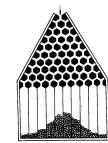


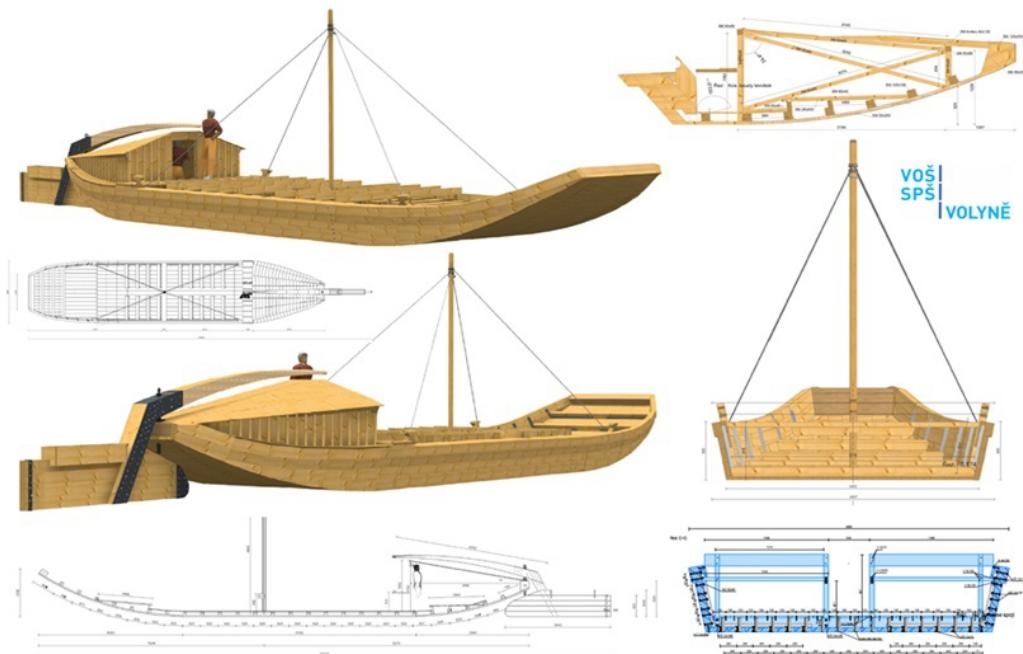
ROBUST 2022
VOLYNĚ
12. – 17. června 2022



RSJ



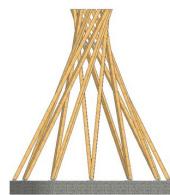
Sborník abstraktů



Organizátoři ROBUSTu 2022 by tímto chtěli poděkovat všem, kteří pomohli při přípravě a realizaci celé akce. Velmi důležitá pro nás byla morální i věcná podpora ČMS JČMF, ČStS, RSJ a KPMS MFF UK. Neméně důležitá byla podpora řady univerzit a sponsorů, ať již jmenovitých či anonymních, která především umožnila účast mnoha studentů, jakož udělení cen za nejlepší vystoupení studentů a/nebo doktorandů.

Nejdůležitější pro zdánlivý průběh ROBUSTu však bylo úsilí všech účastníků, které věnovali jak přípravě a prezentaci svých vystoupení, tak vytvoření skutečné „robustní atmosféry“. K příjemné pohodě též přispěla péče všech pracovníků VOŠ Volyně a laskavost pana ředitele Dr. J. Homolky, s níž se snažili našim přání vyjít vstříc.

Všem děkují a na setkání po čtyřiceti letech na odloženém dvacátém prvním Robustu v létě 2024 v Bardějově se těší organizátoři.



Úvodem několik historických mezníků z kroniky našich hostitelů

Léta Páně tisícího osmstého šedesátého čtvrtého za panování krále Františka Josefa a vladařství náměstka jeho v království Českém Richarda hraběte Belcrediho, když obec Volynskou spravovali purkmistr Jan Hödl a radové Jáchym Minibergr, František Harmach a Dr. Václav Niklas a když nad touto správou bděli výborové: Vojtěch Majer, Kajetán Kafka, Hynek Votava, Jan Alter, Jan Lom, Jindřich Fiala, Josef Prokop, Antonín Prokop, Tomáš Mařík, Filip Vanžura, Josef Prušák, Josef Lom, Felix Uhlík a František Boháč, když o duchovní správu péčí měl děkan P. František Gundermann *bylo ustavičně u veřejném mínění náryky na kleslost zanedbaného řemeslnictva Českého slyšeti a obcím kladeno na srdce, aby školy proň zakládaly.* Tedy obec Volynská byla první v Čechách, která nahlížíc potřebu školy průmyslové na návrh Vojtěcha Majera v zasedání zastupitelstva městského dne 3. září 1864 o zřízení takové školy jednohlasně se usnesla.

Dr. Majer, ředitel pražské průmyslové školy byl o radu požádán, s kterou ochotně radě městské přispěl. V sezéně zastupitelstva dne 15. září 1864 předložil osobně plán průmyslové školy, který jednohlasně schválen byv později tiskem vyšel. Plán se stal *Návrh ku zřízení jednotřídní denní průmyslové školy* sestavený od Dr. Antonína Majera pro město Volyni na školní rok 1865. *Účel takové školy jest, přiměřeně zaměstnávati a na živnosti připravovati mládež, která po zákonu rokem dvanáctým ve venkovských městech našich národní škole odrůstá. Zasedání v nejvyšší třídě městské školy po dvě-tři i více let nevidí se býti dosť prospěšným; na druhé straně ukazuje zkušenost, že není radno před rokem čtrnáctým dětí k řemeslům určovati, jednak proto, že dříve učeníku k řemeslu vlastně ani se užiti nedá a že se jich pak právě za tou příčinou užívá k věcem, které řemeslu neprospívají ba ani s ním nesouvisí, ano které zcela opačný výsledek mívají, poněvadž ducha, v tomto věku zvláště čilého, otupují a ku průmyslnému vedení živnosti pak docela nezpůsobilého činí.*

Jednotřídní průmyslová škola

V době vzniku v roce 1864 zněl název školy *Jednotřídní průmyslová škola ve Volyni.* Škola měla za úkol vzdělávat a vychovávat řemeslníky ze širokého okolí. Zpočátku škola vychovávala dorost pro různá řemesla a později také do ní chodili dospělí řemeslníci doplnit si vzdělání (např. zimní kurzy pro příslušníky řemesel stavebních od roku 1906). Vzhledem k tomu, že v okolí působilo hodně řemeslníků truhlářů, tesařů a zedníků, škola se postupně zaměřila na obory stavební a dřevozpracující.

Zemská průmyslová škola

V roce 1923 byly dosavadní zimní stavební kurzy zreorganizovány na mistrovskou školu stavební se zimním semestrálním vyučováním pro vyučené zedníky od sedmnácti let (později i pro dívky) a řemeslnická škola byla zrušena. Místo ní byla v roce 1924 zřízena odborná škola truhlářská, která poskytovala nejméně čtrnáctiletým žákům celoroční vyučování trvající tři roky, přičemž vysvědčení na odchodnou ze školy bylo rovnocenné s průkazem o vyučení se řemeslu. Přijímání byli hoši i dívky. Obě tyto části byly na sobě nezávislé pod názvem *Zemská průmyslová škola ve Volyni.* V roce 1933 začala výuka v nově vybudovaném komplexu, který byl vystavěn podle plánů Ing. Vladimíra Wallenfelse a Ing. arch. Františka Fialy. V těchto prostorách sídlí škola dodnes. I po roce 1945 zde pokračovala výuka truhlářů, tesařů a stavařů.

Střední průmyslová škola

V roce 1951 začíná výuka ve čtyřletém oboru pozemní stavitelství, v roce 1953 ve čtyřletém oboru technická zařízení budov, v roce 1960 ve čtyřletém oboru truhlářství (od roku 1984 nábytkářství) a v roce 1969 ve čtyřletém oboru dřevařství.

Vyšší odborná škola a Střední průmyslová škola

Od roku 1996 začíná výuka na vyšší odborné škole v oboru tvorba nábytku a dřevěné konstrukce se specializacemi tvorba nábytku a navrhování interiérů, a dřevěné konstrukce (dřevostavby). Počínaje školním rokem 2000/2001 byl na střední průmyslové škole zaveden ke čtyřem stávajícím oborům obor nábytkářská a dřevařská výroba, který umožňoval oddálit profesní specializaci v přípravě na povolání v celém rozsahu problematiky zpracování dřeva. Od školního roku 2006/2007 byly všechny tři dřevozpracující obory nahrazeny školním vzdělávacím programem *Interiérová tvorba, navrhování nábytku a dřevěné konstrukce.* Ve školním roce 2009/2010 byla zahájena výuka v uměleckém oboru *Design interiéru.*

Škola dnes vědomě navazuje na tradičně vysokou úroveň českého odborného školství a na svou historii, k níž se hrdě hlásí. I v budoucnosti chce škola vzdělávat a vychovávat odborníky pro oblast zpracování dřeva a stavebnictví, kteří budou navazovat na úspěchy více než dvanácti tisíc svých předchůdců, zdejších absolventů.

