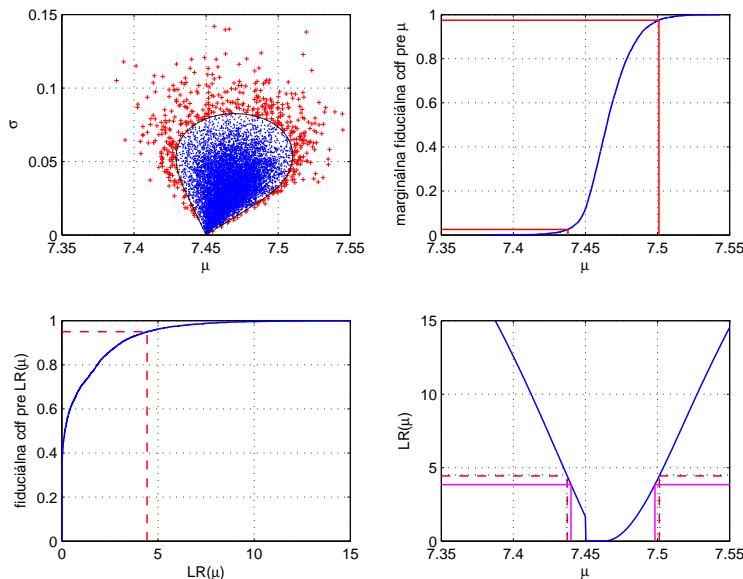
Pre malé rozsahy výberov, je možné urobiť modifikovanú verziu približného LR konfidenčného intervalu (13), nahradením (asymptoticky platného) kvantilu $\chi_1^2(1 - \alpha)$ kvantilom $q(1 - \alpha | X_1, \dots, X_n)$, definovaným pre každú realizáciu pozorovaných meraní (x_1, \dots, x_n) , na základe marginálnej fiduciálnej distribúcie pre parameter μ .

Približný, $(1-\alpha) \times 100\%$ konfidenčný interval pre parameter μ , určený LR metódou (*likelihood-ratio method*, resp. *profile likelihood method*), definujeme ako

$$CI_{\mu}^{LRE}(1-\alpha) := \{\mu : LR(\mu) < q(1-\alpha|X_1, \dots, X_n)\}, \quad (31)$$

kde $q(1-\alpha|x_1, \dots, x_n)$ označuje $(1-\alpha)$ -kvantil rozdelenia štatistiky $LR(\tilde{\mu})$ definovanej v (11), pričom $\tilde{\mu}$ má marginálne fiduciálne rozdelenie pri daných realizáciách meraní $(X_1, \dots, X_n) = (x_1, \dots, x_n)$.

Alternatívnu metódu na určenie kvantilu $q(1-\alpha|X_1, \dots, X_n)$ navrhli Lee a Vardeman, [11]. Ako aproximáciu kvantilu $q(1-\alpha|X_1, \dots, X_n)$ navrhli kvantil $q_n(1-\alpha)$, ktorý závisí iba od veľkosti náhodného výberu n



Obrázok 6: Približné 95%-konfidenčné intervaly pre parameter μ určené pomocou náhodného výberu rozsahu $N = 10000$ zo združenej fiduciálnej distribúcie parametrov μ a σ (vľavo hore) z digitalizovaných pozorovaní $x = [7.5, 7.5, 7.4, 7.5, 7.5, 7.4, 7.5, 7.5, 7.5, 7.4]$. Vpravo hore je ilustrovaná konštrukcia konfidenčného intervalu $CI_{\mu}^{FD}(1-\alpha)$, (29), $\alpha = 0.05$. V tomto prípade sa tento interval zhoduje s modifikovaným konfidenčným intervalom $CI_{\mu}^{FDM}(1-\alpha)$, (30). Vľavo dole je ilustrovaná metóda na určenie kvantilu $q(1-\alpha|x_1, \dots, x_n)$ na základe fiduciálnej distribúcie $LR(\tilde{\mu})$. Vpravo dole je ilustrovaná konštrukcia konfidenčného intervalu $CI_{\mu}^{LRE}(1-\alpha)$, ktorý je definovaný (31) a zobrazený pomocou prerušovanej čiary, v porovnaní s intervalom $CI_{\mu}^{LR}(1-\alpha)$, (13) - plná čiara, a $CI_{\mu}^{LRLV}(1-\alpha)$, (33) - bodkovaná čiara. Intervaly $CI_{\mu}^{LRE}(1-\alpha)$ a $CI_{\mu}^{LRLV}(1-\alpha)$ sú na obrázku nerozlíšiteľné. Odlíšu sa od intervalu $CI_{\mu}^{LR}(1-\alpha)$ určeného pomocou kvantilu $\chi_1^2(1-\alpha)$.

$$q_n(1 - \alpha) = n \ln \left(1 + \frac{F_{1,n-1}(1 - \alpha)}{n - 1} \right), \quad (32)$$

príčom $F_{1,n-1}(1 - \alpha)$ je $(1 - \alpha)$ -kvantil F -rozdelenia s 1 a $n - 1$ stupňami voľnosti. Teda,

$$CI_{\mu}^{LRLV}(1 - \alpha) := \{ \mu : LR(\mu) < q_n(1 - \alpha) \}. \quad (33)$$

Konštrukcia približných konfidenčných intervalov založených na fiduciálnej distribúcii ($CI_{\mu}^{FD}(1 - \alpha)$, $CI_{\mu}^{FDM}(1 - \alpha)$ a $CI_{\mu}^{LRE}(1 - \alpha)$ v porovnaní s intervalmi $CI_{\mu}^{LR}(1 - \alpha)$ a $CI_{\mu}^{LRLV}(1 - \alpha)$) je ilustrovaná pre špeciálny prípad meraní dĺžky pomocou digitálneho posuvného meradla, pozri [13, 17], na obrázku 6.

5 Zhrnutie

V príspevku sme diskutovali problémy spojené s konštrukciou konfidenčných oblastí a intervalov pre parametre regresnej priamky a parameter polohy na základe malých výberov z digitalizovaných meraní. Popísali sme niektoré známe a používané metódy konštrukcie približných konfidenčných intervalov pre parameter polohy. Medzi relatívne nové metódy patrí metóda založená na zovšeobecnenom fiduciálnej inferencii, ktorú sformalizoval J. Hannig v práci [6]. V príspevku sme ilustrovali, že v prípade špeciálnej voľby fiduciálnej distribúcie ide o metódu konštrukcie približnej konfidenčnej oblasti, ktorá je založená na funkcii vierohodnosti a možno nájsť užitočné prepojenie s metódou konštrukcie konfidenčných oblastí a intervalov založenou na štatistike podielu vierohodností (*likelihood ratio*) a profilovej funkcii vierohodnosti (*profile likelihood*). Podrobný popis algoritmu na generovanie výberu zo združenej fiduciálnej distribúcie a výsledky porovnania niektorých vybratých konfidenčných intervalov (na základe simulácii) možno nájsť v práci [20].

Literatúra

- [1] Cordero R., Seckmeyer G., Labbe F. (2006). *Effect of the resolution on the uncertainty evaluation*. Metrologia, **43**, L33–L38.
- [2] Dempster P., Rubin D.B. (1983). *Rounding error in regression: The appropriateness of Sheppard's corrections*. Journal of the Royal Statistical Society, Series B (Methodological) **45**, 51–59.
- [3] Elster C. (2000). *Evaluation of measurement uncertainty in the presence of combined random and analogue-to-digital conversion errors*. Measurement Science and Technology, **11**, 1359–1363.
- [4] Fisher R.A. (1935). *The fiducial argument in statistical inference*. Annals of Eugenics, **6**, 391–398.
- [5] Frenkel R.B., Kirkup L. (2005). *Monte Carlo-based estimation of uncertainty owing to limited resolution of digital instruments*. Metrologia, **42**, L27–L30.

- [6] Hannig J. (2008). *On generalized fiducial inference*. Statistica Sinica, submitted. See also the working paper at http://www.stat.colostate.edu/research/2006_3.pdf.
- [7] Hannig J., Iyer H.K., Patterson P. (2006). *Fiducial generalized confidence intervals*. Journal of the American Statistical Association, **101**, 484–499.
- [8] Hannig J., Iyer H.K., Wang C.M. (2007). *Fiducial approach to uncertainty assessment accounting for error due to instrument resolution*. Metrologia, **44**, 476–483.
- [9] Chvosteková M., Witkovský V. (2009). *Exact likelihood ratio test for the parameters of the linear regression model with normal errors*. In: MEASUREMENT 2009. The 7th International Conference on Measurement, May 20 - 23, 2009, Smolenice Castle, Slovakia.
- [10] Kubáček L. (1965). *Ocenka točnosti izmerenija pri eksperimentálnych metodach*. Izmeriteľná technika, **3**, 6–11.
- [11] Lee C-S., Vardeman S.B. (2001). *Interval estimation of a normal process mean from rounded data*. Journal of Quality Technology, **33**, 335–348.
- [12] Lee C-S., Vardeman S.B. (2002). *Interval estimation of a normal process standard deviation from rounded data*. Communications in Statistics – Simulation and Computation, **31**, 13–34.
- [13] Lira I.H. (2006). *Resolution revisited*. Metrologia, **43**, L14–L17.
- [14] Lira I.H., Wöger, W. (1997). *The evaluation of standard uncertainty in the presence of limited resolution of indicating devices*. Measurement Science and Technology, **8**, 441–443.
- [15] Morales D., Pardo L., Vajda I. (2006). *On efficient estimation in continuous models based on finitely quantized observations*. Communications in Statistics (Theory and Methods), **35**, 1629–1653.
- [16] Sheppard W. (1898). *On the calculation of the most probable values of frequency constants for data arranged according to equidistant divisions of a scale*. Proceedings of the London Mathematical Society, **29**, 353–380.
- [17] Taraldsen G. (2006). *Instrument resolution and measurement accuracy*. Metrologia, **43**, 539–544.
- [18] Vardeman S.B., Lee C-S. (2005). *Likelihood-based statistical estimation from quantized data*. IEEE Transactions on Instrumentation and Measurement, **54**, 409–414.
- [19] Willink R. (2007). *On the uncertainty of the mean of digitized measurements*. Metrologia, **44**, 73–81.
- [20] Wimmer Jr. G. 2008. *Metódy výpočtu približných konfidenčných intervalov parametra polohy z digitalizovaných meraní*. In: 15. letná škola JČMF ROBUST 2008, 8.-12.9. 2008, Hotel Mier, Račková dolina, Pribylina.

- [21] Witkovský V., Chvosteková M. (2009). *Simultaneous tolerance intervals for the linear regression model*. In: MEASUREMENT 2009. The 7th International Conference on Measurement, May 20 - 23, 2009, Smolenice Castle, Slovakia.

PodĎakovanie: Práca bola podporená z projektov VEGA 1/3016/06, APVV RPEU-0008-06 a MMT R LC06024.

Adresa: V. Witkovský, Ústav merania SAV, Dúbravská cesta 9, 84104 Bratislava; G. Wimmer, Fakulta prírodných vied, Univerzita Mateja Bela, Banská Bystrica, Přírodovědecká fakulta, Masarykova univerzita Brno, Matematický ústav SAV, Bratislava

E-mail: witkovsky@savba.sk, wimmer@mat.savba.sk

SHLUKOVÁNÍ VELKÝCH SOUBORŮ DAT POMOCÍ ALGORITMŮ k-PRŮMĚRŮ

Marta Žambochová

Klíčová slova: Shlukování, velké soubory dat, varianty algoritmu k-průměrů.

Abstrakt: Článek obsahuje popis a porovnání několika obměn metod postavených na základě k-means algoritmu. Existuje totiž velké množství variant základních metod shlukové analýzy, neexistuje však dostatečné srovnání vhodnosti použití jednotlivých alternativ. Příspěvek je zaměřen především na postupy, které nejsou obsaženy v žádných statistických softwarových systémech. V práci jsou vyzdvíženy jejich rozdíly oproti základní metodě. Jednotlivé algoritmy byly naprogramovány ve vývojovém prostředí MATLAB. Kritériem pro jejich vzájemné srovnání byla vybrána hodnota vnitroskupinové variability.

Abstract: This article contains a description and comparison of several variations of methods built on the basis of k-means algorithm. There is a large number of varieties of basic clustering techniques, without some sufficient comparison of how to use the alternatives. My contribution is focused primarily on procedures that are not included in any statistical software system. The work tries to highlight the differences against the underlying method. Individual algorithms are programmed in the MATLAB environment. The value of intragroup variability was chosen as the criterion for their mutual comparison.

1 Úvod

Shluková analýza je pojem, který zahrnuje celou řadu metod, jejichž cílem je roztřídění množiny objektů do předem nespécifikovaných skupin, označených jako shluky. Objekty uvnitř shluků mají být maximálně podobné a objekty patřících do různých shluků podobné minimálně.

Velmi oblíbenou skupinou algoritmů shlukové analýzy jsou metody hierarchického shlukování, které jsou v různých variantách implementovány ve všech statistických softwarových systémech. Popis těchto postupů nalezneme například v [2] nebo v [6]. Pro svoji jednoduchost, přehlednost a dobrou interpretovatelnost výsledků jsou používány ve většině běžných průzkumů zabývajících se analýzou dat. Mají ale i jednu velmi nepříjemnou nevýhodu. Tou je špatné, nebo dokonce v některých případech i nemožné, použití těchto metod pro analýzu souborů obsahujících velké množství sledovaných objektů. V učebnicích statistiky se velkým souborem nazývá již soubor obsahující 250 objektů. S rozvojem výpočetní techniky se situace poněkud vylepšila. Pomocí těchto metod lze zpracovat i soubory obsahující přes tisíc objektů. Při analýze dat je však mnohdy nutno zpracovat i mnohem větší soubory.

Jinou alternativou jsou metody rozkladu. Většina postupů z této skupiny je obstojně použitelná pro mnohem větší soubory dat než metody hierarchické. Oblast metod rozkladu je velmi široká a je zde zahrnuto mnoho algoritmů na velmi rozdílných principech. Základem však je vždy hledání nejlepšího shlukování pomocí prvotního rozdělení objektů do shluků a následného postupného přerozdělování objektů ve shlucích.

Jednou velkou podskupinou metod rozkladu jsou algoritmy k -průměrů. Jedná se o optimalizační metody. Základní myšlenkou je hledání rozkladu objektů do předem daného počtu shluků, pro který je součet vzdáleností jednotlivých objektů od centra jejich shluku minimální (tj. hledání minima funkce $Q = \sum_x \|x - c(x)\|^2$.)

V literatuře je uvedeno pod různými názvy mnoho variant základního postupu. Mimo jiné jsou to Forgyova, Janceyova, Llyodova, MacQueenova, Wishartova metoda, ale i mnoho dalších.

2 Postup metody k -průměrů

Algoritmus k -průměrů využívá při svém průběhu výpočtu tzv. centroidů (středů, těžišť). Základní algoritmus k -průměrů probíhá v následujících krocích:

- 0) prvotní rozdělení souboru dat do k shluků,
- 1) výpočet centroidů všech shluků,
- 2) přiřazení všech objektů k centroidům,
- 3) pokud došlo ke změně oproti předchozí iteraci, návrat na 1).

Hlavními výhodami metody k -průměrů jsou především:

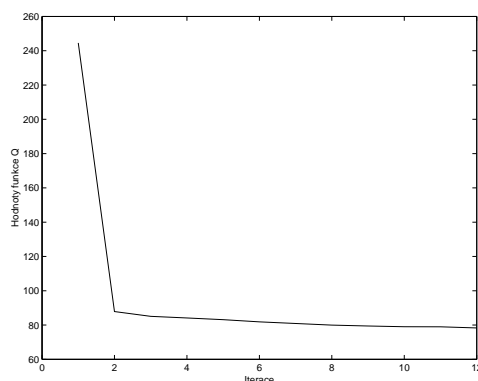
- jednoduchý princip,
- přijatelná rychlost (a tím použitelnost pro velké soubory dat),
- relativně dobré výsledky (vzhledem k minimalizaci vnitroskupinové variability).

Mezi nevýhody metody k -průměrů patří:

- nutnost zadání požadovaného počtu shluků,
- hledání pouze lokálního optima,
- hledání pouze konvexních shluků,
- velká časová náročnost pro obzvláště velké soubory,
- silný vliv inicializačního rozdělení.

Problém nalezení pouze lokálního optima je dán principem algoritmu, a proto jej nelze odstranit. Obdobně i schopnost nalezení pouze konvexních shluků. Problém nutnosti zadání požadovaného počtu shluků je možno řešit vhodným předzpracováním dat pro určení optimálního počtu shluků.

Alternativy metody k -průměrů se zaměřují na odstranění nebo alespoň zmírnění zbývajících dvou nevýhod.

Obrázek 1: Vývoj hodnot optimalizované funkce Q .

3 Varianty zkracující běh algoritmu

Na obrázku 1 je patrný velmi rychlý počáteční pokles hodnot optimalizované funkce Q a nepatrný pokles ve vyšších iteracích. Algoritmus umožňuje nalezení pouze lokálního minima, proto již není velkým ústupkem od přesnosti výsledku, pokud ukončíme běh algoritmu po dostatečném počtu iteračních kroků, ještě v okamžiku, kdy se hodnota funkce Q stále nepatrně snižuje. Toto řešení může znamenat úsporu času při zpracování velmi rozsáhlých souborů dat. Nejčastější variantou je ukončení algoritmu udáním mezní hodnoty zlepšení optimalizované funkce Q , při které se má ještě pokračovat ve výpočtu. V některých variantách (například implementace metody k -průměrů ve statistickém systému STATISTICA) se výpočet ukončuje omezením počtu iterací.