Radim Adolt	
<i>O užití statistiky pro účely monitoringu lesa a krajiny I: Představení NIL</i>	1
Radim Adolt	
<i>O užití statistiky pro účely monitoringu lesa a krajiny II: Matematicko-statistické principy NIL</i>	1
Friday I. Agu, Gábor Szűcs and Ján Mačutek	
<i>On Schröter family of discrete probability distributions</i>	2
Juraj Bodík, Linda Mhalla and Valérie Chavez	
<i>Causal inference for extremal dependence</i>	2
Michal Černý	
<i>Rank-estimátory, minimalizace Jaeckelovy disperze a elipsoidový algoritmus</i>	3
Jan Elgner, Jiří Ambros a Eva Fišerová	
<i>Posouzení efektu dopravních opatření upravujících rychlosť vozidel v intravilánu obcí</i>	3
Zdeněk Fabián	
<i>Geometrie výběrového prostoru parametrických rodin</i>	3
Kamila Fačevicová, Petra Kynčlová a Karel Macků	
<i>Modelování závislosti na vícerozměrné relativní struktuře pomocí prostorově vážené regrese</i>	4
Eva Fišerová a Veronika Římalová	
<i>Permutační testy pro testování hypotéz v prostorovém regresním modelu</i>	4
Vít Fojtík	
<i>Computation of central regions of the halfspace depth</i>	4
Laura Hajzoková	
<i>Metóda striedavých projekcií</i>	5
František Hendrych	
<i>Stochastické besovovské sewing lemma</i>	5
Zdeněk Hlávka	
<i>„Dokonalé“ dvouvýběrové funkcionální permutační testy</i>	5
Daniel Hlubinka a Šárka Hudcová	
<i>Pořadí a testy ve více rozměrech I</i>	6
Tomáš Hobza	
<i>Zobecněné lineární smíšené modely pro odhadování parametru malých oblastí</i>	6
Šárka Hudcová a Daniel Hlubinka	
<i>Pořadí a testy ve více rozměrech II</i>	6
Marie Hušková	
<i>Několik poznámek o funkcionálních datech</i>	6
Pavel Charamza	
<i>Zkažený statistik</i>	6
Markéta Janošová, Stanislav Katina a Jozef Hanes	
<i>Metody odhadu parametrů normálního rozdělení v přítomnosti dat pod měřitelným limitem</i>	6
Paulína Jašková, Karel Hron, Javier Palarea Albaladejo a Aleš Gába	
<i>Funkcionálna izotemporálna kompozičná substitučná analýza</i>	7
Tomáš Jurczyk	
<i>Moderní analytické platformy a analýza dat v praxi</i>	7
Karel Kadlec	
<i>Optimal control of Levy-driven stochastic equations in Hilbert spaces</i>	8
Lev Borisovič Klebanov	
<i>Pravděpodobnostní rozdělení s těžkými chvosty ve společenských vědách</i>	8
Daniel Klein, Jolanta Pielaśkiewicz and Katarzyna Filipiak	
<i>Approximate normality in testing an exchangeable covariance structure</i>	8
Jan Koláček	
<i>Analýza nepravidelně pozorovaných funkcionálních dat</i>	8
Arnošt Komárek	
<i>Jak jsme si nechali ukrást statistiku</i>	9
Miloš Kopa	
<i>Portfolio optimization with distortion risk measures</i>	9
Audrius Kabašinskas, Miloš Kopa, Karel Kozmík and Kristina Šutienė	
<i>Multicriteria crypto-portfolio optimization</i>	9
Petr Lachout	
<i>Obslužný systém (M/M/1)</i>	10
Lukáš Lafférs	
<i>Kauzálné strojové učenie</i>	10

Matúš Maciak, Michal Pešta and Gabriela Ciuperca	10
<i>Real-time changepoint detection in a nonlinear expectile model</i>	10
Patrícia Martinková	
<i>Výpočetní aspekty odhadu reliabilitu v případě vlivu kovariát a další témata z psychometrie</i>	10
Tomáš Masák	
<i>Inference and computation for sparsely sampled random surfaces</i>	11
Viktoria Nesrstová, Karel Hron, Peter Filzmoser and Ines Wilms	
<i>Identifikace významných párových logopodílů v kompozičních datech</i>	11
Jiří Novák	
<i>Simulace syntetických mikrodat z populačního censu</i>	12
Marek Omelka, Irène Gijbels and Vojtěch Kika	
<i>Estimation of tail coefficients based on copulas</i>	12
George Ostrouchov	
<i>High performance statistical computing with R</i>	12
Zbyněk Pawlas	
<i>Statistika krystalografických orientací</i>	12
Katerina Pawlasová	
<i>Úloha klasifikace pro replikované bodové vzorky</i>	13
Michal Pešta	
<i>Infinitely stochastic micro forecasting</i>	13
Ondřej Pokora	
<i>Stochastické modelování degradace technického systému</i>	13
Zuzana Prášková	
<i>Postupné změny v autoregresním modelu</i>	13
Samuel Rosa	
<i>O optimálnych návrhoch experimentov a najmenších elipsoidoch</i>	14
Pál Somogyi	
<i>Optimálna augmentácia experimentu</i>	14
Jan Vávra	
<i>Shluková analýza longitudinálních dat smíšeného typu</i>	14
Petr Vejmělka	
<i>Přesnost predikcí rezerv v neživotním pojištění u vybraných metod</i>	15
Ondřej Vencálek	
<i>Statistika v době covidové</i>	15
Ján Veselý a Jozef Kováč	
<i>Náhodné dynamické systémy pre predpísané stacionárne rozdelenia</i>	15
Petr Volf	
<i>A model of random walk with transition probabilities depending on the walk history</i>	16
Ondřej Vozář a Jaromír Antoch	
<i>Znáhodnené dotazování pro odhad populačního průměru kvantitativní proměnné</i>	16
Gejza Wimmer a Viktor Witkovský	
<i>Lineárny model v prípade nie normálne rozdelených meraní</i>	16
Viktor Witkovský and Gejza Wimmer	
<i>A note on computing the exact distribution of the bootstrap mean</i>	17
Markéta Zoubková	
<i>Modeling Bitcoin price using the attention factor</i>	17
Ivan Žežula a Daniel Klein	
<i>Vícenásobné testování středních hodnot u vícerozměrných dat</i>	17



Radim Adolf

O užití statistiky pro účely monitoringu lesa a krajiny I: Představení Národní inventarizace lesů

Analytické centrum NIL, Ústav pro hospodářskou úpravu lesů Brandýs nad Labem, pobočka Kroměříž, Náměstí Míru 498, 767 01 Kroměříž

adolt.radim@uhul.cz

Cílem přednášky je seznámení s Národní inventarizací lesů (NIL) v České republice [4], její náplní a způsobem provedení na principu výběrového šetření.

V úvodu přednášky jsou popsány různé účely a cíle zjišťování informací o lese. Následuje přehled o počátcích inventarizace lesů ve vybraných zemích a podrobnější informace o vývoji v České republice.

V navazující části je popsána statistická podstata provedení NIL v ČR spočívající v terénním sběru údajů na takzvaných inventarizačních plochách. Ty jsou umístěny na bodech vygenerovaných v rámci celé ČR pomocí vhodného výběrového designu. Kromě nezbytných metadat jsou na inventarizačních plochách zjištovány atributy území a prvků různých populací - stromů, kusů ležícího mrtvého dříví, vodních toků, lesních cest, vzácných nebo invazivních druhů rostlin, půdních horizontů, ale také například mysliveckých zařízení, odpadků a dalších typů objektů.

V rámci přednášky je na příkladu plochy lesa ukázáno, jak výrazně se mohou informace o stejném území pocházející z různých zdrojů lišit, a to nejen s ohledem na hodnoty konkrétních parametrů, ale také z pohledu jejich kvality, spolehlivosti a vhodnosti pro daný účel.

Přednáška je zakončena výčtem nedostatků celoplošného mapování (censu) z pohledu monitoringu lesa, respektive krajiny jako celku. Tyto nedostatky lze překlenout výběrovým šetřením. Poměrně často je možné mapování a výběrová šetření zkombinovat a získat tím nejen výhody obou jednotlivých postupů, ale i dosáhnout synergického efektu v podobě zvýšené přesnosti odhadu, snížených nákladů terénního šetření nebo obojího různou měrou, současně.

Reference

- [1] Kučera, M. and Adolf, R. (2019) Ústav pro hospodářskou úpravu lesů Brandýs nad Labem, Nábřežní 1326, 250 01 Brandýs nad Labem, Národní inventarizace lesů v České republice – výsledky druhého cyklu 2011–2015, ISBN 978-80-88184-23-2, ([dostupné on-line zde](#))

Radim Adolf

O užití statistiky pro účely monitoringu lesa a krajiny II: Matematicko-statistické principy Národní inventarizace lesů

Analytické centrum NIL, Ústav pro hospodářskou úpravu lesů Brandýs nad Labem, pobočka Kroměříž, Náměstí Míru 498, 767 01 Kroměříž

adolt.radim@uhul.cz

Cílem přednášky je představit matematicko-statistické principy Národní inventarizace lesů (NIL) v České republice [4] a současně pojmenovat některé aktuální úlohy, jejichž řešení se stále vyvíjí, přinejmenším z pohledu zavedení do běžné praxe.

V úvodu přednášky je pojednáno o funkci lokální hustoty podle Mandallaze [2], pp. 55–63 a [3], která umožňuje převedení výběru z konečné populace stromů na bodový výběr z populace kontinua. V rámci české NIL je lokální hustota s praktickými výhodami používána univerzálně pro výběry ze všech uvažovaných populací.

Další část přednášky je věnována problematice generování inventarizačních bodů a odhadu cílových parametrů NIL pomocí Horvitz-Thompsonova teorému pro populace kontinua [1]. Pozornost je dále věnována zvláštnostem vyplývajícím z neexistence design-based odhadů variability pro systematické a některé další výběry, standardně využívané v NIL.

Vzrůstající požadavky na aktuálnost a prostorový detail informací poskytovaných NIL, při současné snaze nezvýšit nebo dokonce snížovat náklady, vyžadují hledání nových, spolehlivých a přitom relativně snadno proveritelných typů odhadů. Jako perspektivní se jeví zejména design-based odhady využívající kromě terénních dat i pomocné zdroje informací, které lze získat s přiměřenými náklady. V této části přednášky jsou diskutovány regresní odhady s pomocí map a/nebo zahuštěné sítě inventarizačních bodů, jejich vlastnosti, dosavadní a očekávané využití v české NIL.

V závěru přednášky je nastíněn problém odhadu meziročních změn (např. změny plochy lesa) a dynamických veličin (např. přírůstu a těžby dříví) pro konkrétní rok, při použití trvalých inventarizačních bodů navštěvovaných s delší než roční periodou (v NIL typicky pět až deset let).

Reference

- [1] Cordy, C.B. (1993) An extension of the Horvitz-Thompson theorem to point sampling from a continuous universe. *Statistics and Probability Letters* 18, 353–362.
- [2] Mandallaz, D. (1991) A unified approach to sampling theory for forest inventory based on infinite population and superpopulation models. PhD thesis, Swiss Federal Institute of Technology (ETH), Zurich.
- [3] Mandallaz, D. (2007) Sampling Techniques For Forest Inventories. Chapman and Hall/CRC, Boca Raton.
- [4] Kučera, M. and Adolt, R. (2019) Ústav pro hospodářskou úpravu lesů Brandýs nad Labem, Nábřežní 1326, 250 01 Brandýs nad Labem, Národní inventarizace lesů v České republice - výsledky druhého cyklu 2011 – 2015, ISBN 978-80-88184-23-2. ([dostupné on-line zde](#))

Friday I. Agu^{1,2}, Gábor Szűcs³ and Ján Mačutek^{1,4}

On Schröter family of discrete probability distributions

¹Mathematical Institute, Slovak Academy of Sciences, Štefánikova 49, 81473 Bratislava, Slovakia

²Department of Statistics, University of Calabar, Calabar, Nigeria

³Department of Applied Mathematics and Statistics, Faculty of Mathematics, Physics and Informatics, Comenius University in Bratislava, Mlynská dolina, 84248 Bratislava, Slovakia

⁴Department of Mathematics, Faculty of Natural Sciences and Informatics, Constantine the Philosopher University in Nitra, Trieda Andreja Hlinku 1, 94901 Nitra, Slovakia

agu1@uniba.sk, gabor.szucs@fmph.uniba.sk, jmacutek@yahoo.com

One of standard problems in actuarial mathematics is modelling the aggregate claim distribution in the collective risk model when the claim severity distributions are concentrated on the non-negative integers and the number of claims is discrete. Theoretically, convolution could be used as a tool to derive the model, but given that the number of claims can be high, one often looks for simpler models. One of them was suggested in Schröter [2]. Schröter introduced recursive formula

$$P_n = \left(a + \frac{b}{n} \right) P_{n-1} + \frac{c}{n} P_{n-2}, \quad n = 1, 2, \dots, \quad (1)$$

which is a generalization of the well-known recursion by Panjer [1]. We note that in the formula above $P_{-1} = 0$ by definition. Despite the simplicity of Schröter formula, there are still open problems waiting to be solved. We will mention some of them in our presentation.

Acknowledgement: Supported by research project VEGA 2/0096/21.

References

- [1] Panjer, H.H. (1981) Recursive evaluation of a family of compound distributions. *ASTIN Bulletin* 12(1), 22 – 26.
- [2] Schröter, K.J. (1990) On a family of counting distributions and recursions for related compound distributions. *Scandinavian Actuarial Journal* 1990(2-3), 161 – 175.

Juraj Bodík, Linda Mhalla and Valérie Chavez

Causal inference for extremal dependence

University of Lausanne, Switzerland

jurobodik@gmail.com

This contribution deals with a combination of causal inference and extreme value theory (EVT). The part of EVT that we are dealing with concerns questions such as: “What is the dependence between two extreme events?” or “If the first random variable is large, what can we say about the second one?”. These questions are closely related to GEV distribution and extremal copula.

We want to show the causal causes of these phenomena. What is a causal driving factor of extreme events? Inferring this is (in most cases) impossible from just a random sample. However, if we perturb the system and observe different environments, we can observe some patterns of invariance that lead to causal covariates. We will show details of why and how we can use it for a more robust prediction of extreme events.

Michal Černý

Rank-estimátory, minimalizace Jaeckelovy disperze a elipsoidový algoritmus

VŠE, Katedra ekonometrie, Nám. W. Churchilla 4, 130 67 Praha 3

cernym@vse.cz

Jistou třídu robustních estimátorů v lineární regresi (tzv. rank-estimátory, R-estimátory) lze formulovat pomocí minimalizace Jaeckelovy disperze. Tuto optimalizační úlohu lze formulovat jako lineární program s $n!$ omezujićimi podmínkami, kde n značí počet pozorování. Není tudíž možné užít standardní LP-algoritmy, neboť počet omezujićích podmínek (nebo počet proměnných v duální formulaci) je extrémně velký. Jediný dosud známý polynomiální algoritmus je postaven na tzv. Grötschelově-Lovászově-Schrijverově elipsoidové metodě s oraculem; detaily jsou k dispozici v [M. Č. et al., A class of optimization problems motivated by rank estimators in robust regression, *Optimization, Latest articles*]. V příspěvku pojednáme rovněž o tom, proč je obtížné pro tento problém sestrojit metodu vnitřního bodu. (Poděkování patří také spoluautorům J. Antochovi, M. Radovi a M. Hladíkovi.)

Poděkování: Výzkum je podpořen grantem GA ČR P403/22/19353S.

Jan Elgner^{1, 2}, Jiří Ambros² a Eva Fišerová¹

Posouzení efektu dopravních opatření upravujících rychlosť vozidel v intravilánu obcí

¹ Katedra matematické analýzy a aplikací matematiky, Prírodovedecká fakulta Univerzity Palackého v Olomouci, 17. listopadu 12, 779 00 Olomouc

² Centrum dopravního výzkumu, v.v.i., Líšenská 33a, 636 00 Brno

jan.elgner01@upol.cz

Bezpečný dopravní prostor se týká každého z nás, ať už se v něm pohybujeme jako řidič, spolujezdec, cyklista nebo chodec při přecházení vozovky. Každý rok je v obcích ČR evidováno přibližně 5 900 dopravních nehod zapříčiněných nepřiměřenou rychlostí, pri kterých je registrováno cca 40 usmrcených a 2 000 zraněných účastníků silničního provozu. Nepriměřená rychlosť patří mezi nejvážnější příčiny dopravních nehod (14,8 % dopravních nehod). Správci komunikací používají řadu opatření ke snížení rychlosť vozidel, např. informační radary, ochranné ostruvky na přechodech nebo retroreflexní dopravní znacení. Jejich účinnost však často není známá.