4 Varianta využívající mrkd-stromů

Daleko výraznější obměnou je algoritmus využívající mrkd-stromů. Kd-strom (kvadrantový strom) je datová stromová struktura, která reprezentuje rekurzivní dělení konečné množiny bodů z prostoru dimenze d na k částí (hyperkvádrů dimenze d), pomocí ortogonálních nadrovin dimenze $d - 1$ (např. pomocí rozdělení ortogonálně k nejdelší straně hyperkvádrů na úrovni mediánu ze všech bodů hyperkvádrů). Mrkd-stromy jsou binární formou kd-stromů. Hlavní myšlenkou, které úspěšně využívá tento algoritmus je fakt, že každý z vrcholů obsahuje informaci o všech datových objektech z příslušného hyperkvádrů. Pokud je celý hyperkvádr nejbližší k jednomu z center, přísluší všechny datové objekty obsažené v tomto hyperkvádrů do shluku příslušného k tomuto centru. V jednotlivých krocích výpočtu můžeme mnoho početních i rozhodovacích úkonů s jednotlivými datovými objekty nahradit jediným hromadným úkonem a tím se velmi sníží počet přístupů k datům a aritmetických operací. Algoritmus využívající mrkd-stromy je

velkým zefektivněním klasického přístupu metody k-průměrů. Efektivita tohoto algoritmu spočívá ve faktu, že mrkd-strom je zkonstruovaný pouze jedenkrát pro dané datové objekty a celá struktura nemusí být přepočítávána v každém iteračním kroku algoritmu.

V algoritmu popsáném v [5], [3] či [9] je dělení hyperkvádrů prováděno na úrovni mediánu ze všech datových objektů obsažených v daném hyperkvádrů. V testovacím programu se tento krok jevil poněkud pomalým (možná použitím nevhodného třídícího algoritmu). Z tohoto důvodu bylo použito dělení na úrovni aritmetického průměru. Tato změna oproti originální verzi přinesla velké urychlení v této fázi programu. Přinesla bohužel nerovnoměrné rozdělení datových bodů do jednotlivých podkvádrů. Pro některé ze zkušebních souborů byl tento fakt dosti znatelný. Některé ze vznikajících hyperkvádrů byly velmi přeplněné, v některých zbylo jen velmi málo datových objektů. Závažnost tohoto nedostatku se zatím nepodařilo lépe prozkoumat z důvodu nedostatečného množství různorodých souborů dat.

5 BIRCH algoritmus k-průměrů

Jednou z variant algoritmu k-průměrů je i BIRCH algoritmus k-průměrů. Jedná se o alternativu algoritmu BIRCH, popsáného v [7] či [8]. Algoritmus má dva parametry, a to povolený počet objektů v libovolném shluku a hranici variability libovolného shluku ("poloměr" shluku). Na rozdíl od klasického algoritmu BIRCH nevytváří CF-strom.

Postup BIRCH algoritmu k-průměrů probíhá ve třech fázích.

- 1) V cyklu se vždy vybere objekt z množiny objektů a nalezne se "nejbližší" shluk z množiny shluků, ve kterém se přidáním objektu nepřekročí hranice pro počet prvků ve shluku ani hranice variability shluku. Pokud takovýto shluk neexistuje, vytvoří se pro objekt nový shluk, který obsahuje pouze tento objekt po zařazení objektu do shluku se objekt vymaže z množiny všech objektů. Cyklus se provádí do vyprázdnění množiny objektů.
- 2) Pomocí speciální varianty kvadratické metody k-průměrů jsou shlukovány centroidy všech shluků, které vznikly ve fázi první.
- 3) Všechny původní objekty jsou rozřazeny do shluků, každý objekt je přiřazen k nejbližšímu z centroidů vzniklých ve druhé fázi.

Výsledek algoritmu je velmi závislý na volbě obou parametrů. Při volbě extrémně nízké hodnoty parametru pro povolenou variabilitu shluků se v první fázi algoritmu vytvoří relativně vysoký počet prvotních shluků, které poté vstupují do dalšího zpracování. Pro velmi velké původní datové soubory to může vést k nedostatečné redukci objektů a tím i časové náročnosti algoritmu. Naopak při volbě vysokých hodnot parametru pro povolenou variabilitu, zvláště v kombinaci s volbou velké hodnoty parametru pro povolený počet objektů ve shlucích, vzniká v první fázi algoritmu relativně málo shluků

vstupujících do dalšího zpracování a výsledek druhé a třetí fáze vede k dosti nepřesným výsledkům.

6 Algoritmus k-průměrů⁺⁺

Jednou z variant je i algoritmus k-průměrů⁺⁺, který jeho autoři popsali ve článku [1]. Tato alternativa se snaží vylepšit výsledek shlukování vytvořením speciálního inicializačního rozdělení do shluků. Prvotní centroidy jsou vybírány postupně. První centroid je vybrán náhodně ze všech datových objektů. Další centroidy jsou vybírány jeden po druhém, vždy dle pravděpodobnosti $\frac{D(x^*)^2}{\sum_x D(x)^2}$, kde x^* je zkoumaný objekt, $D(x)$ je nejkratší vzdálenost datového objektu x od nejbližšího centroidu ze všech doposud vybraných.

Dále probíhá standardní algoritmus k-průměrů.

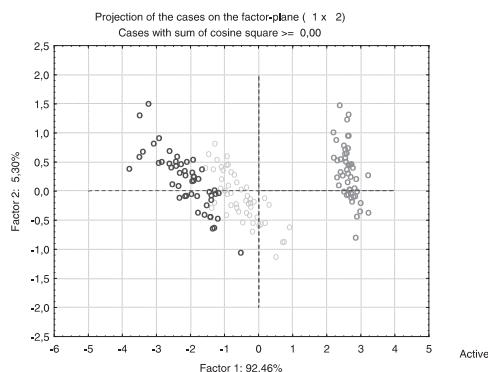
7 Inkrementální metoda k-průměrů

Základní algoritmus inkrementální metody k-průměrů využívá ve svém postupu přesunu jednoho vybraného objektu ze stávajícího shluku do jiného tak, aby došlo k co největšímu poklesu hodnoty optimalizované funkce Q .

V knize [4] je popsán jednak tento základní algoritmus, tak i algoritmus kombinující základní metodu k-průměrů s inkrementální metodou k-průměrů. Tento upravený algoritmus byl vybrán pro testování.

8 Shrnutí

Vybrané alternativy metody k-průměrů byly naprogramovány ve vývojovém prostředí MATLAB. Pro názorné prozkoumání některých vlastností jednotlivých algoritmů byla vybrána aplikace algoritmů na známá data IRIS z [10]. Soubor obsahuje 150 objektů, každý objekt je popsán pomocí čtyř nume-



Obrázek 2: Rozptylový graf komponentního skóre.

rických proměnných (jednotlivé rozměry kališních a korunních lístků květů). Do souboru jsou zařazeny květy tří různých druhů z rodu iris, z každého druhu 50 zástupců. Z průzkumové analýzy je zřejmé, že jeden z druhů se výrazněji odlišuje v popsáních atributech od zbývajících dvou. Ty se sice vzájemně také odlišují, ale ne tak výrazně, dokonce dochází k prolínání. Vše je názorně vidět na obrázku 2. Srovnání výsledků shlukování pomocí vybraných metod je uvedeno v tabulce 1.

Metoda	Počet chybně zařazených objektů ve 2.shluku	Počet chybně zařazených objektů ve 3.shluku	Výsledná hodnota funkce Q
k-průměrů (náhodné počáteční shlukování)	20	46	142,8593
k-průměrů (STATISTICA 6.0)	2	15	77,4107
Kvadratická k-průměrů	3	14	78,2451
Inkrementální k-průměrů	1	14	77,2215
k-průměrů ⁺⁺	1	13	77,0924
k-průměrů (mrkd-stromy)	1	14	77,2215
BIRCH k-průměrů (param. 50; 0,9)	17	1	78,3159
BIRCH k-průměrů (param. 50; 0,2)	4	11	77,6817
BIRCH k-průměrů (param. 50; 0,1)	2	15	77,4107
BIRCH k-průměrů (param. 20; 0,9)	objekty výrazně odlišného druhu (+3 další objekty) byly rozděleny do 1. a 2. shluku, 3. shluk obsahuje zbylých 97 objektů druhých druhů		145,3234
BIRCH k-průměrů (param. 20; 0,1)	2	15	77,4107

Tabulka 1: Srovnání výsledků shlukování pomocí vybraných metod.

9 Závěr

Zaměřila jsem se na popis a otestování několika vybraných variant algoritmu k -průměrů. Většina testů dopadla příznivě pro zkoumané alternativy. Nelze ale jednoznačně určit, který z postupů je lepší. Existuje ještě řada dalších variant, které by bylo dobré prozkoumat. Další šetření bude zaměřeno na testování jednotlivých alternativ algoritmu na speciálně vygenerovaných velkých souborech dat s cílem dokonalejšího porovnání. Konečným cílem je doporučit nejvýhodnější způsob shlukování podle charakteru dat.

Literatura

- [1] Arthur D., Vassilvitskii S. (2007). *k-means++ The advantages of careful seeding*. Symposium on Discrete Algorithms (SODA), New Orleans, Louisiana, 1027–1035.
- [2] Hebák P., Hustopecký J., Pecáková I., Průša M., Řezanková H., Svobodová A., Vlach P. (2005). *Vícerozměrné statistické metody (3)*. Praha, Informatorium.
- [3] Kanungo T., Mount, D.M., Netanzahu N.S., Piatko CH.D., Silverman R., Wu A.Y. (2002). *An efficient k-means clustering algorithm: analysis and implementation*. Proc ACM SIGKDD Int'l Conf. IEEE Transactions on pattern analysis and machina inteligence, **24**, 7.
- [4] Kogan J. (2007). *Introduction to clustering large and high-dimensional data*. New York, Cambridge University Press.
- [5] Pelleg D., Moore A. (1999). *Accelerating exact k-means algorithms with geometric reasoning*. Proc ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining, 277–281.
- [6] Řezanková H., Húsek D., Snášel V. (2007). *Shluková analýza dat*. Praha, Professional Publishing.
- [7] Zhang T., Ramakrishnan R., Livny M. (1996). *An efficient data clustering method for very large databases*. ACM SIGMOD Record, **25**, 2, 103–114.
- [8] Zhang T., Ramakrishnan R., Livny M. (1997). *BIRCH: A new data clustering algorithms and its applications*. TJournal of Data Mining and Knowledge Discovery, **1**, 2, 141–182.
- [9] Žambochová M. (2006). *Rozhodovací stromy a mrkd-stromy v analýze dat*. Sborník prací účastníků vědeckého semináře doktorského studia FIS VŠE, Praha, 196–202.
- [10] <http://archive.ics.uci.edu/ml/datasets/>.

Adresa: M. Žambochová, FSE UJEP, KMS, Moskevská 54, CZ- 400 96, Ústí nad Labem

E-mail: marta.zambochova@ujep.cz

TESTING HOMOGENEITY AND GOODNESS OF FIT IN SURVIVAL DATA

Jana Timková

Keywords: Survival analysis, counting process, MCMC, goodness of fit, homogeneity test, residual analysis.

Abstract: The present paper deals with the goodness of fit and the two-sample problem testing related to the event-history type data. The proposed graphical methods are related to the bayesian nonparametric approach and take advantage of MCMC estimation of the hazard rate. The first technique of the homogeneity testing is based on the method of Arjas [2] who estimated the hazard rate under the null hypothesis (all individuals are governed by the same rule) and then plotted the estimated cumulative hazard rate against the number of failures in one group and in the other group separately. The second method to detect the inhomogeneity in two groups uses the estimation of the hazard rate based on the first - control - group and compare it to the second one. The article conclude with a concrete application of the methods on a vaginal cancer mortality data.

Abstrakt: Súčasný článok sa zaoberá problémom testu dobrej zhody a dvojvýberovým testom v oblasti dát prežitia. Predložené metódy sú spojené s bayesovským neparametrickým prístupom a MCMC odhadmi pre funkciu hazardu. Prvý test homogeneity vychádza z článku Arjas [2], v ktorom je odhadnutá funkcia hazardu za predpokladu platnosti nulovej hypotézy, čiže všetky objekty sa správajú rovnako nezávisle na rozdelení do skupín. Následne sa vykreslí odhadnutá kumulatívna funkcia hazardu proti kumulatívne počtu udalostí v oboch skupinách zvlášť. V druhom prístupe odhadneme funkciu hazardu v jednej skupine a podobne ako v predchádzajúcom porovnáme s druhou. Článok je zakončený konkrétnou aplikáciou na reálnych dátach.

1 Introduction

The aim of the analysis of life history data is to describe the behaviour of the underlying processes what is usually done by assuming a suitable model for the hazard rate. When the model is fitted, the next step is the assessment of quality of the estimation and that leads us to the goodness of fit testing. Over last decades many powerful and convenient methods of detecting, whether such a model is unsatisfactorily specified, arised. Particularly in case of presence of some covariates, the inference is related to the form and significance of the dependence of the observed times on those covariates. The known methods are mostly based on the martingal properties of the differences between observed cumulative number of failures and the estimated cumulative hazard rate.

Let us consider n parallel mutually independent counting processes $N_1(t)$, $N_2(t)$, \dots , $N_n(t)$ observed in the time interval $[0, T]$. We assume that all processes start from zero, $N_i(0) = 0$, and that $N_i(t)$ increases in $+1$ when the i -th object happens to meet an event of interest. The i -th process is expected to behave according to an intensity process $I_i(t) \cdot \lambda(t)$ where $\lambda(t)$ is a bounded nonnegative continuous hazard function and $I_i(t)$ is an indicator process (indicating whether the i -th individual is at risk of event, i.e. $I_i(t) = 1$ when it is at risk, $I_i(t) = 0$ otherwise). The indicator process has its importance when the censoring is present or when the occurrence of event implies the end of the observing of the object. As the result we obtain the multivariate counting process $\mathbf{N}(t) = (N_1(t), N_2(t), \dots, N_n(t))^T$.

The hazard rate related to i -th object, $\lambda_i(t)$, is the instantaneous rate of an event occurring at time t defined as

$$\lambda_i(t) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} P(N_i(t + \Delta t) - N_i(t) = 1 | \mathcal{F}_t),$$

where $\mathcal{F}_t := \sigma\{N_i(s), 0 \leq s \leq t\}$ is σ -algebra of the history of the i -th object up to time moment t . In next we will consider a special case of previous setting - a typical survival data with one possible event on each object (either death or censoring).

Now, let us introduce the generalized residual at given time t as

$$X_i(t) = N_i(t) - \int_0^t I_i(u) \lambda_i(u) du$$

for all i . Then the process $(X_i(t), t \in [0, T])$ is an \mathcal{F}_t -martingale and is called the residual process for i -th individual. Let us have t_1, \dots, t_n observed times of events. By considering $X_i(t)$ at the ordered set of uncensored failure times $t_{(1)}, t_{(2)}, \dots, t_{(K)}$, where K is the total number of uncensored failures, we obtain

$$X_i(t_{(k)}) = N_i(t_{(k)}) - \int_0^{t_{(k)}} I_i(u) \lambda_i(u) du, \quad k = 1, \dots, K.$$

an discrete time martingale of parameter k . In fact, the integrated intensity process approximates the number of uncensored events observed on the individual. Hence, the motivation for the testing technique is straightforward. Let us have the cumulative hazard rate $\Lambda_i(t) = \int_0^t \lambda_i(u) du$ and after denoting $\bar{X}(t) = \sum_{i=1}^n X_i(t)$ and $\bar{N}(t) = \sum_{i=1}^n N_i(t)$,

$$\bar{X}(t_{(k)}) = \bar{N}(t_{(k)}) - \sum_{i=1}^n \int_0^{t_{(k)}} I_i(u) \lambda_i(u) du = k - \sum_{i=1}^n \Lambda_i(t_{(k)} \wedge t_i),$$

is close to zero for all $k \in \{1, \dots, K\}$. Now we can substitute the unknown hazard rate with the appropriate estimation and use it for testing the goodness of fit as well as homogeneity among groups.

Naturally, the martingale property of the residual process is not always preserved after plugging-in the estimated hazard rate but there are cases in which it is true. There is certain number of publications regarding the topic in various models of hazard rate; see e.g. excellent [7] for the generalized residuals in the Cox model, further [11] in Aalen regression model and [6] in Cox-Aalen model. Apparently, once the estimated residual process is a martingale, the central limit theorem for martingales (cf. [4]) gives the weak convergence to zero-mean Gaussian process. This is the way to obtain the confidence bands around the estimated residual process needed for tests, although often tedious and in some cases intractable because of complicated covariance structure of the limiting process.

The alternative which introduces itself is to take the bayesian approach for modelling and estimating, gain the sample of the estimated hazard rate via MCMC algorithms and use the pointwise credibility bands for the estimated residual process. There were developed many bayesian models for hazard rate, for a good overview see [8] or more detailed [9].