V pilotní studii byla použita data z vozidel obsahujících GPS jednotku se záznamem polohy a rychlosti vozidla (frekvence záznamu 2–10 sekund) v obci Sardón de Duero ve Španělsku v letech 2020–2021. V obci byla realizována opatření ke snížení rychlosťi uprostřed uvažovaného období na obou koncích průtahu při vjezdu do obce. Cílem studie je identifikace rozdílu v rychlostních profilech před a po realizaci opatření a evaluace efektu realizovaných opatření. Na jednotlivé průjezdy vozidel lze nahlížet jako na spojitý proces, a proto navrhujeme zacházet se souborem studovaných dat jako s křivkami rychlosťi a pro modelování a analýzu použít metod funkcionální analýzy dat [2, 3]. Pro porovnání rychlostních profilů a identifikaci úseku, na kterých došlo vlivem opatření ke změně rychlosťi, navrhujeme využití metodiky dvouvýběrových funkcionálních intervalových testů [1].

References

- [1] Pini, A., Vantini, S. (2017) Interval-wise testing for functional data. *Journal of Nonparametric Statistics*, 29 (2), 407–424.
- [2] Ramsay, J.O., Silverman, B.W. (2013) Functional data analysis. New York, Springer.
- [3] Římalová, V., a kol. (2022) Modelling the driving speed on expressway ramps based on floating car data. *Measurement* 195, 110995.

Zdeněk Fabián

Geometrie výběrového prostoru parametrických rodin

ÚI AV ČR, Pod Vodárenskou Věží 2, 182 07 Praha 8

zdenek@cs.cas.cz

Indexovaná pravděpodobnostní míra na otevřeném intervalu představuje rodinu pravděpodobnostních rozdělení a generuje influenční funkci určující obecně nelineární geometrii na tomto intervalu. V této geometrii je typickou hodnotou rozdělení těžiště, charakteristikou variability je skórová variance, nabývající konečných hodnot i v případě rozdělení s těžkými konci, a sama geometrie napovídá, jak volit jejich validní estimátory.

Kamila Fačevicová, Petra Kynčlová a Karel Macků

Modelování závislosti na vícerozměrné relativní struktuře pomocí prostorově vážené regrese

Univerzita Palackého v Olomouci a United Nations Industrial Development Organization

kamila.facevicova@gmail.com

V rámci příspěvku se zaměříme na modelování vztahu mezi reálnou proměnnou a vícerozměrnou strukturou v podobě kompoziční tabulky. Představený přístup je prakticky motivován problémem analýzy vztahu mezi rižikem upadnutí do chudoby a vzdělanostní struktury osob ve věkové kategorii 30 až 34 let. K dispozici byly údaje z Německa z let 2000 a 2017, které byly sledovány na úrovni jednotlivých spolkových zemí. Toto členění nám mimo jiné umožňuje sledovat prostorová specifika dat. K analýze tak byla použita prostorově vážená regrese, zohledňující navíc relativní a vícerozměrnou povahu regresorů. Mimo nezbytné teorie, týkající se analýzy vícerozměrných kompozičních dat a prostorově vážené regrese, bude pozornost věnována také výslednému modelu a jeho interpretaci.

Eva Fišerová a Veronika Římalová

Permutační testy pro testování hypotéz v prostorovém regresním modelu s funkcionální závisle proměnnou

Katedra matematické analýzy a aplikací matematiky, Prírodovedecká fakulta Univerzity Palackého v Olomouci, 17. listopadu 12, 779 00 Olomouc

eva.fiserova@upol.cz

Cílem příspěvku je představit permutační přístup k testování hypotéz ve funkcionálním lineárním modelu [1] pro prostorová data. Vzhledem k prostorové závislosti mezi daty nejsou rezidua regresního modelu zaměnitelná, což poruší základní předpoklady permutačního schématu Freedmana a Lanea. Proto navrhujeme permutační test založit na približně zaměnitelných prostorově filtrovaných reziduích [2]. Simulační studie ukazuje, že zanedbání prostorové struktury reziduí v permutačním schématu (tj. přímá permutace korelovaných reziduí) vede k velmi liberálnímu testu, zatímco empirická velikost navrhovaného testu založeného na prostorově filtrovaných reziduích se blíží nominální hodnotě. Navržená metodika bude demonstrována na reálném datovém souboru o množství produkce odpadu v provincii Benátky v Itálii.

References

- [1] Ramsay, J.O., Silverman, B.W. (2013) Functional data analysis. New York, Springer.
- [2] Římalová, V. et al. A. (2022) Inference for spatial regression models with functional response using a permutational approach. JMVA 189, 104893.

Vít Fojtík

Computation of central regions of the halfspace depth

MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8

fojtik.vit@gmail.com

In one dimension, quantile methods provide a powerful toolkit for non-parametric statistical analysis. However, generalizing these tools into multiple dimensions is highly non-trivial. One such generalization is the well-known halfspace (or Tukey) depth, which is however notoriously difficult to compute and therefore to use in practice. In our paper, co-authored by P. Laketa, S. Nagy and P. Mozharovskyi, we study theoretical properties of a recent fast algorithm.

The algorithm provided in [1] for computation of central depth regions, corresponding to inter-quantile regions in one dimension, had been shown to return an exact result in a large simulation study. Surprisingly, we showed that this algorithm is not exact in general, but we proved exactness in some special cases. Additionally, we provided an extension of the algorithm that is exact and comparable to the original algorithm in the task of finding all central regions up to a given order.

References

- [1] Liu, X., Mosler, K. and Mozharovskyi, P. (2019) Fast computation of Tukey trimmed regions and median in dimension $p > 2$. Journal of Computational and Graphical Statistics 28 (3), 682–697.

Laura Hajzoková

Metóda striedavých projekcií

Ústav merania, Slovenská akadémia vied, Dúbravská cesta 9, 841 04, Bratislava

Fakulta matematiky, fyziky a informatiky, Katedra aplikovanej matematiky a štatistiky, Univerzita Komenského v Bratislave, Mlynská dolina F1, 842 48, Bratislava

laura.hajzokova@gmail.com

Cieľom príspevku je predstaviť a popísať metódu striedavých projekcií. Ide o iteračný algoritmus na nájdenie prieniku dvoch konvexných množín, ktorý využíva postupnosť ortogonálnych projekcií na tieto množiny. V krátkosti vysvetlíme teoretické pozadie metódy, popíšeme jej algoritmus a jeho fungovanie následne demonštrujeme na jednoduchých príkladoch. Uvedieme tiež pár špeciálnych prípadov, kedy sa dá projekcia na dané množiny vyjádriť explicitne, čím sa algoritmus značne zefektívni. Metóda nachádza využitie napríklad pri riešení lineárnych systémov alebo semidefinitných úloh prípustnosti, čo zahŕňa napríklad úlohy na doplnenie matice. Tie sa často vyskytujú pri dopĺňaní chýbajúcich dát alebo získavaní stratených informácií (z angl. *information recovery*). Tiež je jednou z metód, ktorá sa dá pre vhodné zvolený štartovací bod uplatniť aj pri nekonvexných úlohách.

František Hendrych

Stochastické besovovské sewing lemma

MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8

hendrych.frantisek@gmail.com

Sewing lemmata jsou užitečné nástroje při definici abstraktního integrálu skrze zadání lokální approximace. Nachází uplatnění například v teorii tzv. „rough paths“. Je známo mnoho různých verzí takových lemmat. Stochastické sewing lemma je speciální zobecnění, které umožnuje snížit požadavky na regularitu příslušných objektů. Ukážeme, že známé stochastické sewing lemma pro stochastické procesy chápáno jako Hölderovské funkce s hodnotami v $L^m(\Omega)$ predstavené v článku Lé [Electric Journal of Probability, 25:1-55, 2020] a besovovské sewing lemma pro funkce besovovského typu predstavené v článku Friz and Seeger, 2021, lze zobecnit. Přirozeným zobecněním je stochastické besovovské sewing lemma pro stochastické procesy vnímané jako funkce s hodnotami v $L^m(\Omega)$, jejichž regularita je besovovského typu.

Zdeněk Hlávka

„Dokonalé“ dvouvýběrové funkcionální permutační testy

MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8

hlavka@karlin.mff.cuni.cz

V odborné literatuře již bylo navrženo větší množství dvouvýběrových testů pro funkcionální nebo vysocerozměrná data [2, 3, 5, 6], nulové rozdělení testové statistiky přitom obvykle bývá značně komplikované a praktické provedení testu proto často spoléhá na tzv. permutační princip [1].

V příspěvku se zaměříme na funkcionální data a pojednáme o dvouvýběrové testové statistice založené na vhodné vzdálenosti empirických charakteristických funkcionálů, kterou integrujeme přes předem zvolenou rodinu pravděpodobnostních měr. Volba vhodné (Gaussovské) rodiny pak ovlivňuje sílu výsledného testu proti alternativám posunutí nebo změně měřítka [4].

Vlastnosti navrženého dvouvýběrového funkcionálního permutačního testu zkoumáme pomocí metod Monte Carlo, při jistém nastavení simulační studie a vhodné volbě parametrů použitého testu pak dokážeme nalézt i „dokonalý“ test, který dodržuje zvolenou hladinu a přitom prakticky vždy zamítá neplatnou nulovou hypotézu.

References

- [1] Good P.I. (2013) *Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypotheses*, Springer-Verlag, New York.
- [2] Górecki T., Smaga Ł. (2015) A comparison of tests for the one-way ANOVA problem for functional data. *Comput. Statist.* **30**(4), 987–1010.
- [3] Hall P., van Keilegom I. (2007) Two-sample tests in functional data analysis starting from discrete data. *Statist. Sinica* **17**(4), 1511–1531.
- [4] Hlávka, Z., Hlubinka, D., Koňasová, K. (2021) Functional ANOVA based on empirical characteristic functionals. *J. Multivariate Anal.* **189** 104878.
- [5] Székely G.J., Rizzo M.L. (2004) Testing for equal distributions in high dimension. *InterStat* **5**(16.10), 1249–1272.
- [6] Székely G.J., Rizzo M.L. (2010) Disco analysis: A nonparametric extension of analysis of variance. *Ann. Appl. Stat.* **4**(2), 1034–1055.

Daniel Hlubinka a Šárka Hudecová

Pořadí a testy ve více rozměrech I

MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8

hlubinka@karlin.mff.cuni.cz

Pořadové testy jsou již po mnoho let velmi oblíbené pro svou neparametrickou povahu a zajímavé vlastnosti. Jejich rozšíření na mnohorozměrné náhodné veličiny však stále narází na problém s definicí pořadí ve více než jenom rozměru. Již dlouho je zřejmé, že pořadí ve více než jednom rozměru musíme brát nikoliv ve smyslu „menší – větší“, ale v uspořádání „zevnitř – ven“, nejlépe ještě doprovázeného směrem.

Představíme pomérne nedávný koncept mnohorozměrných pořadí a znamének a ukážeme, jak odvodit asymptotickou normalitu lieárních statistik pomocí Hájkovy reprezentace a také jak vypadají mnohorozměrné analogie známých testů.

Tomáš Hobza

Zobecněné lineární smíšené modely pro odhadování parametrů malých oblastí

FJFI ČVUT, Katedra matematiky, Trojanova 13, 120 00 Praha 2

tomas.hobza@fjfi.cvut.cz

Tento příspěvek bude věnován problematice použití statistických modelů pro odhadování parametrů tzv. malých oblastí. Nejdříve bude stručně nastíněno, cím se obor statistiky nazývaný odhadování v malých oblastech (small area estimation) zabývá. Problematika pak bude ilustrována na konkrétním problému odhadování průměrného ročního příjmu nebo míry příjmové chudoby v oblastech španělské provincie Valencia. Protože příjem jako proměnná má asymetrické rozdělení a nelze ji většinou adekvátně modelovat přímo pomocí normálního rozdělení, je třeba enjdříve provést vhodnou transformaci nebo použít asymetrické rozdělení, jako např. gama rozdělení. V tomto příspěvku navrhujeme použití smíšeného gama modelu na úrovni individuů a na něm založených prediktorů charakteristik malých oblastí. Pro odhad střední kvadratické chyby prediktorů je navržen parametrický bootstrap. Závěrem bude prezentována reálná aplikace na datech ze španělského výběrového šetření příjmu a životních podmínek (SILC) v roce 2013.

Šárka Hudecová a Daniel Hlubinka

Pořadí a testy ve více rozměrech II

MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8

hudecova@karlin.mff.cuni.cz

V příspěvcích bude představena konstrukce mnohorozměrných pořadí a znamének pomocí optimálního přenosu míry a následně bude popsáno využití těchto konceptů ve statických testech jako je dvouvýběrový problém a mnohorozměrná analýza rozptylu.

Marie Hušková

Několik poznámek o funkcionálních datech

MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8

huskova@karlin.mff.cuni.cz

Jestliže jsou jednotlivá pozorování reprezentovaná funkcemi na nějakém intervalu či nějaké obecnější množině, obvykle mluvíme o funkcionálních datech. Jejich statistické zpracování přináší řadu problémů. Jedním z častých problémů je vyhodnocování funkcionálních dat, jestliže jsou pozorována jen v diskrétních bodech.

Příspěvek se bude týkat dvou metod pro řešení této situace. Jedna je založena na dizertaci nedávno obhájené na Univerzitě v Grazu, druhá využívá některé ideje z neparametrické regrese.

Pavel Charamza

Zkažený statistik

pavel.charamza@seznam.cz

Přednášející probere úlohy (primárně statistické), se kterými se setkal během své praxe. Půjde hlavně o problematiku ze světa financí, konkrétně kreditních rizik, ale drobné odbočky do světa výzkumu veřejného mínění nebo oblasti tzv. nových médií mu snad budou odpuštěny. Pokusí se zároveň vzpomenout si na metody, kterými se dané úlohy řešily, resp. o kterých se říká, že by mohly dané úlohy řešit. Pochopitelně půjde převážně o standardní statistické přístupy, ale dráždit bude i drobnými náznaky ve směru tzv. strojového učení.

Markéta Janošová^a, Stanislav Katina^a a Jozef Hanes^b**Metody odhadu parametrů normálního rozdělení v přítomnosti dat pod měřitelným limitem**^aÚstav matematiky a statistiky, PřF MU, Brno^bNeuroimunologický ústav, SAV, Bratislava

xjanosovam@math.muni.cz

Při použití laboratorního vybavení jsou měření omezena jeho limity. Obecně by pro každý přístroj měly být definovány limit slepého pokusu (limit of blank), detekční limit (limit of detection) a kvantifikační limit (limit of quantitation). Pokud pozorování spadá mimo měřitelný rozsah, vytvárá problém odhadu parametrů rozdělení. V tomto příspěvku zkoumáme použití čtyř metod: ignorování nezměřitelných pozorování, nahrazení nezměřitelných pozorování násobkem limitu, použití uříznutého normálního rozdělení a použití normálního rozdělení s cenzorovanými pozorováními.

Pro srovnání metod byla navržena simulační studie, kde vygenerované náhodné výběry z normálního rozdělení různých rozsahů byly zleva uříznuty ve vybraných kvantilech. Odhad parametrů získané pomocí všech metod byly porovnány pomocí relativní chyby $\frac{\hat{\mu} - \mu}{\hat{\mu}}$ a poměru $\frac{\hat{\sigma}}{\sigma}$.

Metody byly aplikovány na reálná laboratorní data (TnT, HCG, CRP, DDMPL, pT217) měřená z krve pacientů v laboratoři SK-Lab spol. s.r.o. v Lučenci a na Neuroimunologickém ústavu SAV v Bratislavě. Na základě výsledků simulační studie formulujeme doporučení pro praxi.

Příspěvek byl podpořen specifickým výzkumem Masarykovy univerzity na podporu studentských projektů (MUNI/A/1342/2021). Sběr dat a měření byly financovány firmou AXON Neuroscience SE, Bratislava.

Paulína Jašková¹, Karel Hron¹, Javier Palarea Albaladejo² a Aleš Gába³**Funkcionálna izotemporálna kompozičná substitučná analýza**

¹ Katedra matematické analýzy a aplikací matematiky, Přírodovědecká fakulta Univerzita Palackého v Olomouci, 17. listopadu 12, 771 46 Olomouc

² POLITÈCNICA IV, Campus Montilivi, 17003 - GIRONA

³ Fakulta tělesné kultury, Univerzity Palackého v Olomouci, třída Míru 117, 771 11 Olomouc

jaskovapaulina@gmail.com

Existuje vzťah medzi časom stráveným vo fyzickej aktivite (FA) a zdravím. Pri vyššom relatívnom podiele vysokointenzívnej FA na celkovú FA sa očakáva lepší zdravotný stav a naopak. Intenzitu FA možno opísť na spojitej stupnici prostredníctvom rastúceho gravitačného zrýchlenia pohybu [1]. Otázkou je, ako relatívne prerozdelenie času medzi FA rôznej intenzity súvisí so zdravotným stavom, teda ako nájsť vhodný regresný model tejto závislosti [2]. Relatívne údaje o využívaní času, označované ako kompozičné údaje s časťami zodpovedajúcimi kategóriám FA, možno rozšíriť na spojity prípad ako funkcie hustoty pravdepodobnosti (PDF). Podobne ako kompozičné údaje, aj PDF sa vyznačujú vlastnosťou invariantnosti mierky a sú geometricky reprezentované pomocou tzv. bayesovských priestorov. Po vhodnej transformácii PDF do priestoru L^2 možno použiť štandardný funkčný regresný model s reálnou premennou. Na opis vplyvu určitých podintervalov PDF gravitačného zrýchlenia možno použiť izotemporálnu substitúciu na zosilnenie vplyvu konkrétnych úrovní FA na zdravotný výsledok. Prezentovaná metóda bola aplikovaná na súbore údajov pozostávajúcich z funkčných pozorovaní, ktoré sú výsledkom rozsiahlej štúdie realizovanej medzi deťmi školského veku z Českej republiky. Premenné FA bola dosiahnutá pomocou trojosových akcelerometrov a modelovaná ako spojité funkčné premenné. Ako zdravotný ukazovateľ sa použilo percento hmotnosti telesného tuku. Výsledky potvrdzujú prirodzené očakávania, keďže viac času stráveného vo FA vyšej intenzite je spojené s nižším percentom tukovej hmoty.