Useful graphical methods based on the residual processes were introduced by e.g Andersen [1] who considered various plots mainly focused on the checking the proportional-hazards assumption of Cox model, and later by Arjas [2] who came up with the plot of the estimated cumulative hazard against number of failures in stratas, in which the observed individuals were splitted.

In next sections, a bayesian model for the hazard rate will be introduced and the testing techniques for two-sample problem will be carried out. The results will be presented in example of a real data analysis.

2 Model and estimation of the hazard rate

Here we describe one possible model which can be used in bayesian setting. Let us for now assume that all observed objects are governed by the intensity processes with the same hazard rate $\lambda(t)$ without taking any covariate process into account. To estimate the hazard rate we use nonparametric method firstly proposed by Arjas and Gasbarra [3]. They assumed the piecewise constant function for the hazard rate. All the individuals were expected to behave according to a common hazard rate $\{\lambda_t, t \in [0, T]\}$. Further in their model the hazard function was constant within the $(m + 1)$ intervals which came from dividing the whole time interval $[0, T]$ by m jump times T_1, \dots, T_m . The level of hazard function within the interval $(T_{j-1}, T_j]$ was denoted as λ_j . The number of jump times m could vary. To reach the estimates the MCMC were used.

Let us denote $H = (\lambda_1, T_1, \dots, \lambda_m, T_m, \lambda_{m+1})$. The main idea is to provide as many iterations as necessary in which the piecewise constant hazard function defined by vector of parameters H is upgraded according to the data and prior information. As soon as we possess the Markov chain $(H^{(i)})_{i=1}^R = (\lambda_1^{(i)}, T_1^{(i)}, \dots, \lambda_{m^{(i)}}^{(i)}, T_{m^{(i)}}^{(i)}, \lambda_{m^{(i)}+1}^{(i)})_{i=1}^R$, where R is the number

of provided iteration and parameters denotes with $^{(i)}$ are those connected with i -th iteration and we are ensured about the ergodicity of the chain, we could compute the resulting estimate of the hazard function as the mean of the members of the chain. Finally, even though the simulated hazard rates are piecewise constant functions, asymptotically the mean of them yields to a continuous function.

Except few changes we adopt the similar structure of the model for the hazard rate. The chosen prior distribution of the levels of hazard rate λ_i , $i = 1, \dots, m + 1$ is gamma distribution with particular setting of parameters dependent on parameter α to fold in the martingale structure. The jump times are distributed as the homogeneous Poisson process in $[0, T]$ with unknown constant rate μ . Further, we extended Arjas and Gasbarra's model by accompanying the hierarchical structure into model. For details about the estimation and the discussion on the ergodicity of the chain see [10].

3 Testing

As it was announced before the main focus of the paper is in a simple two-sample problem which occurs when the observed objects come from two different environments. Also more complicated cases could be transformed to this situation by dividing the individual processes in two stratas according to certain values of covariates or the observed times themselves (as it was done in [2, sec. 3.3]).

3.1 Arjas's plot

Arjas [2] suggested to provide the estimation under the assumption that the individuals are divided in two or more stratas according to some covariates observed alongside (more clearly, in his case he worked with the Cox model with p covariates and wanted to check whether $(p + 1)$ -th covariate should be included in model). We accommodate his approach in our simple case of two stratas. The null hypothesis for us is that the observed individuals (i.e. the counting processes) behave in the same way in both stratas. We provide the estimation of the hazard rate according to the null hypothesis what means to use all observations from all objects together, as if they were replicates of each other, to estimate the hazard rate. Once we have the estimated hazard rate $\hat{\lambda}_i(t) \equiv \hat{\lambda}(t)$, $0 \leq t \leq T$, for all i , we can compute the estimated cumulative hazard function for $0 \leq t \leq T$ in the first strata as follows

$$\hat{H}_{(strata1)}(t) = \sum_{i \in strata1} \int_0^t I_i(s) \hat{\lambda}(s) ds \quad (1)$$

and analogically in the second one. Now we base plotting technique on the zero-mean martingale property of difference between observed numbers of failures and the real cumulative hazard function. Let $t_{(1)} \leq t_{(2)} \leq \dots \leq t_{(N)}$ be ordered observed times of failures. The idea is to plot $\hat{H}_{(strata1)}(t_{(i)})$

and $\hat{H}_{(strata2)}(t_{(i)})$ against the observed cumulative number of failures in the first strata and the second strata respectively. In case the null hypothesis of homogeneity of both groups is correct, both lines should be placed near the diagonal line $f(x) = x$. If the lines are far from each other and from the diagonale then the null hypothesis is rejected. The significant departure from the null hypothesis is proposed to be detected by the pointwise 95% credibility bands created from the sample of $\hat{H}_{strata1}$ calculated from the posterior sample of the hazard rate.

3.2 Two-sample test

The basis of the second proposed approach is to provide the estimation of the hazard rate from the observations belonging to one - control - group of objects. Then we calculate the estimated cumulative hazard function $\hat{H}_{(strata1)}$ accordingly to (1) and plot $\hat{H}_{(strata1)}(t_{(i),2})$ against the observed cumulative number of failures in the second strata. In the previous we used the notation $t_{(1),2} \leq t_{(2),2} \leq \dots \leq t_{(N_2),2}$ for the observed times of failures occurring in the second group of objects with the total number of observations denoted as N_2 . If both groups of objects possess the same behaviour, the line is near the imaginary line $f(x) = x$. Again, to detect a significant departure from the null hypothesis we use the simulated pointwise confidence bands drawn from the MCMC simulation output.

Apart from the homogeneity testing, the introduced idea is plausible for testing of overall goodness of fit. Similarly, as in (1) for stratas, we can calculate $\hat{H}(t) = \sum_{i=1}^n \int_0^t I_i(s) \hat{\lambda}(s) ds$ for whole dataset. Plotting the values of $\hat{H}(\cdot)$ in observed uncensored failure times versus number of observed failures is close to diagonale $f(x) = x$ if the fit is sufficient.

4 Application

In following the application on a real data is presented. The data set was analyzed in [5]. The observations give the times from insult with the carcinogen DMBA to mortality from vaginal cancer in rats. Two group were distinguished by a pretreatment regimen. We want to compare the impact of different regimens on the waiting times to death. The data represents the usual time-to-event setting with one observation (failure or censored event) belonging to each individual.

First, the hazard rate was estimated separately for each group as well as for the whole set of observations. The estimated hazard rates are plotted in figure 1. In the bottom part of the figure the observed times of events for both groups are displayed, the failures are denoted with \circ and censored events with $+$. The upper row of observations is from the first group and bellow is the row of the second group's observations. Figure 2 shows the plots related to the estimated cumulative hazard rate in the first group which acts as the control group here. In the left side the goodness of fit of the estimated hazard

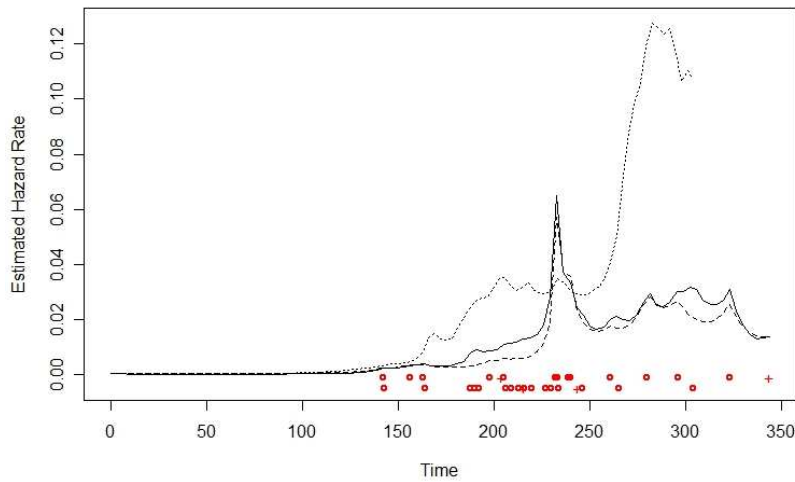


Figure 1: The plot of the estimated hazard rate in the whole set of observations (solid line) and in the two groups separately (dashed lines). At the bottom of the figure the data are displayed in two lines - the observations from the first group in the upper level and from the second group below.

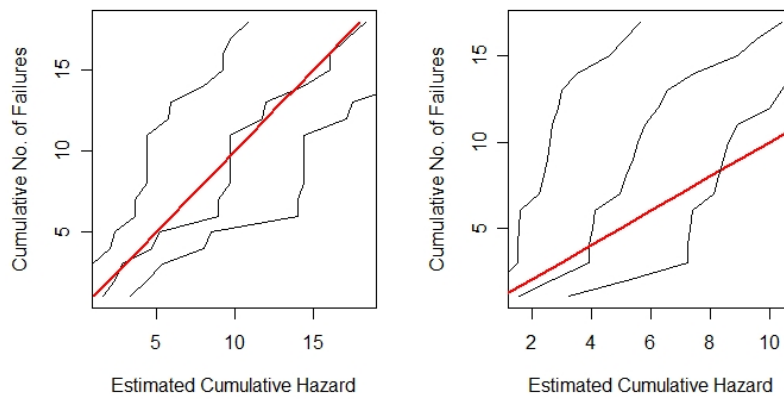


Figure 2: Plots of the cumulative estimated hazard rate against the cumulative number of failures; in the left side the goodness of fit test related to the control group, the right-hand plot displays the test of homogeneity of both groups. The 95% pointwise credibility bands are included.

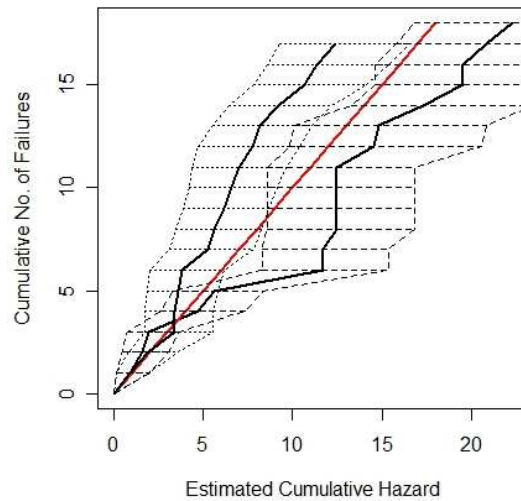


Figure 3: Arjas's plot of the homogeneity of two groups. The 95% pointwise credibility bands are included.

rate to the control group is shown and the right-hand plot displays the test of homogeneity described in the section 3.2. Further, in figure 3 Arjas's plot is presented. Obviously, we can reject the hypothesis of homogeneity of mortality in two groups of rats.

The parameters chosen for the MCMC estimation of the hazard rate was $\mu_a = 10$, $\mu_b = 100$, $\alpha_a = 1000$, $\alpha_b = 100$, $\alpha_0 = 1$ and $\beta_0 = 10000$. The number of iterations was 5000 and first 1000 were erased.

5 Conclusion and discussion

The main idea of the previous pages was to think of alternative way of assessing the goodness of fit and homogeneity of two samples in survival data. The proposed methods are related to bayesian methodology and work with a sample of simulated hazard rates from MCMC procedure. The advantage is concentrated in the fact that existence of the sample allows us to create the testing machinery based on the calculated pointwise credibility bands instead of the rather cumbersome asymptotics. This type of approach can also be a good starting point for testing in more complicated models in which the key martingale property after plugging-in the estimator is lost.

The disadvantage of the method is that the result of the test depends also on the precision of the estimation which is connected with the choice

of parameters. Theoretically, as long as the graphical goodness of fit test approves the fit of the estimated hazard rate to the observed data of the control group, the test of the homogeneity of the second group with the control group should be valid. Nevertheless, as the results of test may differ among the group of the approved fits, it is important to decide which from those estimators suits the data and situation the best and accordingly to this provide the test.

References

- [1] Andersen P.K. (1982). *Testing goodness of fit of Cox's regression and life model*. Biometrics **38**, 67–77.
- [2] Arjas E. (1988). *A graphical method for assessing goodness of fit in Cox's proportional hazard models*. Journal of the American Statistical Association **83**, 401, 204–212.
- [3] Arjas E., Gasbarra D. (1994). *Nonparametric Bayesian inference from right censored survival data, using Gibbs sampler*. Statist. Sinica **4**, 505–524.
- [4] Fleming T.R., Harrington D.P. (2005). *Counting processes and survival analysis*. Wiley Series in Probability and Science.
- [5] Kalbfleisch J.D., Prentice R.L. (2002). *The statistical analysis of failure time data*. Wiley Series in Probability and Science.
- [6] Kraus D. (2004). *Testing goodness of fit of hazard regression models*. WDS 2004 Proceedings of contributed papers, Part I: Mathematics and Computer Sciences, 6–12.
- [7] Marzec L., Marzec P. (1997). *Generalized martingale-residual processes for goodness-of-fit inference in Cox's type regression models*. The Annals of Statistics **25**, 2, 683–714.
- [8] Sinha D., Dey D.K. (1997). *Semiparametric Bayesian analysis of survival data*. Journal of the American Statistical Association **92**, 439, 1195–1212.
- [9] Ibrahim J.G., Chen M., Sinha D. (2001). *Bayesian survival analysis*. Springer-Verlag.
- [10] Timková J. (2008). *Bayesian nonparametric estimation of hazard rate in survival analysis using Gibbs sampler*. WDS 2008 Proceedings of contributed papers, Part I: Mathematics and Computer Sciences, 80–87.
- [11] Volf P. (1996). *Analysis of generalized residuals in hazard regression models*. Kybernetika **32**, 5, 501–510.

Poděkování: The present work was supported by the grants GA ČR 201/05/H007 and GA AV IAA101120604.

Adresa: J. Timková, MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8 – Karlín
E-mail: timkova@karlin.mff.cuni.cz

A NEW R PACKAGE FOR BAYESIAN ESTIMATION OF MULTIVARIATE NORMAL MIXTURES ALLOWING FOR SELECTION OF THE NUMBER OF COMPONENTS AND INTERVAL-CENSORED DATA

Arnošt Komárek

Keywords: Density estimation, deviance information criterion, Markov chain Monte Carlo, penalized expected deviance, reversible jump, software.

Abstract: An R package `mixAK` is introduced which implements routines for a semiparametric density estimation through normal mixtures using the Markov chain Monte Carlo (MCMC) methodology. Besides producing the MCMC output, the package computes posterior summary statistics for important characteristics of the fitted distribution or computes and visualizes the posterior predictive density. For the estimated models, penalized expected deviance (PED) and deviance information criterion (DIC) is directly computed which allows for a selection of mixture components. Additionally, multivariate right-, left- and interval-censored observations are allowed. For univariate problems, the reversible jump MCMC algorithm has been implemented and can be used for a joint estimation of the mixture parameters and the number of mixture components. The core MCMC routines have been implemented in C++ and linked to R to ensure a reasonable computational speed. In the paper, implemented algorithms are briefly reviewed and the use of the package illustrated on three real examples of different complexity.

1 Introduction

It has been proven on several places that normal mixtures are a suitable semi-parametric structure to model unknown distributions and at the same time are a natural tool for clustering or modelling heterogeneity, see [30], [17], [18] or [1] for a comprehensive discussion of the topic. Starting with a paper by [6], in which the Gibbs algorithm has been introduced for normal mixtures with pre-specified number of mixture components, their use for Bayesian density estimation became relatively routine. Subsequently, [23] showed how the reversible jump Markov chain Monte Carlo (RJ-MCMC) algorithm of [9] can be used for a joint estimation of mixture parameters and a number of mixture components in a univariate case. Since then, several attempts have been made to develop a RJ-MCMC algorithm for multivariate normal mixtures, see, e.g., [5].

Alternatively, selection of the number of mixture components can be based on comparison of models fitted with different numbers of components by the means of some joint measure of model complexity and fit. A popular such measure, especially in Bayesian modelling, is the deviance information criterion (DIC, [26]). In rejoinder to discussion, [3] however conclude that it remains controversial to apply DIC beyond the exponential family case due to lack in theoretical foundation. Possible solution to this problem is offered by [20] who suggests to use penalized loss function for Bayesian model comparison and shows that DIC is an approximation to a penalized loss function based on the deviance, with a penalty derived from a cross-validation argument. Particularly in mixture context, [20] recommends to use penalized expected deviance (PED). Another possibility for Bayesian model comparison is provided by the Bayes factor [13] which however has some practical limitations (see, e.g., [20]).

At this point, mixtures with a pre-specified number of components can be easily handled by WinBUGS [16] however DIC or PED can be only manually and with some programming effort computed for mixtures in the current version. Additionally, reversible jump, birth death or other transdimensional MCMC algorithms for mixture problems do not seem to be fully implemented in any of nowadays standard publicly available packages like R [22] or WinBUGS. See also conclusion section of this paper, where we discuss some WinBUGS suggestions due to [15]. The purpose of this paper is to present a new R package called `mixAK` which allows for the initial analysis of mixtures either with pre-specified number of components where selection of the number of components is based on DIC or PED, or with the number of components estimated jointly with the remaining model parameters using the RJ-MCMC algorithm. To make the applicability of the package even broader, the package allows to estimate the density if the (part of) data are right-, left-, or generally interval-censored. To ensure a reasonable computational speed, the core parts of the MCMC routines have been implemented in C++ and linked to R.