References

- [1] Karas, M. et al. (2019) Accelerometry data in health research: challenges and opportunities. *Statistics in biosciences*, Springer, 210–237.
- [2] Talská, R. and Hron, K. and Grygar, T.M. (2021) Compositional scalar-on-function regression with application to sediment particle size distributions. *Mathematical Geosciences*, Springer, 1667–1695.

Tomáš Jurczyk**Moderní analytické platformy a analýza dat v praxi**

TIBCO Software

tjurczyk@tibco.com

Nároky a směr vývoje moderních nástrojů pro zpracování dat se v posledních letech dramaticky mění. K typickým výzvám patří například zpracování obrovského množství dat, dostupnost a připravenost dat k analýzám, zkrácení

času procesu od sběru dat až k využití získané informace k finálnímu rozhodnutí, problémy s nasazením finálních modelů do produkce nebo nedostatek kvalifikovaných statistiků. Cílem příspěvku je probrat tyto praktické výzvy a ukázat, jak se snaží analytické platformy tyto problémy řešit.

Karel Kadlec

Optimal control of Levy-driven stochastic equations in Hilbert spaces

TODO

Karel.Kadlec.9@seznam.cz

Controlled linear stochastic evolution equations driven by Lévy processes are studied in the Hilbert space setting. The control operator may be unbounded which makes the results obtained in the abstract setting applicable to parabolic SPDEs with boundary or point control. The first part contains some preliminary technical results, notably a version of Itô formula which is applicable to weak/mild solutions of controlled equations. In the second part, the ergodic control problem is solved: The feedback form of the optimal control and the formula for the optimal cost are found. The control problem is solved in the mean-value sense and, under selective conditions, in the pathwise sense. As examples, various parabolic type controlled SPDEs are studied.

Lev Borisovič Klebanov

Pravděpodobnostní rozdělení s těžkými chvosty ve společenských vědách

MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8

klebanov@karlin.mff.cuni.cz

V přednášce budou diskutovány vybrané důvody vedoucí k výskytu rozdělení pravděpodobnosti s těžkými chvosty při studiu vybraných sociálních procesů. Mezi tyto distribuce patří například Paretovo či Lotkovo rozdělení, jakož i některé další. Získané výsledky ukazují velmi významnou roli náhody v sociálních procesech.

Daniel Klein, Jolanta Pielaśkiewicz and Katarzyna Filipiak

Approximate normality in testing an exchangeable covariance structure

PF UPJŠ, Jesenná 5, 040 11, Košice, SR

daniel.klein@upjs.sk, jolanta.pielaskiewicz@liu.se, katarzyna.filipiak@put.poznan.pl

The Rao score test and likelihood ratio test for testing the hypothesis of exchangeable (called sometimes compound symmetry) structure in multivariate data are studied. The normal approximation of Rao score test under large- and high-dimensionality setting is derived as well as the exact and approximate distribution of the likelihood ratio test is determined. The tests are compared via simulations and it is demonstrated that the normal approximation of Rao score test matches well its distribution in high-dimensionality case. Thus, the normal approximation could be recommended for practice.

Acknowledgements: This work was supported by the Slovak Research and Development Agency under the contract no. APVV-17-0568 (Daniel Klein).

Jan Koláček

Analýza nepravidelně pozorovaných funkcionálních dat

Masarykova univerzita, ÚMS, Kotlářská 2, 611 37 Brno

kolacek@math.muni.cz

Funkcionální data jsou typicky tvořena funkcemi popisujícími křivky a plochy, a statistické metodologie pro práci s takovými daty se označují jako analýza funkcionálních dat. Příspěvek je motivován faktom, že reálně pozorovaná funkcionální data jsou často nestandardní povahy. Jedním z takových případů jsou situace, kdy trajektorie jsou pozorovány jen v relativně malém počtu okamžiků. Zaměříme se na využití vyhlazovacích technik při modelování kovariančních struktur v datech. Cílem je navrhnout rychlou a efektivní proceduru pro doplnování takovýchto dat.

Arnošt Komárek

Jak jsme si nechali ukrást statistiku

Matematicko-fyzikální fakulta Univerzity Karlovy, KPMS, Sokolovská 49/83, 186 75 Praha 8 – Karlín

komarek@karlin.mff.cuni.cz

Tato přednáška se měla původně uskutečnit v červnu 2020 v Bardejově. Pokud by se tak stalo, téma by se týkalo tzv. modelově založeného shlukování (*Model Based Clustering – MBC*) pro data s longitudinální strukturou, po přednášce by padlo několik málo dotazů a téma by utichlo s přednáškou dalšího řečníka. V červnu 2022 se již MBC s longitudinálními daty může zasvěceně věnovat i někdo jiný (Jan Vávra) a já bych tedy mohl začít z jiného soudku, s potenciálem pokračovat v jeho konzumaci i dlouho po skončení oficiálního denního programu. Události posledních dvou let spojené s šířením nového koronaviru a s ním souvisejícím onemocněním Covid-19 netřeba dlouze připomínat. Z pohledu statistiky a *statistické analýzy* dat je alarmující (alespoň pro mě), jak snadno na její místo nastoupili nejrůznější datoví analytici, matematictí modeláři, matematictí biologové a další se svými „daty podloženými“ závěry, které významně ovlivňovali dění okolo nás, avšak mnohdy zcela ignorovaly i ty nejzákladnější principy statistické inference. Skutečným statistikům to až na výjimky příliš nevadilo. Situaci velmi dobře ilustruje následující výrok Simona Wooda (v letech 2018–21 editora *JRSS B*): *“Something that the Covid-19 crisis has emphasised is the fact that statisticians have not managed to adequately communicate how fundamental random sampling is to proper measurement of things like infection rates. That data somehow related to the thing we want to measure are not the same as data that actually measure it.”*

Pokusím se o historické ohlédnutí za dobu uplynulých dvou let z pohledu statistika s důrazem na dění v Zemích koruny české a markrabství moravském. Přednáška bude mít i čistě odbornou část, ve které představím jeden možný přístup k hodnocení průběhu epidemie a vlivu „opatření“ na její průběh navržený výše zmíněným Simonem Woodem (Wood, 2021). Prezentované výsledky budou založeny na českých datech shromažďovaných pečí Ústavu pro Zkreslování Informací a Statistiky. Za spolupráci na této části přednášky děkuji Robertu Strakoví (Akademie Górnictvo-Hutnicza im. Stanisława Staszica, Kraków).

Obsah přednášky bude koordinován se související přednáškou Ondřeje Vencálka (*Statistika v době covidové*), aby nedocházelo k významnému překryvu.

References

- [1] Wood, S.N. (2021) Inferring UK COVID-19 fatal infection trajectories from daily mortality data: Were infections already in decline before the UK lockdowns? *Biometrics*, [on-line first](#).

Poznámka JA: Zde s panem kolegou nesouhlasím. Statistikům to vadilo, ale bylo jim to, jak se říká, prd platné. Dovolím si k tomu citovat z knihy M. Bártý, Sedm zákonů, JOTA, 2021, strana 144. . . . Televizní seriály stanovují obecné normy a tzv. tvůrci obecných názorů (opinion makers), často z řad zpěváků, herců a politiků, se bez bázně a hany vyjadřují mnohdy k věcem, o kterých většinou nic moc neví. Jejich názory pak vnucují média všem ostatním. . . Možná by pomohlo přidat při příští akreditaci v nabízeném curriculu vitae též něco z tohoto oboru.

Miloš Kopa

Portfolio optimization with distortion risk measures

MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8

kopa@karlin.mff.cuni.cz

The paper deals with distortion risk measures with a special focus on coherent distortion risk measures. Computationally attractive forms are presented under discrete distribution of returns (losses) of underlying assets. These measures are applied in the min-risk portfolio selection problems. The sensitivity analysis with respect to changes in distortion functions completes the paper.