The rest of the paper is organized as follows. In Section 2, the normal mixture as a model for the unknown density is described together with Bayesian specification of the problem. Section 3 provides an overview of the MCMC methods used in the package `mixAK` to sample from the corresponding posterior distribution. Section 4 describes possibility how to initialize two chains to start MCMC. The posterior inference is discussed in Section 5. A practical analysis using the package is illustrated on three real data examples in Section 6. The paper is finalized by concluding remarks in Section 7 where we also discuss limitations of the package.

2 Model

2.1 Basic model for a density

Let $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ be p -dimensional i.i.d. random vectors with a density $g_y(\mathbf{y})$. To incorporate right-, left- and interval-censoring, we will assume that we observe $[l_1, \mathbf{u}_1], \dots, [l_n, \mathbf{u}_n]$, where

$$[l_i, \mathbf{u}_i] = \begin{pmatrix} [l_{i,1}, u_{i,1}] \\ \vdots \\ [l_{i,p}, u_{i,p}] \end{pmatrix} \quad (i = 1, \dots, n). \quad (1)$$

In the expression (1), $-\infty \leq l_{i,j} < \infty$ are lower limits of observed intervals, $-\infty < u_{i,j} \leq \infty$ are upper limits of observed intervals ($l_{i,j} \leq u_{i,j}$, $i = 1, \dots, n$, $j = 1, \dots, p$) and $[]$ is open-, closed-, or half-open interval according to the context. Note that $y_{i,j} = l_{i,j} = u_{i,j}$ if the observation is not censored, $-\infty = l_{i,j} < u_{i,j} < \infty$ indicates left-censored observation, $-\infty < l_{i,j} < u_{i,j} = \infty$ indicates right-censored observation and $-\infty < l_{i,j} < u_{i,j} < \infty$ indicates interval-censored observation. Censoring leading to observed intervals is assumed to be non-informative. Density $g_y(\mathbf{y})$ is modelled as the following shifted and scaled normal mixture:

$$\begin{aligned} g_y(\mathbf{y}) &= |\mathbf{S}|^{-1} \sum_{k=1}^K w_k \varphi(\mathbf{S}^{-1}(\mathbf{y} - \mathbf{m}) \mid \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \\ &= \sum_{k=1}^K w_k \varphi(\mathbf{y} \mid \mathbf{m} + \mathbf{S}\boldsymbol{\mu}_k, \mathbf{S}\boldsymbol{\Sigma}_k\mathbf{S}'), \end{aligned} \quad (2)$$

where $\varphi(\cdot \mid \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ is a density of the normal distribution $\mathcal{N}_p(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$, $\mathbf{w} = (w_1, \dots, w_K)'$, $0 \leq w_k \leq 1$, $\sum_{k=1}^K w_k = 1$, is a vector of unknown mixture weights, $\boldsymbol{\mu} = \{\boldsymbol{\mu}_k : k = 1, \dots, K\}$ are unknown mixture means, and $\boldsymbol{\Sigma} = \{\boldsymbol{\Sigma}_k : k = 1, \dots, K\}$ are unknown mixture variance-covariance matrices. Further, $\mathbf{m} = (m_1, \dots, m_p)'$ is a fixed shift vector and $\mathbf{S} = \text{diag}(s_1, \dots, s_p)$ is a fixed scale matrix. Inclusion of the shift vector \mathbf{m} and the scale matrix \mathbf{S} in the model is here mainly due to a possibility to improve the mixing and numerical stability of the MCMC algorithm described in section 3, especially needed when different margins are measured in considerably different scales. Finally, let $\mathbf{Q} = \{\mathbf{Q}_1, \dots, \mathbf{Q}_K\}$ be mixture precision matrices, i.e., $\mathbf{Q}_k = \boldsymbol{\Sigma}_k^{-1}$ ($k = 1, \dots, K$). In the remainder of the paper, let $\mathbf{y}^* = \mathbf{S}^{-1}(\mathbf{y} - \mathbf{m})$ be shifted and scaled values of \mathbf{y} and let

$$g_{y^*}(\mathbf{y}^*) = \sum_{k=1}^K w_k \varphi(\mathbf{y}^* \mid \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \quad (3)$$

be a density of $\mathbf{Y}^* = \mathbf{S}^{-1}(\mathbf{Y} - \mathbf{m})$. Finally, let $\mathbf{l}_i^* = \mathbf{S}^{-1}(\mathbf{l}_i - \mathbf{m})$ and $\mathbf{u}_i^* = \mathbf{S}^{-1}(\mathbf{u}_i - \mathbf{m})$ ($i = 1, \dots, n$) be shifted and scaled lower and upper limits of observed intervals.

2.2 Bayesian specification

Bayesian technique and MCMC are used to estimate unknown parameters and make a corresponding inference in the package `mixAK`. Hence a prior distribution has to be specified for model parameters. Let $f(\cdot)$ and $f(\cdot|\cdot)$, respectively, be a generic symbol for a (conditional) probability density and let

$$\boldsymbol{\theta} = (w_1, \dots, w_K, \boldsymbol{\mu}'_1, \dots, \boldsymbol{\mu}'_K, \text{vec}(\mathbf{Q}_1)', \dots, \text{vec}(\mathbf{Q}_K)', \boldsymbol{\gamma}')' \quad (4)$$

be a vector of unknown parameters provided the number of mixture components is known or fixed in advance. In expression (4), $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_p)'$ is the variance hyperparameter and its meaning will be explained below. Further, let [data] denote a set $\{\mathbf{l}_i, \mathbf{u}_i : i = 1, \dots, n\}$ of observed intervals. Following expression (2), the *observed data* likelihood of the model equals

$$\begin{aligned} L_{\boldsymbol{\theta}}(\boldsymbol{\theta}, K) &= f([\text{data}] | \boldsymbol{\theta}, K) \\ &= |\mathbf{S}|^{-n} \prod_{i=1}^n \left\{ \int_{\mathbf{l}_{i,1}}^{\mathbf{u}_{i,1}} \cdots \int_{\mathbf{l}_{i,p}}^{\mathbf{u}_{i,p}} \sum_{k=1}^K w_k \varphi(\mathbf{S}^{-1}(\mathbf{y}_i - \mathbf{m}) | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) dy_{i,p} \cdots dy_{i,1} \right\}, \end{aligned} \quad (5)$$

with the convention that $\int_l^u f(y)dy = \int_l^u f(y)dy$, whenever $l < u$ and $\int_l^u f(y)dy = f(l) = f(u)$ whenever $l = u$ (uncensored observation).

The following prior specifications are implemented in the package `mixAK` and specified default values are used if not determined by the user. Note that default values attempt to use weakly informative prior distribution.

Mixture weights \mathbf{w} : a Dirichlet distribution $D(\delta, \dots, \delta)$, i.e.,

$$f(\mathbf{w} | K) = \{\Gamma(\delta)\}^{-K} \Gamma(K\delta) \prod_{k=1}^K w_k^{\delta-1}, \quad (6)$$

where δ is a fixed hyperparameter. A default value is $\delta = 1$.

Mixture means $\boldsymbol{\mu}$ and precision matrices \mathbf{Q} : it is possible to choose

1. *Semiconjugate independent* Normal and Wishart prior independently for the K components, i.e.,

$$\begin{aligned}
 f(\boldsymbol{\mu}, \mathbf{Q} | \boldsymbol{\gamma}, K) &= \prod_{k=1}^K \{f(\boldsymbol{\mu}_k) f(\mathbf{Q}_k | \boldsymbol{\gamma})\} \\
 &\propto \prod_{k=1}^K \left[|\mathbf{D}_k|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\boldsymbol{\mu}_k - \boldsymbol{\xi}_k)' \mathbf{D}_k^{-1} (\boldsymbol{\mu}_k - \boldsymbol{\xi}_k) \right\} \right. \\
 &\quad \left. \times |\boldsymbol{\Xi}|^{-\frac{\zeta}{2}} |\mathbf{Q}_k|^{\frac{\zeta-p-1}{2}} \exp \left\{ -\frac{1}{2} \text{tr}(\boldsymbol{\Xi}^{-1} \mathbf{Q}_k) \right\} \right].
 \end{aligned} \tag{7}$$

That is, $f(\boldsymbol{\mu}_k)$ is a density of the normal $\mathcal{N}_p(\boldsymbol{\xi}_k, \mathbf{D}_k)$ and $f(\mathbf{Q}_k | \boldsymbol{\gamma})$ is a density of the Wishart $W_p(\zeta, \boldsymbol{\Xi})$, where $\boldsymbol{\Xi} = \text{diag}(\gamma_1, \dots, \gamma_p)$. The semiconjugate prior is also the default choice in the package `mixAK`.

2. *Natural-conjugate* Normal-Wishart prior independently for the K components, i.e.,

$$\begin{aligned}
 f(\boldsymbol{\mu}, \mathbf{Q} | \boldsymbol{\gamma}, K) &= \prod_{k=1}^K \{f(\boldsymbol{\mu}_k | \mathbf{Q}_k) f(\mathbf{Q}_k | \boldsymbol{\gamma})\} \\
 &\propto \prod_{k=1}^K \left[c_k^{\frac{p}{2}} |\mathbf{Q}_k|^{\frac{1}{2}} \exp \left\{ -\frac{c_k}{2} (\boldsymbol{\mu}_k - \boldsymbol{\xi}_k)' \mathbf{Q}_k (\boldsymbol{\mu}_k - \boldsymbol{\xi}_k) \right\} \right. \\
 &\quad \left. \times |\boldsymbol{\Xi}|^{-\frac{\zeta}{2}} |\mathbf{Q}_k|^{\frac{\zeta-p-1}{2}} \exp \left\{ -\frac{1}{2} \text{tr}(\boldsymbol{\Xi}^{-1} \mathbf{Q}_k) \right\} \right].
 \end{aligned} \tag{8}$$

That is, $f(\boldsymbol{\mu}_k | \mathbf{Q}_k)$ is a density of the normal $\mathcal{N}_p(\boldsymbol{\xi}_k, c_k^{-1} \mathbf{Q}_k^{-1})$ and $f(\mathbf{Q}_k | \boldsymbol{\gamma})$ is a density of the Wishart $W_p(\zeta, \boldsymbol{\Xi})$ with $\boldsymbol{\Xi} = \text{diag}(\gamma_1, \dots, \gamma_p)$.

In prior distributions (7) and (8), $\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_K$ are prior means, $\zeta > p - 1$ are Wishart prior degrees of freedom and $\boldsymbol{\Xi}$ is Wishart prior scale matrix. Further, $\mathbf{D}_1, \dots, \mathbf{D}_K$ in (7) are prior variance-covariance matrices and $c_1 > 0, \dots, c_K > 0$ in (8) are prior precision parameters. Let $y_{\min,j}^* = \min\{y_{i,j}^* : i = 1, \dots, n\}$, $y_{\max,j}^* = \max\{y_{i,j}^* : i = 1, \dots, n\}$, $R_j = y_{\max,j}^* - y_{\min,j}^*$ ($j = 1, \dots, p$) be minimum, maximum and range of the shifted- and scaled-data in margin j where in the case of censoring unobserved values of y^* are replaced by their reasonable initial values (e.g., midpoints of observed intervals). Possible choices for prior hyperparameters are then $\xi_{k,j} = 0.5(y_{\min,j}^* + y_{\max,j}^*)$, $\mathbf{D}_k = \text{diag}(d_{k,1}, \dots, d_{k,p})$, $d_{k,j} = R_j^2$ ($k = 1, \dots, K$, $j = 1, \dots, p$) which

follows suggestions of [23]. Weakly informative Wishart prior is obtained, e.g., with $\zeta = p + 1$. Default values for precisions c_1, \dots, c_K with a natural-conjugate prior are $c_k = 1$ ($k = 1, \dots, K$).

Variance hyperparameter γ As noted by [23] in the univariate case ($p = 1$), it seems restrictive to suppose that knowledge of the range of the data (R_j) implies much about the size of \mathbf{Q}_k and suggest to add an additional hierarchical level and use a gamma hyperprior for Ξ . Multivariately, we will assume that $\Xi = \text{diag}(\gamma_1, \dots, \gamma_p)$ and a priori $\gamma_j^{-1} \sim \text{Gamma}(g_j, h_j)$ independently for $j = 1, \dots, p$ which is the prior construction used also by [5]. That is,

$$f(\gamma_1^{-1}, \dots, \gamma_p^{-1}) = \prod_{j=1}^p \left\{ \frac{h_j^{g_j}}{\Gamma(g_j)} (\gamma_j^{-1})^{g_j-1} \exp(-h_j \gamma_j^{-1}) \right\}. \quad (9)$$

Following [23], a weak prior is obtained with g_j being a small positive value and h_j being a small multiple of $1/R_j^2$. Our default values are $g_j = 0.2$, $h_j = 10/R_j^2$ ($j = 1, \dots, K$).

Suppose, we want to estimate the number of mixture components K jointly with the remaining mixture parameters. A possible approach is to follow [23] where a prior distribution is assumed for K and the inference is based on a sample from the joint posterior distribution of $\boldsymbol{\theta}$ and K obtained using the RJ-MCMC of [9]. In the package `mixAK`, two prior distributions have been implemented for K .

1. *Uniform* prior on $\{1, \dots, K_{\max}\}$, i.e.,

$$f(K) \equiv \text{P}(K = K^*) = \frac{1}{K_{\max}} \quad (K^* = 1, \dots, K_{\max}), \quad (10)$$

where K_{\max} is a chosen maximal number of mixture components.

2. *Truncated Poisson* prior, i.e.,

$$f(K) \equiv \text{P}(K = K^*) \propto \frac{\lambda^{K^*} \exp(-\lambda)}{K^*!} \quad (K = 1, \dots, K_{\max}), \quad (11)$$

where K_{\max} is a chosen maximal number of mixture components and λ chosen untruncated Poisson mean.

Default approach in package `mixAK` is with K fixed in advance (i.e., $\text{P}(K = K_{\max}) = 1$) in which case the number of mixture components must be chosen according to the value of PED or DIC. Note that at this moment, joint estimation of K and $\boldsymbol{\theta}$ using RJ-MCMC is implemented for univariate data ($p = 1$) only.

The prior distribution of our model is then hierarchically specified as

$$f(\boldsymbol{\theta}, K) = f(\boldsymbol{\theta} | K) \times f(K), \quad (12)$$

where $f(\boldsymbol{\theta} | K) = f(\mathbf{w} | K) \times f(\boldsymbol{\mu}, \mathbf{Q} | \boldsymbol{\gamma}, K) \times f(\boldsymbol{\gamma})$ follows from (6)–(9) and $f(K)$ follows from (10) or (11). Using the Bayes’ formula, the posterior distribution combines (5) and (12) into

$$f(\boldsymbol{\theta}, K | [\text{data}]) \propto L_{\boldsymbol{\theta}}(\boldsymbol{\theta}, K) \times f(\boldsymbol{\theta}, K). \tag{13}$$

2.3 Latent parameters

Computation of the posterior distribution is simplified by introduction of latent (additional) parameters as generally explained by [29]. Let $\boldsymbol{\psi}$ be the vector of latent parameters. A joint prior is specified for $(\boldsymbol{\psi}', \boldsymbol{\theta}', K)'$ hierarchically as

$$f(\boldsymbol{\psi}, \boldsymbol{\theta}, K) = f(\boldsymbol{\psi} | \boldsymbol{\theta}, K) \times f(\boldsymbol{\theta} | K) \times f(K), \tag{14}$$

the sample from the posterior distribution $f(\boldsymbol{\psi}, \boldsymbol{\theta}, K | [\text{data}])$ is obtained using the (RJ-)MCMC methodology and the inference is based on a sample from the (marginal) posterior distribution $f(\boldsymbol{\theta}, K | [\text{data}])$ which is directly available as a subset of the complete sample. In this context, the first term in the decomposition (14), $f(\boldsymbol{\psi} | \boldsymbol{\theta}, K)$, is sometimes called as the *complete data* likelihood.

In mixture context with censored data, one often considers two sorts of latent parameters: (i) component allocations denoted by $\mathbf{r} = (r_1, \dots, r_n)'$, $r_i \in \{1, \dots, K\}$, (ii) unobserved values of (shifted and scaled) censored data, i.e., values of $\mathbf{y}_1^*, \dots, \mathbf{y}_n^*$ for those observations which are censored. For convenience in notation we will provide explanation for the situation when all observations are censored and hence $\boldsymbol{\psi} = (\mathbf{y}_1^{*'}, \dots, \mathbf{y}_n^{*'}, r_1, \dots, r_n)'$. It is easily seen that the complete data likelihood is decomposed as

$$f(\boldsymbol{\psi} | \boldsymbol{\theta}, K) = \prod_{i=1}^n f(\mathbf{y}_i^*, r_i | \boldsymbol{\theta}, K) = \prod_{i=1}^n \{f(\mathbf{y}_i^* | r_i, \boldsymbol{\theta}, K) \times f(r_i | \boldsymbol{\theta}, K)\}, \tag{15}$$

where

$$\begin{aligned} f(\mathbf{y}_i^* | r_i, \boldsymbol{\theta}, K) &= f(\mathbf{y}_i^* | r_i, \boldsymbol{\theta}) = \varphi(\mathbf{y}_i^* | \boldsymbol{\mu}_{r_i}, \boldsymbol{\Sigma}_{r_i}) \\ f(r_i | \boldsymbol{\theta}, K) &\equiv P(r_i = k | \boldsymbol{\theta}, K) = w_k \quad (k = 1, \dots, K). \end{aligned}$$

Together with the likelihood

$$\begin{aligned} L_{\boldsymbol{\psi}, \boldsymbol{\theta}}(\boldsymbol{\psi}, \boldsymbol{\theta}, K) &= f([\text{data}] | \boldsymbol{\psi}, \boldsymbol{\theta}, K) \\ &= \prod_{i=1}^n f(\mathbf{l}_i^*, \mathbf{u}_i^* | \mathbf{y}_i^*) \propto \prod_{i=1}^n I(y_{i,1}^* \in [l_{i,1}^*, u_{i,1}^*], \dots, y_{i,p}^* \in [l_{i,p}^*, u_{i,p}^*]), \end{aligned}$$

where I denotes an indicator function, this leads to the joint posterior distribution

$$f(\boldsymbol{\psi}, \boldsymbol{\theta}, K | [\text{data}]) \propto L_{\boldsymbol{\psi}, \boldsymbol{\theta}}(\boldsymbol{\psi}, \boldsymbol{\theta}, K) \times f(\boldsymbol{\psi}, \boldsymbol{\theta}, K) \tag{16}$$

for which the marginalization over $\boldsymbol{\psi}$ leads to the desirable (marginal) posterior (13).

3 Markov chain Monte Carlo

The sample $\{\boldsymbol{\psi}^{(t)}, \boldsymbol{\theta}^{(t)}, K^{(t)} : t = 1, \dots, T\}$ from the posterior distribution is obtained using a (reversible jump) Markov chain Monte Carlo simulation in which one iteration consists of the following move types depending on whether the number of mixture components K is prespecified or not. With K prespecified, the algorithm iterates between

1. updating the latent (censored) observations, see section 3.1;
2. updating the mixture related parameters, see section 3.2.

With K random (allowed currently only when $p = 1$), we iterate between

1. updating the latent (censored) observations, see section 3.1;
2. update of mixture related parameters as with fixed K as described in section 3.2;
3. split-combine move, see section 3.3;
4. birth-death move, see section 3.3.

Improvement in the mixing of the chains can be achieved if within one MCMC sweep, only one of the steps (2)–(4) is performed, each with a given probability π_a^{mix} , π_b^{mix} or π_c^{mix} , respectively ($\pi_a^{mix} + \pi_b^{mix} + \pi_c^{mix} = 1$). This is also the user's option in the package `mixAK`.

3.1 Update of latent (censored) observations

For each i , when \mathbf{y}_i^* contains some censored components, it is updated by sampling from the full conditional distribution which is a (multivariate) normal distribution $\mathcal{N}_p(\boldsymbol{\mu}_{r_i}, \boldsymbol{\Sigma}_{r_i})$ constrained by the limits of the observed intervals \mathbf{l}_i^* and \mathbf{u}_i^* . In a univariate case, this is done by the inverse cdf sampling, in a multivariate case a method described by [8] is used.

3.2 Update of mixture related parameters without a change of K

The moves where the mixture parameters are updated, however when the number of mixture components K remains unaltered, follow largely the proposal of [6]. In fact, all parameters are updated in blocks using a Gibbs kernel by sampling from the full conditional distribution. In the following, let $f(\cdot | \dots)$ denote a density of the full conditional distribution. In summary we perform the following sub-moves:

Update of component allocations: sampling from

$$f(r_i | \dots) \equiv \text{P}(r_i = k | \dots) \propto w_k \varphi(\mathbf{y}_i^* | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \quad (k = 1, \dots, K)$$

sequentially for $i = 1, \dots, n$.

Update of mixture weights: sampling from the Dirichlet distribution $D(\delta + N_1, \dots, \delta + N_k)$, where

$$N_k = \sum_{i=1}^n I[r_i = k] \quad (k = 1, \dots, K).$$

Update of mixture precision matrices: sampling from the Wishart distribution $W_p(\tilde{\zeta}_k, \tilde{\Xi}_k)$ sequentially for $k = 1, \dots, K$ to update \mathbf{Q}_k . For both priors (7) and (8), degrees of freedom of the k -th full conditional Wishart distribution are equal $\tilde{\zeta}_k = \zeta + N_k$. In the case of semiconjugate independent prior (7), the scale matrix of the k -th full conditional Wishart distribution is

$$\tilde{\Xi}_k = \left\{ \Xi^{-1} + \sum_{i:r_i=k} (\mathbf{y}_i^* - \boldsymbol{\mu}_k)(\mathbf{y}_i^* - \boldsymbol{\mu}_k)' \right\}^{-1}.$$

In the case of natural-conjugate prior (8), the scale matrix of the k -th full conditional Wishart distribution is

$$\tilde{\Xi}_k = \left\{ \Xi^{-1} + \sum_{i:r_i=k} (\mathbf{y}_i^* - \bar{\mathbf{y}}_k^*)(\mathbf{y}_i^* - \bar{\mathbf{y}}_k^*)' + \frac{N_k c_k}{N_k + c_k} (\bar{\mathbf{y}}_k^* - \boldsymbol{\xi}_k)(\bar{\mathbf{y}}_k^* - \boldsymbol{\xi}_k)' \right\}^{-1},$$

where $\bar{\mathbf{y}}_k^* = N_k^{-1} \sum_{i:r_i=k} \mathbf{y}_i^*$.

Update of mixture means: sampling from the normal distribution $\mathcal{N}_p(\mathbf{m}_k, \mathbf{S}_k)$ sequentially for $k = 1, \dots, K$ to update $\boldsymbol{\mu}_k$. For semiconjugate independent prior (7), parameters of the k -th full conditional normal distribution are

$$\mathbf{S}_k = (N_k \mathbf{Q}_k + \mathbf{D}_k^{-1})^{-1}, \quad \mathbf{m}_k = \mathbf{S}_k (N_k \mathbf{Q}_k \bar{\mathbf{y}}_k^* + \mathbf{D}_k^{-1} \boldsymbol{\xi}_k).$$

For natural-conjugate prior (8), parameters of the k -th full conditional normal distribution are

$$\mathbf{S}_k = \{(N_k + c_k) \mathbf{Q}_k\}^{-1}, \quad \mathbf{m}_k = (N_k + c_k)^{-1} (N_k \bar{\mathbf{y}}_k^* + c_k \boldsymbol{\xi}_k).$$

Update of variance hyperparameter: sampling from gamma distributions $\text{Gamma}(\tilde{g}_j, \tilde{h}_j)$ sequentially for $j = 1, \dots, p$ to update γ_j . Parameters of the j -th full conditional gamma distribution are

$$\tilde{g}_j = g_j + \frac{K\zeta}{2}, \quad \tilde{h}_j = h_j + \frac{1}{2} \sum_{k=1}^K \mathbf{Q}_k(j, j),$$

where $\mathbf{Q}_k(j, j)$ denotes the j -th diagonal element of matrix \mathbf{Q}_k .

3.3 Moves allowing a change of the number of mixture components

For univariate data ($p = 1$), moves allowing a change of the number of mixture components follow the RJ-MCMC approach taken in [23]. Let μ_1, \dots, μ_K be the (univariate) mixture means and $\sigma_1^2, \dots, \sigma_K^2$ be the (univariate) mixture variances. Two move types are proposed.

Split–combine move consists of either splitting an existing component into two new components or combining two existing components into a new one. Firstly a random choice is made whether to perform the split or combine move. Namely, given K , the probability of attempting the split move is π_K^{split} and the probability of attempting the combine move is $\pi_K^{combine} = 1 - \pi_K^{split}$. Default values in the package `mixAK` are $\pi_1^{split} = 1$, $\pi_2^{split} = \dots = \pi_{K_{\max}-1}^{split} = 0.5$ and $\pi_{K_{\max}}^{split} = 0$ and can possibly be modified by the user.

When a *combine* move is attempted, a pair of neighboring components is chosen, let say k_1 and k_2 such that $\mu_{k_1} < \mu_{k_2}$ and there is no other mixture mean in the interval $[\mu_{k_1}, \mu_{k_2}]$. The two components are combined such that a new component, let say k^* , has the same 0th, 1st and 2nd moment as the original two components. That is, the weight w_{k^*} , the mean μ_{k^*} and the variance $\sigma_{k^*}^2$ of the new component are given by

$$w_{k^*} = w_{k_1} + w_{k_2}, \quad \mu_{k^*} = \frac{w_{k_1}\mu_{k_1} + w_{k_2}\mu_{k_2}}{w_{k^*}},$$

$$\sigma_{k^*}^2 = \frac{w_{k_1}(\mu_{k_1}^2 + \sigma_{k_1}^2) + w_{k_2}(\mu_{k_2}^2 + \sigma_{k_2}^2)}{w_{k^*}} - \mu_{k^*}^2.$$

The move into a new state with components k_1 and k_2 replaced by the component k^* and the number of mixture components K decreased by 1 is then either accepted or rejected with a certain Metropolis-Hastings probability, see [23] for more details. In the case of acceptance, (latent) observations allocated originally into the components k_1 and k_2 are re-allocated in the new component k^* .

In the *split* move, a component, let say k^* is chosen at random from existing K components and two new components, let say k_1 and k_2 are proposed such that the split move is reversible to the combine move in a sense described generally in [9]. The weights w_{k_1} , w_{k_2} , the means μ_{k_1} , μ_{k_2} , and the variances $\sigma_{k_1}^2$, $\sigma_{k_2}^2$ of the two new components are given by

$$w_{k_1} = w_{k^*}u_1, \quad w_{k_2} = w_{k^*}(1 - u_1),$$

$$\mu_{k_1} = \mu_{k^*} - u_2\sigma_{k^*}\sqrt{\frac{w_{k_2}}{w_{k_1}}}, \quad \mu_{k_2} = \mu_{k^*} + u_2\sigma_{k^*}\sqrt{\frac{w_{k_1}}{w_{k_2}}},$$

$$\sigma_{k_1}^2 = u_3(1 - u_2^2)\sigma_{k^*}^2\frac{w_{k^*}}{w_{k_1}}, \quad \sigma_{k_2}^2 = (1 - u_3)(1 - u_2^2)\sigma_{k^*}^2\frac{w_{k^*}}{w_{k_2}},$$

where $\mathbf{u} = (u_1, u_2, u_3)'$ is an auxiliary random vector whose components are generated randomly independently, u_l from $\text{Beta}(a_l, b_l)$ ($l = 1, 2, 3$). Default values in package `mixAK` are $a_1 = b_1 = 2$, $a_2 = 1$, $b_2 = 2$, $a_3 = b_3 = 1$. Additionally, for (latent) observations allocated in the component k^* a new allocation, either k_1 or k_2 is randomly proposed. The whole proposal is accepted with a certain Metropolis-Hastings probability as described in detail by [23]. Note that the proposal is rejected with probability one if there is other mixture mean in the interval $[\mu_{k_1}, \mu_{k_2}]$ to ensure reversibility with the combine move.

Birth–death move consists of either proposing a new component (birth) or deleting one of the empty components (death). Similarly to the split–combine move, a random choice is made whether to perform the birth or the death move. Namely, given K , the probability of attempting the birth move is π_K^{birth} and the probability of attempting the death move is $\pi_K^{death} = 1 - \pi_K^{birth}$. Default values in the package `mixAK` are $\pi_1^{birth} = 1$, $\pi_2^{birth} = \dots = \pi_{K_{\max}-1}^{birth} = 0.5$ and $\pi_{K_{\max}}^{birth} = 0$ and can possibly be modified by the user.

When a *birth* move is attempted a new component weight w_{k^*} , mean μ_{k^*} and variance $\sigma_{k^*}^2$ are sampled from the prior distribution, the weights of existing components are rescaled to satisfy the sum-up-to-one constraint and the whole proposal is accepted with a certain Metropolis-Hastings probability computed along the lines described in [23].

When a *death* move is attempted one of empty components, i.e. components into which currently there are no observations allocated and hence $N_k = 0$, is chosen at random. The proposal consists of deleting this component and rescaling the remaining weights to satisfy the sum-up-to-one constraint. Analogously to the previous cases, the whole proposal is accepted with a certain Metropolis-Hastings probability, see [23] for details.

4 Initial values to start MCMC

To start the MCMC simulation, initial values are required for the model parameters. In the package `mixAK`, all initial values can either be supplied by the user or generated automatically by the program. Up to two chains can automatically be initialized using two different strategies. Initial values for latent (censored) observations from chain 1 are also used to determine data driven priors described in Sec. 2.2. In a sequel, let \hat{s}_j^* be sample standard deviations of uncensored shifted and scaled observations, lower bounds of shifted and scaled right- and interval-censored observations and upper bounds of shifted and scaled left-censored observations in the j -th margin ($j = 1, \dots, p$).

Latent (censored) observations For uncensored observation, both chains start indeed from $y_{i,j}^* = l_{i,j}^* = u_{i,j}^*$. For right-censored observation,

chain 1 starts from $y_{i,j}^* = l_{i,j}^* + \hat{s}_j^*$ and chain 2 from $y_{i,j}^* = l_{i,j}^* + |z_{i,j}^*|$, where $z_{i,j}^*$ is sampled from $\mathcal{N}(0, (\hat{s}_j^*)^2)$. Similarly, for left-censored observation, chain 1 starts from $y_{i,j}^* = u_{i,j}^* - \hat{s}_j^*$ and chain 2 from $y_{i,j}^* = u_{i,j}^* - |z_{i,j}^*|$, where $z_{i,j}^*$ is sampled from $\mathcal{N}(0, (\hat{s}_j^*)^2)$. For interval-censored observation, chain 1 starts from the midpoint $y_{i,j}^* = 0.5(l_{i,j}^* + u_{i,j}^*)$ and chain 2 from a value chosen uniformly at random from interval $(l_{i,j}^*, u_{i,j}^*)$ ($i = 1, \dots, n, j = 1, \dots, p$).

Number of mixture components When the number of mixture components is estimated using the RJ-MCMC, the first chain is initialized with $K = 1$ and the second chain with $K = \min(2, K_{\max})$.

Mixture weights The first chain starts with $w_1 = \dots = w_K = K^{-1}$, where K is the initial value for the number of mixture components. The initial weights for the second chain are sampled from the prior Dirichlet distribution $D(\delta, \dots, \delta)$.

Mixture means For chain 1, initial mixture means in the j -th margin are chosen equidistantly in the interval $(y_{\min,j}^*, y_{\max,j}^*)$ ($j = 1, \dots, p$). For chain 2, initial means in the j -th margin are independently sampled from the normal distributions with means equidistantly splitted in the interval $(y_{\min,j}^*, y_{\max,j}^*)$ and standard deviations equal to \hat{s}_j^*/K , where K is the initial number of mixture components for the second chain.

Mixture precision matrices The initial mixture precision matrices for the first chain are all the same and are equal to inverted sample variance-covariance matrix computed from initial values of (latent) observations. For chain 2, the initial mixture precision matrices are all diagonal. Diagonal of the initial value of matrix \mathbf{Q}_k is equal to inverted sample variances of margins of (latent) observations multiplied by z_k^{-1} , where z_k is sampled from the uniform distribution $\mathcal{U}(0.1, 1.1)$.

Component allocations For both chains, observation is allocated to component showing the highest posterior probability given the initial mixture and initial value of latent (censored) observation.

Variance hyperparameter The initial value of γ for the first chain is equal to the diagonal of the sample variance-covariance matrix computed from initial values of (latent) observations multiplied by the prior hyperparameter ζ . In the second chain, the same approach is exploited together with multiplication of each initial value of γ_j ($j = 1, \dots, p$) by a random variate sampled independently from the uniform distribution $\mathcal{U}(0, p)$.

5 Posterior inference

5.1 Label switching problem

It is well known that in mixture problems the posterior distributions (13) or (16) are invariant against the switching of the labelling of mixture components. For example, in [6], [23] artificial identifiability constraints of the type $\mu_1 < \dots < \mu_K$ are used in a univariate setting to ensure identifiability of the posterior distribution. The problem becomes rather complicated in higher dimensions since the number of identifiability constraints on the parameter space is very large. For a general discussion of this problem, see [11]. In the MCMC implemented in the package `mixAK`, identification of the posterior distribution can be achieved, e.g., by using relabelling techniques retrospectively, by post-processing the MCMC output [28]. However, several important quantities related to the posterior inference, e.g. those of the posterior predictive inference (section 5.2) are invariant to label switching and hence relabelling is redundant in the situations when solely the predictive inference is of interest.