Audrius Kabašinskas, Miloš Kopa, Karel Kozmík and Kristina Šutienė

Multicriteria crypto-portfolio optimization

MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8

kozmikk@karlin.mff.cuni.cz

We use a new consumer preference criterion based on stochastic dominance (SD). We add this criterion to the well-known mean-CVaR model to get mean-CVaR-SD model. The concept is tested on real data of returns of the 20 most valued cryptocurrencies. On a grid of weights for mean, CVaR and SD, we get portfolios of cryptocurrencies, which we evaluate using out-of-time observations. The evaluation is then performed using mean,

CVaR and maximum drawdown. We evaluate whether the portfolios are optimal for at least some combination of the evaluation criteria. The optimality of portfolios modelled by machine learning methods and the resulting models tested using out-of-time dataset. We also compare 2 scenarios, one which does not include the covid month April 2020 and one which does include it in the training dataset.

Petr Lachout

Obslužný systém (M/M/1)

MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8

lachout@karlin.mff.cuni.cz

Obslužný systém (M/M/1) je obslužný systém s jednou obslužnou linkou, s exponenciálně rozdelenými příchody zákazníku a s exponenciálně rozdelenou dobou obsluhy. Jedná se tedy o relativně jednoduchý homogenní Markovův řetězec se spojitym casem, ale není, alespoň pokud vím, dosud plně prozkoumán. V příspěvku se pokusím shrnout to, co je o nem známo, a trochu se pozabývám hledáním jeho rozdílení pravdepodobnosti.

Lukáš Lafférs

Kauzálné strojové učenie

Katedra matematiky, Fakulta prírodných vied, Univerzita Mateja Bela v Banskej Bystrici

lukas.laffers@gmail.com

Identifikácia a odhad kauzálnych efektov z neexperimentálnych dát je cieľom veľkej časti ekonometrie. S rastúcim objemom dát v súčasnosti prichádzajú do popredia metódy, ktoré vedia spracovať veľké množstvo informácií pomocou metód strojového učenia. Typicky sú metódy strojového učenia používané na predikciu, teraz je však objektom záujmu odhad parametrov a kvantifikácia štatistickej neistoty. Taktiež je zaujímavé skúmať heterogenitu efektov; napríklad pokúsiť sa identifikovať z dát, ktoré subjekty citlivejšie reagujú na danú intervenciu. V prezentácii predstavíme niektoré z týchto metód, nedávne pokroky ale aj ich limitácie.

Matúš Maciak, Michal Pešta and Gabriela Ciuperca

Real-time changepoint detection in a nonlinear expectile model

MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8; University Lyon-I

maciak@karlin.mff.cuni.cz

Regime switching within advanced stochastic models attracts a lot of interest over the last years with many different strategies being applied in this direction. We introduce a complex online changepoint detection procedure based on conditional expectiles. Nonlinearity of the underlying model improves the overall flexibility of the overall model, the conditional expectiles—well-known in econometrics for being the only coherent and elicitable risk measure—bring in some additional robustness, and the proposed changepoint detection test is proved to be consistent while the distribution under the null hypothesis depends on neither the functional form of the underlying model nor the unknown parameters which ensure very simple and straightforward applicability for real-life situations. Important theoretical details are summarized and finite sample empirical properties are presented.

Patrícia Martinková

Výpočetní aspekty odhadu reliability v případě vlivu kovariát a další téma z psychometrie

Ústav informatiky AV ČR, Oddělení statistického modelování, Pod Vodárenskou věží 2, 182 07 Praha 8

martinkova@cs.cas.cz

Měření v psychologii, pedagogice a v dalších sociálních, ale i lékařských vědách jsou často značně nepřesná a využívají většího množství položek či hodnotitelů. Z toho plynou některá specifika statistických úloh, které se v těchto oblastech vyskytují. V příspěvku pojednáme o výpočetních aspektech metod odhadů reliability a o dalších statistických úlohách. Nejprve se zaměříme na model analýzy rozptylu jednoduchého třídění s náhodnými efekty a na případy, kdy jsou odhady reliability nulové, viz [1]. Poté navrheme flexibilní přístup k odhadu reliability v případě heterogenity způsobené větším množstvím kovariát. V závěru pojednáme o dalších tématech z psychometrie s důrazem na analýzu jednotlivých položek vícepoložkových měření.

References

- [1] Erosheva, E., Martinková, P., Lee, C. (2021) When zero may not be zero: A cautionary note on the use of inter-rater reliability in evaluating grant peer review. *Journal of the Royal Statistical Society Series A (Statistics in Society)*, 184(2), 904–919. <<https://doi.org/10.1111/rssa.12681>>

Poděkování: Výzkum je podpořen grantem GA ČR 21-03658S.

Tomáš Masák

Inference and computation for sparsely sampled random surfaces

Institut de Mathématiques, École polytechnique fédérale de Lausanne

tom.masak@gmail.com

Non-parametric inference for functional data over two-dimensional domains entails additional computational and statistical challenges, compared to the one-dimensional case. Separability of the covariance is commonly assumed to address these issues in the densely observed regime. Instead, we consider the sparse regime, where the latent surfaces are observed only at few irregular locations with additive measurement error, and propose an estimator of covariance based on local linear smoothers. Consequently, the assumption of separability reduces the intrinsically four-dimensional smoothing problem into several two-dimensional smoothers and allows the proposed estimator to retain the classical minimax-optimal convergence rate for two-dimensional smoothers. Even when separability fails to hold, imposing it can be still advantageous as a form of regularization. A simulation study reveals a favorable bias-variance trade-off and massive speed-ups achieved by our approach. Finally, the proposed methodology is used for qualitative analysis of implied volatility surfaces corresponding to call options, and for prediction of the latent surfaces based on information from the entire data set, allowing for uncertainty quantification. Our cross-validated out-of-sample quantitative results show that the proposed methodology outperforms the common approach of pre-smoothing every implied volatility surface separately.

Viktoria Nesrstová¹, Karel Hron¹, Peter Filzmoser² a Ines Wilms³

Identifikace významných párových logopodílů v kompozičních datech založená na řídké metodě hlavních komponent

¹Katedra matematické analýzy a aplikací matematiky, Přírodovědecká fakulta Univerzity Palackého v Olomouci, 17. listopadu 12, 771 46, Olomouc

²Institute of Statistics and Mathematical Methods in Economics, Vienna University of Technology, Wiedner Hauptstrasse 8-10, 1040 Vienna, Austria

³Department of Quantitative Economics, Maastricht University, Tongersestraat 53, Maastricht, The Netherlands
viktoria.nesrstova@gmail.com, karel.hron@upol.cz, peter.filzmoser@tuwien.ac.at,
i.wilms@maastrichtuniversity.nl

Kompoziční data jsou tvořena pozorováními, jejichž relevantní informace je obsažena v podílech mezi složkami [3]. Častou reprezentací těchto dat jsou logopodílové souřadnice sestrojené vzhledem k dané ortonormální bázi, které vždy agregují více logopodílů formou obecnějšího matematického objektu, tzv. logkontrastu.

Pro vysoce dimenzionální data je navíc analýza všech párových logopodílů problematická, neboť s rostoucím počtem proměnných nejméně roste počet párových logopodílů, což je náročné z hlediska výpočtů i interpretace výsledků. Tento příspěvek se zabývá analýzou párových logopodílů a identifikací těch významných pro mnohorozměrnou datovou strukturu v kontextu redukce dimenze. Toho je dosaženo využitím řídké metody hlavních komponent [2] založené na konstrukci tzv. zpětných pivotových souřadnic [1], které z celého souřadnicového systému zdůrazňují právě párové logopodíly. Identifikace signifikantních logopodílů je výsledkem rovnováhy mezi dostatečnou řídkostí vektorů zátěží a dostatečným množstvím vysvětlené celkové variability. Výsledky budou demonstrovány na reálných i simulovaných datech.

References

- [1] Hron, K. et al. (2021) Analysing pairwise logratios revisited, *Mathematical Geosciences* 53, 1643–1666.
[2] Erichson, N.B. et al. (2020) Sparse principal component analysis via variable projection, *SIAM Journal on Applied Mathematics* 80, 977–1002.
[3] Filzmoser, P., Hron, K. and Templ, M. (2018) *Applied Compositional Data Analysis*. Springer, Heidelberg.