5.2 Predictive density

Very often and especially in situations when the normal mixture is used as a convenient semiparametric structure to model the unknown distribution, the estimate of the density (2) or (3) is of primary interest. A suitable estimator is given by the posterior predictive densities $E[g_y(\mathbf{y}) \mid \text{[data]}]$ or $E[g_{y^*}(\mathbf{y}^*) \mid \text{[data]}]$ which can easily be approximated from the MCMC output as

$$\hat{g}_y(\mathbf{y}) = \frac{1}{T} \sum_{t=1}^T g_y(\mathbf{y} \mid \boldsymbol{\theta}^{(t)}, K^{(t)}) \approx E[g_y(\mathbf{y}) \mid \text{[data]}], \quad (17)$$

where

$$g_y(\mathbf{y} \mid \boldsymbol{\theta}^{(t)}, K^{(t)}) = |\mathbf{S}|^{-1} \sum_{k=1}^{K^{(t)}} w_k^{(t)} \varphi(\mathbf{S}^{-1}(\mathbf{y} - \mathbf{m}) \mid \boldsymbol{\mu}_k^{(t)}, \boldsymbol{\Sigma}_k^{(t)}).$$

The expression for $\hat{g}_{y^*}(\mathbf{y}^*) \approx E[g_{y^*}(\mathbf{y}^*) \mid \text{[data]}]$ is analogous. In the package `mixAK`, functions have been implemented to compute the values of all uni- and bivariate marginal densities derived from \hat{g}_y or \hat{g}_{y^*} in a prespecified grid of \mathbf{y} or \mathbf{y}^* values and visualize them on plots.

5.3 Convergence of the chains

Due to the label switching, we know in advance that the posterior distribution of the K -component model has $K!$ symmetric modes and converging MCMC should visit all of them. This knowledge may be exploited when checking convergence by exploring the chains for mixture weights, means and variances before any possible relabelling, see, e.g. [11]. Additionally, it is

possible to compare posterior distributions of component weights, means or variances which should be identical among components (see subsection 6.2 for illustration).

Another strategy is to base the convergence diagnostic on quantities which are invariant against label switching. For example, moments or quantiles of the mixtures (2) and (3) satisfy this condition. In the package `mixAK`, first two moments, i.e.,

$$\mathbf{E}(\mathbf{Y}^*) = \sum_{k=1}^K w_k \boldsymbol{\mu}_k, \quad \mathbf{E}(\mathbf{Y}) = \mathbf{m} + \mathbf{S}\mathbf{E}(\mathbf{Y}^*),$$

$$\text{var}(\mathbf{Y}^*) = \sum_{k=1}^K w_k \left[\boldsymbol{\Sigma}_k + \{\boldsymbol{\mu}_k - \mathbf{E}(\mathbf{Y}^*)\} \{\boldsymbol{\mu}_k - \mathbf{E}(\mathbf{Y}^*)\}' \right], \quad \text{var}(\mathbf{Y}) = \mathbf{S} \text{var}(\mathbf{Y}^*) \mathbf{S}'.$$

are computed and stored at each iteration of the MCMC and their traceplots and other tools can be used to check the convergence.

Additionally, deviance based quantities are stored at each iteration of the MCMC and can also be used to evaluate the convergence. Namely, the following quantities are computed and stored for $t = 1, \dots, T$:

$$D_{obs}^{(t)} = -2 \sum_{i=1}^n \log \left\{ g_y(\mathbf{y}_i^{(t)} \mid \boldsymbol{\theta}^{(t)}, K^{(t)}) \right\}, \quad (18)$$

$$\ell_{compl,0}^{(t)} = \sum_{i=1}^n \log \left\{ |\mathbf{S}|^{-1} \varphi(\mathbf{S}^{-1}(\mathbf{y}_i^{(t)} - \mathbf{m}) \mid \boldsymbol{\mu}_{r_i}^{(t)}, \boldsymbol{\Sigma}_{r_i}^{(t)}) \right\}, \quad (19)$$

$$\ell_{compl,1}^{(t)} = \sum_{i=1}^n \log(w_{r_i}^{(t)}), \quad (20)$$

$$D_{compl}^{(t)} = -2(\ell_{compl,0}^{(t)} + \ell_{compl,1}^{(t)}). \quad (21)$$

Note that if there are no censored observations, $D_{obs}^{(t)}$ in (18) is twice the observed data logarithmic likelihood (5). In the following, $D_{obs}^{(t)}$ will be called *observed deviance* even in the presence of censoring. Further, $\ell_{compl,0}^{(t)}$ in (19) and $\ell_{compl,1}^{(t)}$ in (20) form together the complete data logarithmic likelihood (15) and hence $D_{compl}^{(t)}$ in (21) is twice the complete data logarithmic likelihood. In the following $D_{compl}^{(t)}$ will be called *complete deviance*.

5.4 Deviance information criterion and penalized expected deviance

A general approach to comparison of complex models based on the samples from the posterior distribution has been suggested by [26] who introduced the deviance information criterion (DIC). In the discussion section of their

paper, Richardson showed how the DIC based on the predictive density could be used to discriminate between mixture models with different numbers of components. Other versions of DIC for mixture and in general missing data models have been discussed by [3]. In package `mixAK`, the Richardson's version of DIC, denoted as DIC_3 in [3] has been implemented and can be used to compare the mixture models with different numbers of components, especially in multivariate situations when the reversible jump MCMC allowing for a joint estimation of mixture parameters and the number of mixture components has not been implemented. That is, our DIC is computed as

$$\text{DIC} = \bar{D} + p_D, \quad p_D = \bar{D} - \tilde{D},$$

where \bar{D} is the approximation to the posterior mean of the deviance, where the posterior expectation is taken with respect to $\boldsymbol{\theta}$, K (if random) and \mathbf{y} if there is censoring present. Further, \tilde{D} is the deviance evaluated in the "estimate" to the model parameters and p_D is the effective dimension. The deviance D is based on the normal mixture (2) and a predictive density is taken as the "estimate" to the model parameters. Hence using the MCMC sample

$$\bar{D} = \frac{1}{T} \sum_{t=1}^T D_{obs}^{(t)}, \quad \tilde{D} = -2 \log \left[\frac{1}{T} \sum_{t=1}^T \prod_{i=1}^n \left\{ g_y(\mathbf{y}_i^{(t)} \mid \boldsymbol{\theta}^{(t)}, K^{(t)}) \right\} \right].$$

We admit that there is some controversy in applying the DIC in mixture context (see discussion in [3]) and hence another possibility for comparing models with different numbers of mixture components is offered by the package `mixAK` – penalized expected deviance (PED) derived from a cross-validation argument introduced by [20]. Basically, $2n$ separate MCMC runs, with a single observation deleted in each of two runs are required to compute the PED. This computationally demanding task can be avoided by the use of importance sampling where only two parallel chains, $\{\boldsymbol{\theta}^{(1,t)}, K^{(1,t)} : t = 1, \dots, T\}$, $\{\boldsymbol{\theta}^{(2,t)}, K^{(2,t)} : t = 1, \dots, T\}$ of genuine parameters of interest are needed. Consequently, PED is computed as

$$\text{PED} = \hat{D}_e + \hat{p}_{opt},$$

where

$$\hat{D}_e = \frac{1}{2T} \sum_{t=1}^T (D_{obs}^{(1,t)} + D_{obs}^{(2,t)}),$$

$$D_{obs}^{(c,t)} = -2 \sum_{i=1}^n \log \left\{ g_y(\mathbf{y}_i^{(c,t)} \mid \boldsymbol{\theta}^{(c,t)}, K^{(c,t)}) \right\} \quad (c = 1, 2).$$

In the case of censoring, $\mathbf{y}_i^{(c,t)}$ is sampled from the normal mixture given by $\boldsymbol{\theta}^{(c,t)}$, $K^{(c,t)}$ truncated on the observed intervals $[\mathbf{l}_i, \mathbf{u}_i]$. Further, in both

censored and uncensored cases, replicated observations $\mathbf{y}^{(rep1,t)}$ and $\mathbf{y}^{(rep2,t)}$ are sampled from the (untruncated) normal mixture given by $\boldsymbol{\theta}^{(1,t)}$, $K^{(1,t)}$ and $\boldsymbol{\theta}^{(2,t)}$, $K^{(2,t)}$, respectively and used to calculate the optimism \hat{p}_{opt} of \hat{D}_e as

$$\hat{p}_{opt} = \sum_{i=1}^n \hat{p}_{opt_i},$$

$$\hat{p}_{opt_i} = \left(\sum_{t=1}^T w_i^{(t)} \right)^{-1} \sum_{t=1}^T w_i^{(t)} \left[\log \left\{ \frac{g_y(\mathbf{y}^{(rep1,t)} | \boldsymbol{\theta}^{(1,t)}, K^{(1,t)})}{g_y(\mathbf{y}^{(rep1,t)} | \boldsymbol{\theta}^{(2,t)}, K^{(2,t)})} \right\} \right. \\ \left. + \log \left\{ \frac{g_y(\mathbf{y}^{(rep2,t)} | \boldsymbol{\theta}^{(2,t)}, K^{(2,t)})}{g_y(\mathbf{y}^{(rep2,t)} | \boldsymbol{\theta}^{(1,t)}, K^{(1,t)})} \right\} \right],$$

where

$$w_i^{(t)} = \left\{ g_y(\mathbf{y}^{(1,t)} | \boldsymbol{\theta}^{(1,t)}, K^{(1,t)}) \cdot g_y(\mathbf{y}^{(2,t)} | \boldsymbol{\theta}^{(2,t)}, K^{(2,t)}) \right\}^{-1} \\ (i = 1, \dots, n, t = 1, \dots, T)$$

are importance sampling weights.

6 Examples

All examples in this section have been run on Intel Core 2 Duo 3 GHz CPU with 3.25 GB RAM. Convergence of the MCMC has been evaluated using the R package `coda` [21]. Selected part of the R code for the examples is shown in the appendix. Additional output and more detailed explanation on how to use the package are offered in the on-line supplementary material and package vignettes.

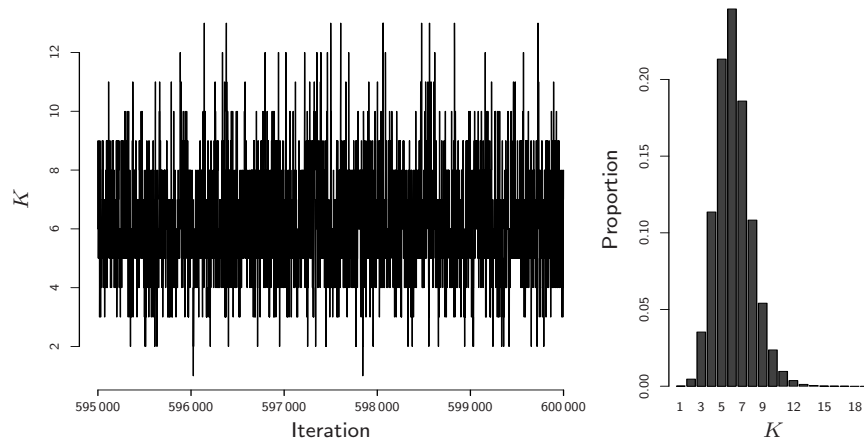


Figure 1: Galaxy data: traceplot (last 5 000 iterations) and histogram for the number of mixture components K .

6.1 1 dimension: Galaxy data

The galaxy data which give velocities (in km/sec) of 82 distant galaxies, diverging from our own galaxy were in the context of mixture modelling introduced by [24]. [23] estimated the velocity density using the RJ-MCMC and we repeat their analysis using the package `mixAK`. Hence the following prior distributions, their parameters and parameters of the proposal densities have been used: uniform prior on K with $K_{\max} = 30$, $\delta = 1$, semiconjugate prior on $\boldsymbol{\mu}$ and \boldsymbol{Q} with $\boldsymbol{\xi}_k = 21.73$, $\boldsymbol{D}_k = 630.5121$ for all k , $\zeta = 4$, $g_1 = 0.2$, $h_1 = 0.008$, $a_1 = b_1 = 2$, $a_2 = b_2 = 2$, $a_3 = b_3 = 1$. The data were neither shifted nor scaled before running the RJ-MCMC, i.e., $\boldsymbol{m} = 0$, $\boldsymbol{S} = 1$. We report results based on 500 000 iterations of 1:10 thinned RJ-MCMC obtained after a burn-in period of 100 000 iterations which took about 11 min.

During the course of the RJ-MCMC the chain visited models with the number of mixture components K ranging from 1 to 19 with the highest posterior probabilities of 0.11, 0.21, 0.25, 0.19, 0.11, and 0.05 for $K = 4, 5, 6, 7, 8$, and 9, respectively. For the remaining values of K , the posterior

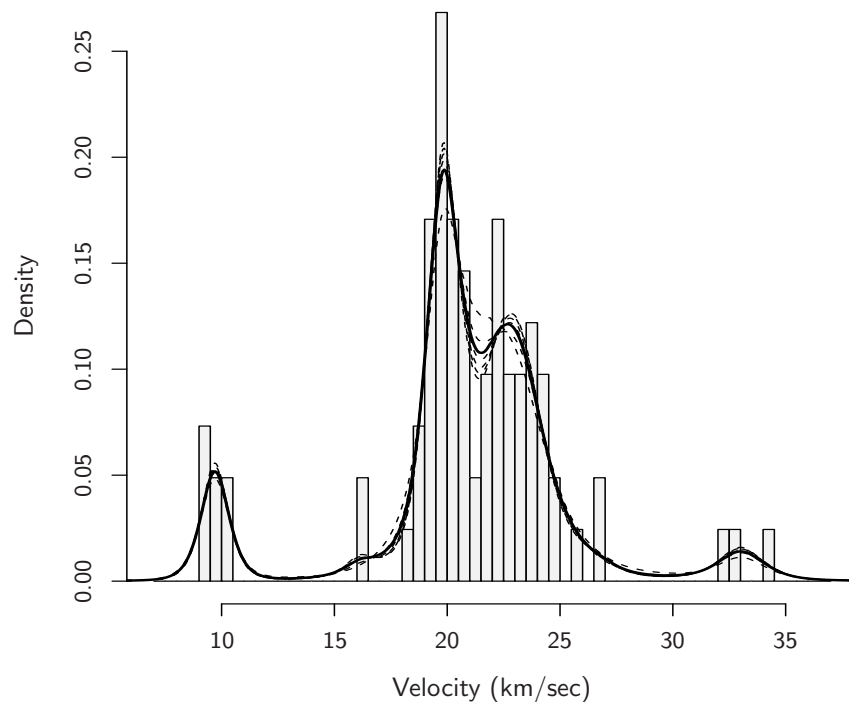


Figure 2: Galaxy data: histogram and predictive densities. Solid line: overall (unconditional) predictive density, dashed lines: conditional predictive densities for $K = 4, \dots, 9$.

probability was lower than 0.04, see Figure 1. The split-combine move has been accepted in 15% of cases whereas the birth-death move in 17% of cases. Consequently, a good mixing of the chain with respect to the number of mixture components has been obtained as is illustrated on the traceplot of K in Figure 1. From the posterior output, we further show for the purpose of comparison the posterior predictive density (17) and additionally, conditional (given K) posterior predictive densities of the velocity for $K = 4, \dots, 9$ in Figure 2 which corresponds to Fig. 2 (c) in [23].

6.2 2 dimensions: Old Faithful data

For our second example, we consider the Old Faithful data (version from [10]) analyzed using mixtures, e.g., by [28] or [5]. In sequence, we fitted one to ten component bivariate mixture to the data using the package `mixAK` with the following prior distributions: $\delta = 1$, semiconjugate prior on $\boldsymbol{\mu}$ and \boldsymbol{Q} with $\boldsymbol{\xi}_k = (-0.1207, -0.1028)'$, $\boldsymbol{D}_k = \text{diag}(9.4033, 15.1983)$, $\zeta = 3$, $g_1 = g_2 = 0.2$, $\boldsymbol{h} = (1.0635, 0.6580)'$. Due to the fact that the two margins are measured in two quite different scales, we have shifted and scaled observations in both margins by corresponding sample means and standard deviations, i.e., $\boldsymbol{m} = (3.488, 70.897)'$, $\boldsymbol{S} = \text{diag}(1.141, 13.595)'$. Reported results are based on 500 000 iterations of 1:10 thinned MCMC obtained after a burn-in period of 100 000 iterations. Sampling of one chain took between 6 min for a model with $K = 1$ up to 54 min for a model with $K = 10$.

Predictive densities based on the models with $K = 1, \dots, 6$ are shown in Figure 3 together with the values of PED and DIC. It is clear that at least two components are needed. Values of PED are quite similar for $K = 2$ and $K = 3$. The same is true for values of DIC. This coincides with previous results of others, see Fig. 8 (c) in [28] where according to the chosen prior model with $K = 2$ or 3 or 4 reached the highest posterior probability. In the analysis of [5], highest posterior probabilities of 0.3035 and 0.5854 have been obtained for $K = 2$ and $K = 3$, respectively.

Further, Figure 4 shows scatterplots of sampled mixture means in a three-component model and estimated posterior densities of mixture means without imposing any identifiability constraints or relabelling techniques. All three scatterplots are quite similar as well as estimated marginal densities and hence there is no serious indication that the chain would not visit all the modes of the posterior distribution.

6.3 6 dimensions with interval censoring: Signal Tandmobiël® data

Our third example considers the data from the Signal Tandmobiël® study [31] which was a dental longitudinal study conducted in Flanders in 1996–2001 involving 4 430 children born in 1989. We will analyze the emergence times of first six permanent teeth from the maxillary right quadrant of the mouth

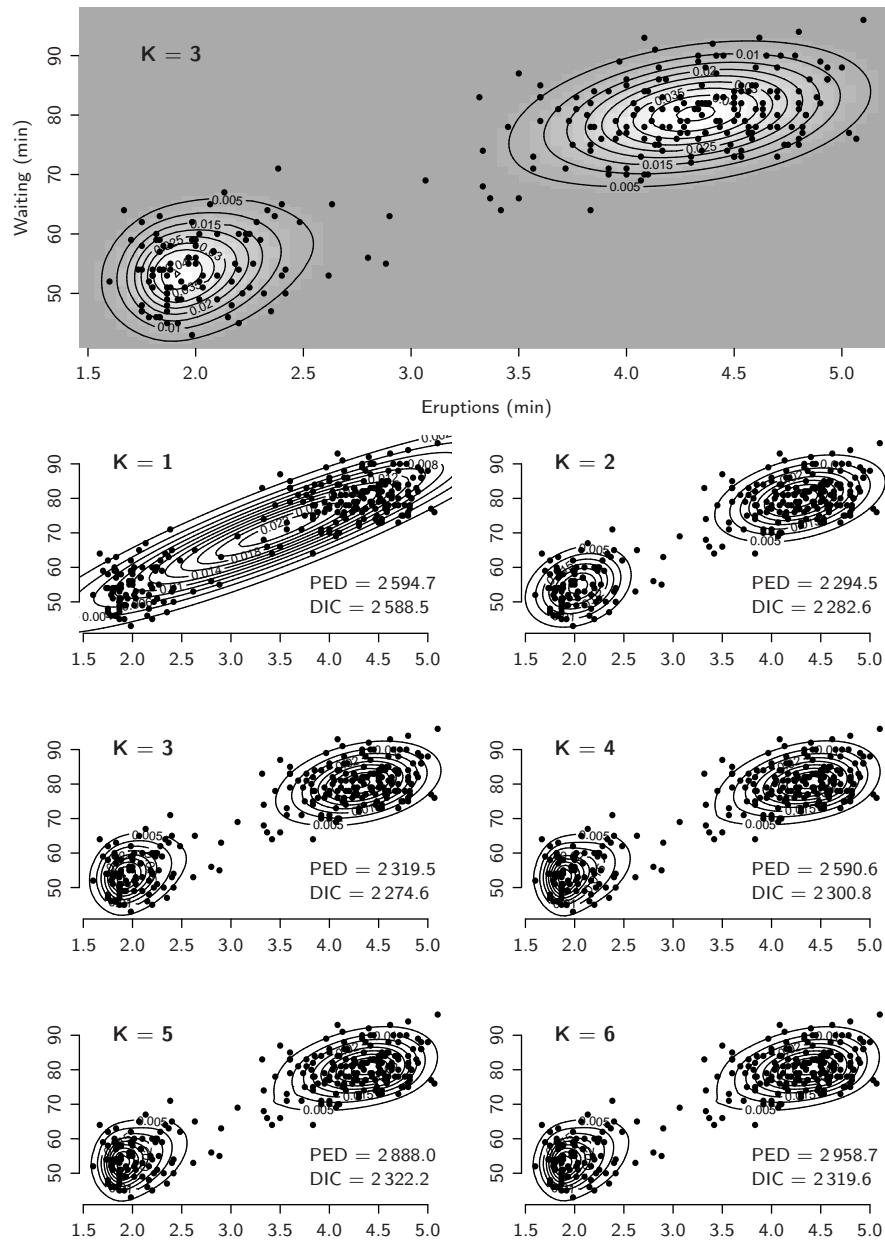


Figure 3: Old Faithful data: scatterplot and predictive densities for different values of K and obtained values of PED and DIC.

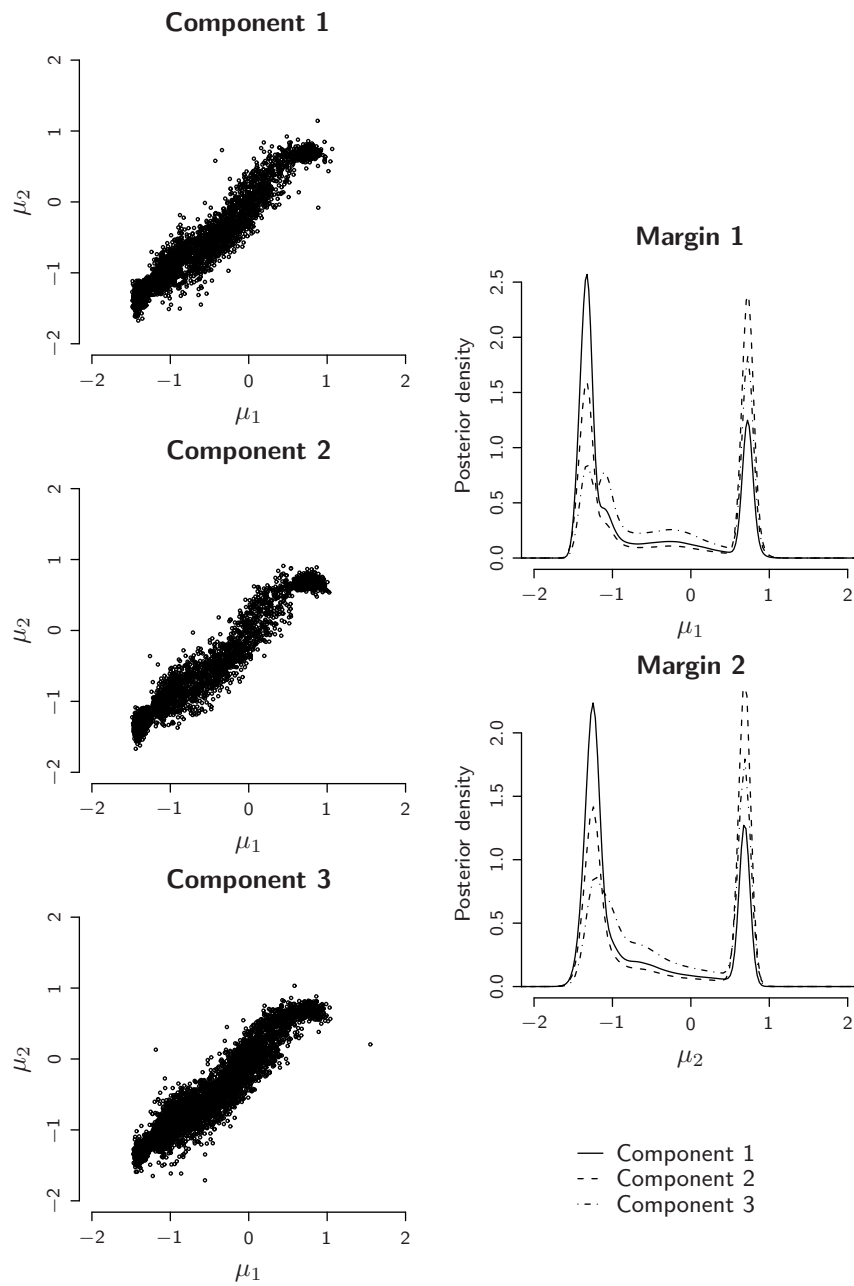


Figure 4: Old Faithful data: scatterplots of sampled mixture means (randomly selected 10 000 iterations) – left panel and estimated posterior densities of marginal mixture means of three components – right panel.

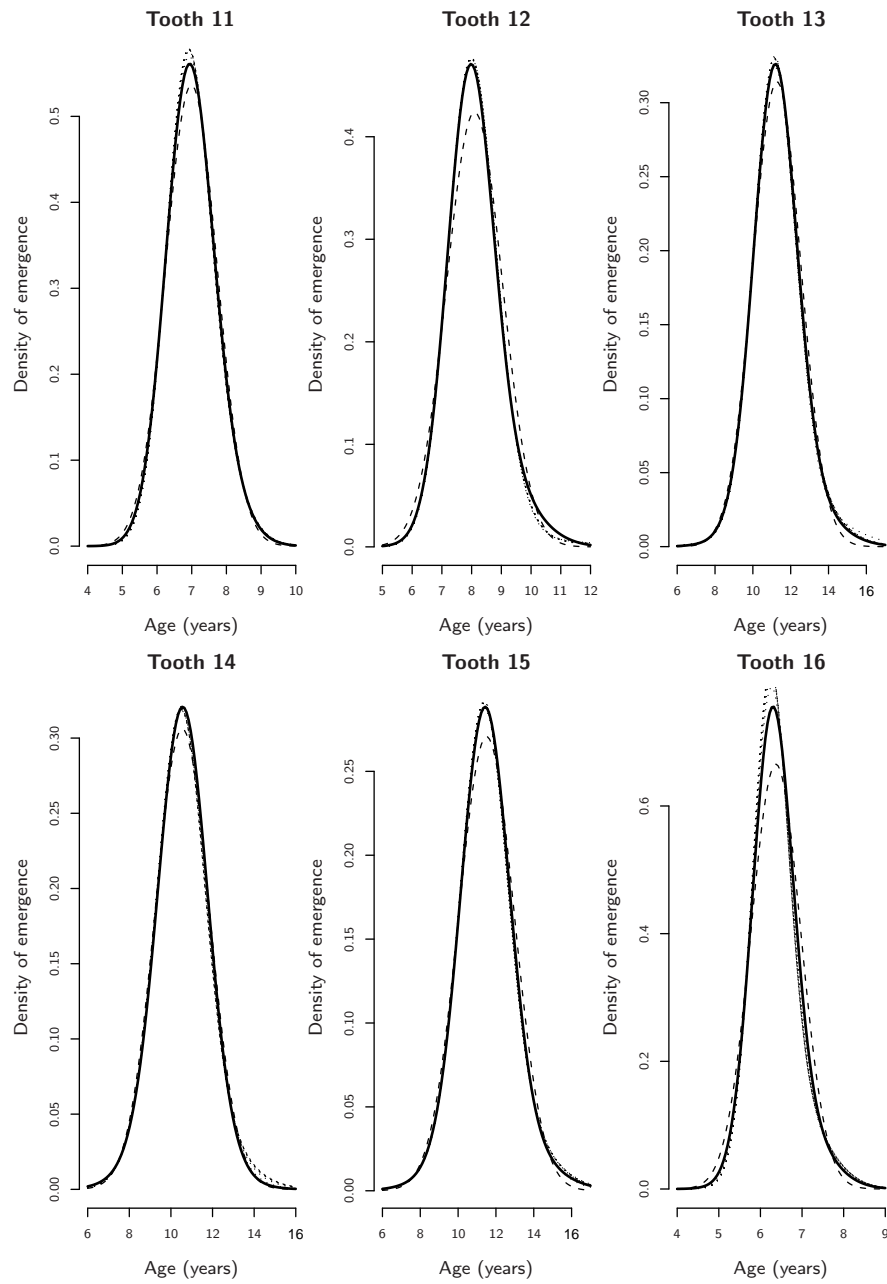


Figure 5: Signal Tandmobiél® data. Marginal predictive densities based on the models with different values of K , solid line for $K = 2$, dashed line for $K = 1$, dotted lines for $K = 3, \dots, 10$.

Tooth	Proportion censoring			Mean of emergence	Std. dev. of emergence
	Left	Interval	Right	Poster. 95% cred.	Poster. 95% cred.
				median interval	median interval
11	55.6%	39.4%	5.1%	7.00 (6.97, 7.03)	0.73 (0.71, 0.76)
12	14.7%	72.8%	12.6%	8.11 (8.08, 8.14)	0.93 (0.90, 0.96)
13	0.3%	47.3%	52.4%	11.30 (11.25, 11.36)	1.34 (1.27, 1.42)
14	0.7%	61.7%	37.6%	10.55 (10.50, 10.59)	1.30 (1.26, 1.35)
15	0.4%	42.0%	57.6%	11.52 (11.45, 11.59)	1.53 (1.45, 1.63)
16	86.0%	12.5%	1.5%	6.38 (6.35, 6.40)	0.58 (0.56, 0.60)

Table 1: Signal Tandmobiel[®] data. Posterior medians and 95% credible intervals for the mean and standard deviation of the emergence times.

(teeth 11, 12, 13, 14, 15, 16 in the European dental notation). In the course of the study children underwent annual dental examinations when emergence (among other things) was recorded. Hence, the emergence times are interval-censored with observed intervals of length of approximately 1 year or left-censored if the tooth was already present at the first examination or right-censored if the tooth has not emerged by the end of the study, see Table 1 for the amount of different types of censoring in the data. However, due to the fact that clinically, the permanent teeth hardly emerge before the age of 5 years [7] we have changed all left-censored observations into interval-censored ones with the lower limit of the observed intervals equal to 5 years for the purpose of computation.

Six-dimensional mixture estimated using the package `mixAK` is used as a tool for semiparametric density estimation and consequent inference on some characteristics of the joint emergence distribution of several teeth. In a sequel, models with fixed number of components $K = 1, \dots, 10$ have been fitted, resulting PED, DIC and other quantities compared. The following prior distributions have been used: $\delta = 1$, semiconjugate prior on $\boldsymbol{\mu}$ and \boldsymbol{Q} with $\boldsymbol{\xi}_k = (8.43, 9.60, 9.77, 9.76, 9.80, 7.98)'$, $\boldsymbol{D}_k = \text{diag}(31.9, 62.5, 67.9, 67.4, 68.9, 18.1)$ for all k , $\zeta = 7$, $g_1 = \dots = g_6 = 0.2$, $\boldsymbol{h} = (0.31, 0.16, 0.15, 0.15, 0.15, 0.55)'$. The data were neither shifted nor scaled before running the MCMC, i.e., $\boldsymbol{m} = (0, \dots, 0)'$, $\boldsymbol{S} = \text{diag}(1, \dots, 1)$. Results based on 20 000 iterations of the 1:10 thinned MCMC obtained after a burn-in period of 10 000 iterations are reported. Sampling of one chain took between 95 min for a model with $K = 1$ up to 135 min for a model with $K = 10$.

PED for a model with $K = 1$ reached the value of 78 853, dropped to 75 237 for $K = 2$ and started to increase through 75 784 ($K = 3$), 76 643 ($K = 4$), 77 224 ($K = 5$), 79 147 ($K = 6$), 80 758 ($K = 7$), 81 007 ($K = 8$), 82 212 ($K = 9$), 94 642 ($K = 10$). Graphical examination of the marginal predictive densities, see Figure 5, revealed that marginally the remarkable

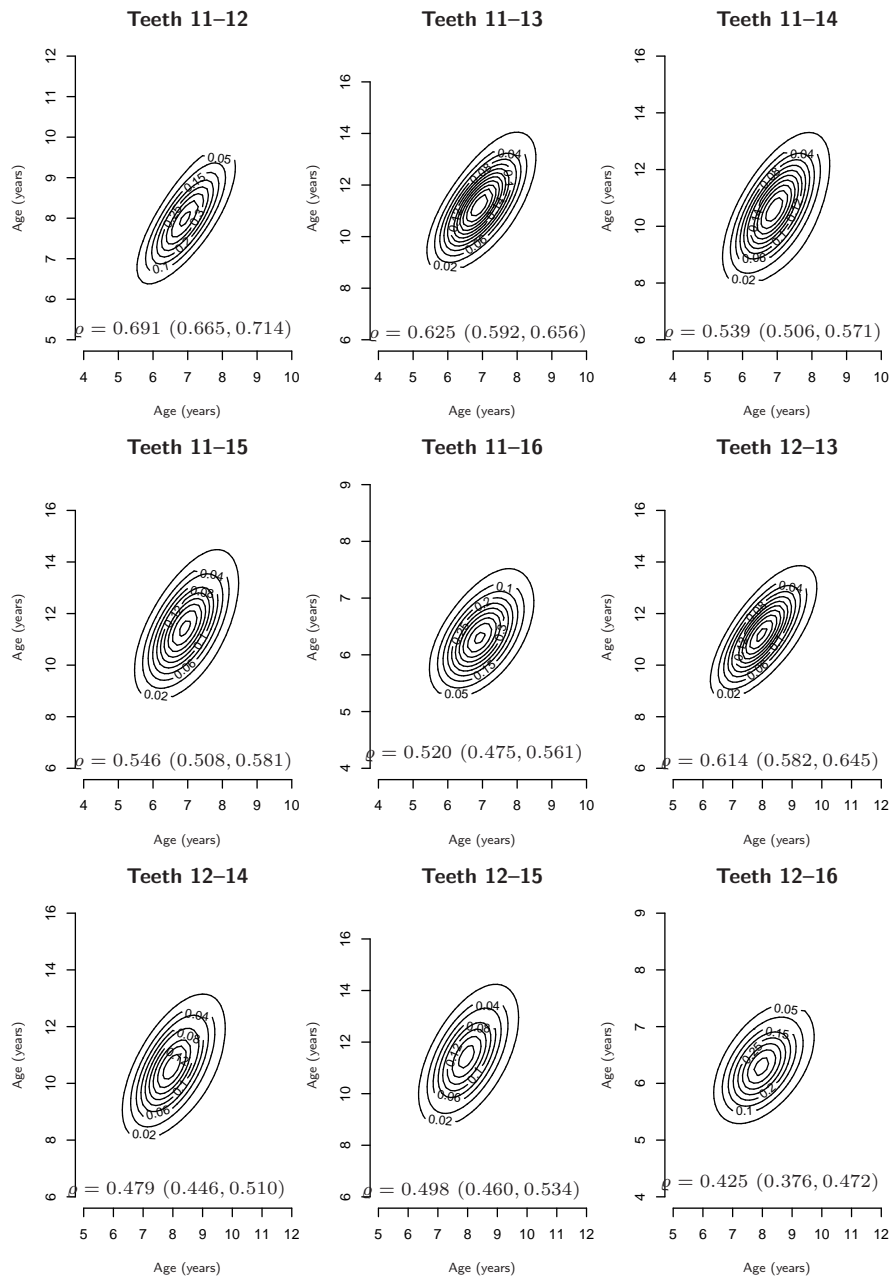


Figure 6: Signal Tandmobiél® data. Bivariate joint predictive densities (part I) based on the model with $K = 2$ and posterior median with 95% credible interval for the correlation between the two emergence times.

change (corresponding also to a considerable decrease of PED) in the shape of the estimated densities happens only when we switch from a model with $K = 1$ to a model with two or more mixture components. In the light of above findings, we will use the model with two mixture components as a suitable semiparametric structure to fit the density of the six emergence times.

For clinicians, it is useful to have an information concerning the timing and distribution of the emergence times, as well as the idea on how the emergence times of different teeth relate to each other. Estimated marginal densities of the emergence for different teeth have already been shown in Figure 5. From the MCMC output, it is quite easy to calculate posterior summary statistics for, e.g., the mean and the standard deviation of the emergence time of each tooth, see Table 1 where we report posterior medians and 95% credible inter-

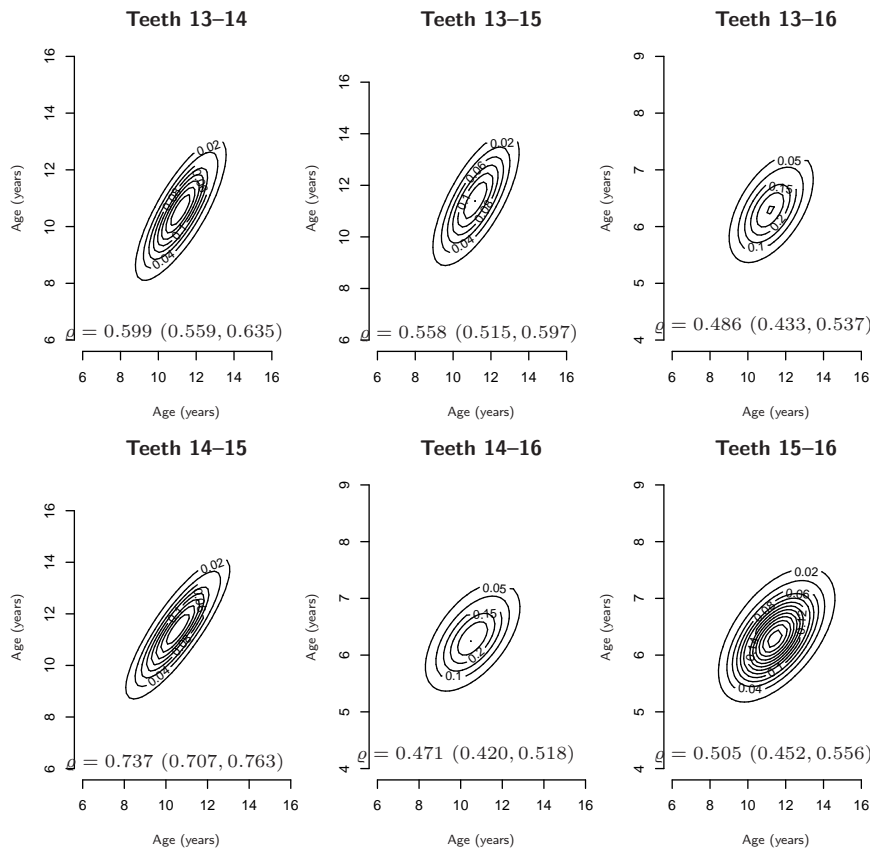


Figure 7: Signal Tandmobiel[®] data. Bivariate joint predictive densities (part II) based on the model with $K = 2$ and posterior median with 95% credible interval for the correlation between the two emergence times.

vals. Further, the idea on how the emergence times of different teeth relate to each other can be obtained from the posterior predictive densities of each pair of teeth or from the posterior summary statistics for the Pearson correlation coefficient, all shown in Figures 6 and 7. All mentioned quantities can be obtained in a straightforward manner using the functions of the package `mixAK`.

7 Concluding remarks

In this paper, we have introduced an R package which can be used in a straightforward manner for a density estimation using normal mixtures where at the same time, multivariate (interval-)censored observations can be considered. The package provides not only the core part of the estimation, i.e. MCMC, but also some routines for a consequent processing of the chains. That is, posterior summaries for several mixture related parameters are computed and posterior predictive densities can be easily computed and visualized. For mixture models of an arbitrary dimension, selection of the number of mixture components can be based on the penalized expected deviance or deviance information criterion, directly produced by the package routines.

For univariate problems, the package implements also the reversible jump MCMC algorithm allowing for a joint selection of the number of mixture components. According to our best knowledge, this has not been implemented in any of standardly used software packages. Even though quite recently, [15] presented a generic methodology for RJ-MCMC and implemented it in a WinBUGS framework. However, in the mixture context, they consider only birth-death moves. Due to the fact that their methodology aims to be generic for transdimensional models, split-combine moves, specific in a mixture setting and rather crucial for the success of the RJ-MCMC in this context are not considered.

It should be highlighted that it is not the ambition of the package to cover all nowadays available approaches to the (Bayesian) analysis of mixtures. Rather, it is the main intention of the package to allow inexperienced users easily get started with Bayesian mixture analyzes or allow for initial analyzes before switching to often time consuming coding of more advanced methods more suitable for a particular problem. The most important limitations of the package include the following. Firstly, in the selection of prior distributions, the user is limited to these described in subsection 2.2 and hence except the choice between the independent and natural conjugate priors for mixture means and variances, any sensitivity analysis is limited to changes of prior hyperparameters. Possible extension of the list of prior distributions and implementation of slightly different models is only possible by extension of the C++ code. Nevertheless, the list of offered prior distributions cover these mostly appearing in the literature on Bayesian analyzes of mixtures and should suffice for most initial analyzes. Further, it has been shown that usage of latent component allocations and corresponding Gibbs

algorithm decelerate convergence of the algorithm by the drastic increase in the dimensionality of the sampling space, see, e.g., [4, 2, 11]. Alternatives not requiring the use of latent allocations include, e.g., tempering MCMC [19] used by [4] and [11] or population and evolutionary MCMC (EMC) [14] used by [12]. In contrast to the Gibbs algorithm, both tempering MCMC and EMC require selection of several tuning parameters and hence their use might be too complicated to get started. Finally, there are other algorithms available for sampling from distributions of varying dimension than RJ-MCMC implemented in the package `mixAK`. For example, birth-and-death MCMC (BD-MCMC) introduced by [27] is a valuable competitor to RJ-MCMC. See also [2] for the link between RJ-MCMC and BD-MCMC and [25] for an overview of available algorithms for transdimensional sampling.

A R package

This appendix shows briefly how to use the package `mixAK` to obtain the results presented in Section 6. Detailed explanation can be found in on-line supplementary materials and package vignettes.

A.1 Galaxy data

```
### Load the data
> data(Galaxy)

### Specify prior and parameters for densities of u
> GalaxyPrior <- list(priorK="uniform", Kmax=30, delta=1,
+   priormuQ="independentC", xi=21.73, D=630.5121,
+   zeta=2*2, g=0.2, h=0.016/2)
> parRJMCMC <- list(par.u1=c(2, 2), par.u2=c(2, 2), par.u3=c(1, 1))

### Run MCMC
> GalaxyModel <- NMixMCMC(y0=Galaxy, prior=GalaxyPrior, RJMCMC=parRJMCMC,
+   nMCMC=c(burn=100000, keep=500000, thin=10, info=10000),
+   scale=list(shift=0, scale=1), PED=TRUE)

### Basic posterior summary
### and predictive density (computed from chain 1)
> print(GalaxyModel)
> GalaxyPDens <- NMixPredDensMarg(GalaxyModel[[1]])
> plot(GalaxyPDens)
```

A.2 Old Faithful data

```
### Load the data
> data(Faithful)

### Specify prior for a model with K = 3
> FaithfulPrior <- list(priorK="fixed", Kmax=3, delta=1,
+   priormuQ="independentC", xi=c(-0.1207, -0.1028),
+   D=diag(c(9.4033, 15.1983)), zeta=3, g=0.2, h=c(1.0635, 0.6580))
```



```

### Run MCMC with shifted and scaled data
> FaithfulModel <- NMixMCMC(y0=Faithful, prior=FaithfulPrior,
+   nMCMC=c(burn=100000, keep=500000, thin=10, info=10000), PED=TRUE)

### Basic posterior summary (including PED, DIC)
### and predictive density (computed from chain 1)
> print(FaithfulModel)
> FaithfulPDens <- NMixPredDensJoint2(FaithfulModel[[1]])
> plot(FaithfulPDens)

### Scatterplot of sampled mixture means in component 2 (chain 1)
> j <- 2
> plot(FaithfulModel[[1]]$mu[(j-1)*2+1], FaithfulModel[[1]]$mu[j*2])

```

A.3 Signal Tandmobiel[®] data

```

### Load the data, select only needed columns
> data(TandmobEmer)
> y0 <- TandmobEmer[, paste("EBEG.", 10+1:6, sep="")]
> y1 <- TandmobEmer[, paste("EEND.", 10+1:6, sep="")]
> censor <- TandmobEmer[, paste("CENSOR.", 10+1:6, sep="")]

### Specify prior for a model with K = 2
> TandmobPrior <- list(priorK="fixed", Kmax=2, delta=1,
+   priormuQ="independentC",
+   xi=c(8.43, 9.60, 9.77, 9.76, 9.80, 7.98),
+   D=diag(c(31.9, 62.5, 67.9, 67.4, 68.9, 18.1)),
+   zeta=7, g=0.2, h=c(0.31, 0.16, 0.15, 0.15, 0.15, 0.55))

### Run MCMC
> TandmobModel <- NMixMCMC(y0=y0, y1=y1, censor=censor,
+   prior=TandmobPrior,
+   nMCMC=c(burn=10000, keep=20000, thin=10, info=1000),
+   scale=list(shift=0, scale=1), PED=TRUE)

### Basic posterior summary (including PED, DIC), marginal univariate
### and pairwise bivariate predictive densities (computed from chain 1)
> print(TandmobModel)
> TandmobPDensUni <- NMixPredDensMarg(TandmobModel[[1]])
> plot(TandmobPDensUni)
> TandmobPDensBi <- NMixPredDensJoint2(TandmobModel[[1]])
> plot(TandmobPDensBi)

```

References

- [1] Böhning D., Seidel W., Alfó M., Garel B., Patilea V., Walther G. (2007). *Editorial: Advances in mixture models*. Computational Statistics and Data Analysis, **51**, 5205–5210.
- [2] Cappé O., Robert C.P., Rydén T. (2003). *Reversible jump, birth-and-death and more general continuous time Markov chain Monte Carlo samplers*. Journal of the Royal Statistical Society, Series B, **65**, 679–700.

- [3] Celeux G., Forbes F., Robert C.P., Titterton D.M. (2006). *Deviance information criteria for missing data models (with discussion)*. Bayesian Analysis, **1**, 651–706.
- [4] Celeux G., Hurn M., Robert C.P. (2000). *Computational and inferential difficulties with mixture posterior distributions*. Journal of the American Statistical Association, **95**, 957–970.
- [5] Dellaportas P., Papageorgiou I. (2006). *Multivariate mixtures of normals with unknown number of components*. Statistics and Computing, **16**, 57–68.
- [6] Diebolt J., Robert C.P. (1994). *Estimation of finite mixture distributions through Bayesian sampling*. Journal of the Royal Statistical Society, Series B, **56**, 363–375.
- [7] Ekstrand K.R., Christiansen J., Christiansen M.E. (2003). *Time and duration of eruption of first and second permanent molars: a longitudinal investigation*. Community Dentistry and Oral Epidemiology, **31**, 344–350.
- [8] Geweke J. (1991). *Efficient simulation from the multivariate normal and Student-t distributions subject to linear constraints and the evaluation of constraint probabilities*. Computer Sciences and Statistics, **23**, 571–578.
- [9] Green P.J. (1995). *Reversible jump Markov chain computation and Bayesian model determination*. Biometrika, **82**, 711–732.
- [10] Härdle W. (1991). *Smoothing Techniques with Implementation in S*. Springer Verlag, New York. ISBN 978-0-387-97367-8.
- [11] Jasra A., Holmes C.C., Stephens D.A. (2005). *Markov chain Monte Carlo methods and the label switching problem in Bayesian mixture modeling*. Statistical Science, **20**, 50–67.
- [12] Jasra A., Stephens D.A., Holmes C.C. (2007). *Population-based reversible jump Markov chain Monte Carlo*. Biometrika, **94**, 787–807.
- [13] Kass R.E., Raftery A.E. (1995). *Bayes factors*. Journal of the American Statistical Association, **90**, 773–795.
- [14] Liang F., Wong W.H. (2001). *Real parameter evolutionary Monte Carlo with applications to Bayesian mixture models*. Journal of the American Statistical Association, **96**, 653–666.
- [15] Lunn D.J., Best N., Whittaker J. (2005). *Generic reversible jump MCMC using graphical models*. Technical Report EPH-2005-01, Department of Epidemiology and Public Health, Imperial College, London.
- [16] Lunn D.J., Thomas A., Best N., Spiegelhalter D. (2000). *WinBUGS – A Bayesian modelling framework: Concepts, structure, and extensibility*. Statistics and Computing, **10**, 325–337.
- [17] McLachlan G.J., Basford K.E. (1988). *Mixture Models: Inference and Applications to Clustering*. Marcel Dekker, Inc., New York. ISBN 0-8247-7691-7.
- [18] McLachlan G.J., Peel D. (2000). *Finite Mixture Models*. John Wiley & Sons, New York.

- [19] Neal R. (1996). *Sampling from multimodal distributions using tempered transitions*. *Statistics and Computing*, **6**, 353–366.
- [20] Plummer M. (2008). *Penalized loss functions for Bayesian model comparison*. *Biostatistics*, **9**, 523–539.
- [21] Plummer M., Best N., Cowles K., Vines K. (2007). *coda: Output analysis and diagnostics for MCMC*. R package version 0.13-1.
- [22] R Development Core Team (2008). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org>. ISBN 3-900051-07-0.
- [23] Richardson S., Green P.J. (1997). *On Bayesian analysis of mixtures with unknown number of components (with discussion)*. *Journal of the Royal Statistical Society, Series B*, **59**, 731–792.
- [24] Roeder K. (1990). *Density estimation with confidence sets exemplified by superclusters and voids in the galaxies*. *Journal of the American Statistical Association*, **85**, 617–624.
- [25] Sisson S. (2005). *Transdimensional Markov chains: A decade of progress and future perspectives*. *Journal of the American Statistical Association*, **100**, 1077–1089.
- [26] Spiegelhalter D.J., Best N.G., Carlin B.P., van der Linde A. (2002). *Bayesian measures of model complexity and fit (with discussion)*. *Journal of the Royal Statistical Society, Series B*, **64**, 583–639.
- [27] Stephens M. (2000a). *Bayesian analysis of mixture models with an unknown number of components – an alternative to reversible jump methods*. *The Annals of Statistics*, **28**, 40–74.
- [28] Stephens M. (2000b). *Dealing with label switching in mixture models*. *Journal of the Royal Statistical Society, Series B*, **62**, 795–809.
- [29] Tanner M.A., Wong W.H. (1987). *The calculation of posterior distributions by data augmentation*. *Journal of the American Statistical Association*, **82**, 528–550.
- [30] Titterton D.M., Smith A.F.M., Makov U.E. (1985). *Statistical Analysis of Finite Mixture Distributions*. John Wiley & Sons, Chichester. ISBN 0-471-90763-4.
- [31] Vanobbergen J., Martens L., Lesaffre E., Declerck D. (2000). *The Signal-Tandmobiel[®] project – a longitudinal intervention health promotion study in Flanders (Belgium): baseline and first year results*. *European Journal of Paediatric Dentistry*, **2**, 87–96.

Acknowledgement: This work was supported by a grant MSM 0021620839. I want to thank two anonymous referees whose comments led to a considerable improvement of the manuscript. Data collection for the Signal Tandmobiel[®] project introduced in subsection 6.3 was supported by Unilever, Belgium. The Signal Tandmobiel[®] project comprises the following partners: D. Declerck (Dental School, Catholic University Leuven), L. Martens (Dental School,

University Ghent), J. Vanobbergen (Oral Health Promotion and Prevention, Flemish Dental Association), P. Bottenberg (Dental School, University Brussels), E. Lesaffre (Biostatistical Centre, Catholic University Leuven), K. Hoppenbrouwers (Youth Health Department, Catholic University Leuven; Flemish Association for Youth Health Care).

Address: A. Komárek, Dept. of Probability and Mathematical Statistics, Charles University, Sokolovská 83, CZ-186 75 Praha 8, Czech Republic

E-mail: arnost.komarek@mff.cuni.cz