Jiří Novák

Simulace syntetických mikrodat z populačního censu

FIS VŠE, KEST, nám. Winstona Churchilla 1938/4, 120 00 Praha 3

xnovj159@vse.cz

Populační census je unikátním zdrojem dat, který obsahuje řadu cenných informací o populaci v jedinečném geografickém detailu. Mikrodata, která obsahují sebrané údaje na úrovni jedince, jsou však z pohledu ochrany osobních údajů velice citlivá a z daného důvodu jsou diseminovaná pouze agregovaně ve formě tabulek. Příspěvek bude pojednávat o možnosti využití metod simulace syntetických mikrodat pro data z populačního censu. V rámci těchto metod je vytvářen syntetický (nový/umělý) dataset mikrodat, ve kterém jsou ovšem zachovány vztahy a struktura jako v původních mikrodatabech. Cílem je vytvoření syntetického datasetu, který bude analyticky validní, ale absence reálných hodnot umožní jeho diseminaci přes SafeCentrum na Českém statistickém úřadě.

Marek Omelka, Irène Gijbels and Vojtěch Kika

Estimation of tail coefficients based on copulas

MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8 and KU Leuven, Statistics and Risk, Belgie

omelka@karlin.mff.cuni.cz

Multivariate tail coefficients are of interest when one is trying to summarize dependence of extremes. The difficulty is that these coefficients are defined as limits and thus to estimate them from data one needs to choose a kind of smoothing parameter k_n . In this talk we will discuss several ways how one can choose this smoothing parameter k_n .

George Ostrouchov

High performance statistical computing with R

Oak Ridge National Laboratory and University of Tennessee and MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8

ostrouchovg@ornl.gov

The tutorial is aimed at a basic understanding of high performance computing and its use from the R language. We begin with basic concepts of parallel computing hardware that make up today's devices ranging from your laptop to a cluster computer, like Karolina at the IT4 Innovations Czech Supercomputing Center. We introduce native language interfaces to parallel computing concepts with a focus on their access from the R language. Many packages in R provide abstracted access to parallel computing, yet there are only a few underlying concepts. We take a narrow path that looks at the underlying concepts and their direct and high-level access from R. Specifically, these include shared-memory multithreaded libraries and the unix fork, and distributed-memory MPI computation. Some examples will be demonstrated on a 4-core laptop while other examples shown will need access to a cluster computer, which the participants can later try on their own. Provided materials will include examples of submission scripts for cluster computers.

Zbyněk Pawlas

Statistika krystalografických orientací

MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8

pawlas@karlin.mff.cuni.cz

V různých vědních oblastech se vyskytují data ve formě rotací kolem počátku souřadnic v trojrozměrném eukleidovském prostoru \mathbb{R}^3 . Takové rotace spolu s operací skládání tvoří grupu $SO(3)$. V krystalografii lze orientaci krystalické mřížky popsat rotací referenčního souřadného systému do souřadného systému mřížky. Vzhledem k symetrii mřížky v dané krystalografické soustavě nelze některé rotace rozlišit, neboť vedou na stejnou mřížku. Označíme-li jako K podgrupu všech těchto symetrických rotací, pak prostor krystalografických orientací je příslušná faktorová grupa $SO(3)/K$. Při statistické analýze je potřeba zohlednit specifika spojená s tímto prostorem. V přednášce se zaměříme na některé statistické úlohy pro data představující krystalografické orientace. Především se budeme věnovat testování nezávislosti.

Katerina Pawlasová

Úloha klasifikace pro replikované bodové vzorky

MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8

pawlasova@karlin.mff.cuni.cz

Bodovým vzorkem rozumíme kolekci bodů v d -rozměrném euklidovském prostoru. Tyto body mohou reprezentovat například pozorované středy buněk ve vzorku tkáně nebo středy bublinek vzduchu ve zkoumané pevné látce. V naší úloze předpokládáme, že bodové vzorky, které máme k dispozici, pocházejí z J skupin, kde J je předem známo. Úkolem je poté nově pozorovaný bodový vzorek zařadit do jedné ze skupin. Dále předpokládáme, že máme k dispozici kolekci bodových vzorků, pro které známe jejich skupiny. Tuto kolekci nazýváme trénovací data.

Pro úlohy strojového učení, mezi které úloha klasifikace spadá, jsou v dnešní době typická řešení s využitím neuronových sítí. V příspěvku představíme sérii simulačních experimentů, které ilustrují využití neuronových sítí pro bodové vzorky. V prvním případě použijeme reprezentaci bodového vzorku pomocí binárního obrázku. Ke klasifikaci poté využijeme konvoluční neuronové sítě. Ve druhém případě se budeme zajímat o neuronové sítě, které na vstupní vrstvě umí zpracovávat funkcionální data. V tomto případě můžeme pro popis bodového vzorku využít funkcionální charakteristiky, jako je například párová korelační funkce. Řešení za pomoci neuronových sítí porovnáme s naším předchozím výzkumem, kde byla stejná úloha řešena za pomoci jádrové regrese. Příspěvek je založen na spolupráci s Jiřím Dvořákem.

Michal Pešta

Infinitely stochastic micro forecasting

MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8

michal.pesta@mff.cuni.cz

Stochastic forecasting and risk valuation are now front burners in a list of applied and theoretical sciences. In this work, we propose an unconventional tool for stochastic prediction of future expenses based on the individual (micro) developments of recorded events. Considering a firm, enterprise, institution, or any entity, which possesses knowledge about particular historical events, there might be a whole series of several related subevents: payments or losses spread over time. This all leads to an infinitely stochastic process at the end. The aim, therefore, lies in predicting future subevent flows coming from already reported, occurred but not reported, and yet not occurred events. The emerging forecasting methodology involves marked time-varying Hawkes process with marks being other time-varying Hawkes processes. The estimated parameters of the model are proved to be consistent and asymptotically normal under simple and easily verifiable assumptions. The empirical properties are investigated through a simulation study. In the practical part of our exploration, we elaborate a specific actuarial application for insurance claims reserving.

Ondřej Pokora

Stochastické modelování degradace technického systému

Ústav matematiky a statistiky, Přírodovědecká fakulta, Masarykova univerzita, Kotlářská 2, 611 37 Brno

pokora@math.muni.cz

Analýza a modelování degradace technického systému pomocí diagnostických signálů je jedním ze způsobů zajištění jeho spolehlivosti a plánování jeho údržby. Příspěvek se zabývá statistickou analýzou vlastností oleje v systému pohonné jednotky těžkého nákladního vozidla pro predikci degradace, odhad rizika výskytu poruchy a zbytkové životnosti systému. Ukážeme stochastický model diagnostických charakteristik oleje založený na metodách analýzy funkcionálních dat, zobecněných aditivních modelech a difuzních procesech.

Zuzana Prášková

Postupné změny v autoregresním modelu

MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8

praskova@karlin.mff.cuni.cz

V příspěvku budeme uvažovat autoregresní proces s časově se měnícími koeficienty, který je obecně nestacionární, ale může být lokálně stabilní. Speciálně se budeme zabývat autoregresním procesem 1. rádu, jehož parametr (autokorelace 1. rádu) se mění postupně v závislosti na určité funkci, a odhadem bodu této postupné změny.

Vedle asymptotických vlastností takového odhadu (konzistence, limitní rozdělení) se budeme zabývat rovněž numerickými vlastnostmi a bootstrapovými variantami.

Poděkování: Příspěvek vznikl ze spolupráce s Prof. RNDr. Marií Huškovou, DrSc. (MFF UK) a prof. Josefem Steinebachem (Univerzita Kolín nad Rýnem). Práce Z. Práškové byla podpořena grantem EXPRO GX19-28231X.

Samuel Rosa

O optimálnych návrhoch experimentov a najmenších elipsoidoch

Univerzita Komenského v Bratislave, Bratislava

samuel.rosa@fmph.uniba.sk

Problém nájsť elipsoid s čo najmenším objemom, ktorý pokrýva danú množinu bodov, a problém nájsť čo najinfomatívnejší návrh experimentu (konkrétnie D-optimálny approximatívny návrh) sa zdajú byť úplne nesúvisiacimi úlohami, predsa však spolu súvisia a dokonca sú ekvivalentné. V príspevku najprv priblížime alebo pripomeneieme túto známu ekvivalenciu. Následne sa budeme venovať metódam, ktoré hľadajú počiatočné riešenie pre algoritmy na výpočet optimálnych návrhov, resp. najmenších elipsoidov. Ukážeme, že konštrukcia kvalitného počiatočného riešenia má zaujímavé technické úskalia a sledujúc [1] budeme analyzovať heuristiky na konštrukciu spomínaných počiatočných riešení, napríklad ukážeme ich prekvapivo pekné geometrické interpretácie.

Podakovanie: Práca bola podporená grantom VEGA 1/0362/22 Vedeckej grantovej agentúry SR.

Literatúra

- [1] Harman R., Rosa S. (2020) On greedy heuristics for computing D -efficient saturated subsets, *Operations Research Letters* **28**, 122–129.

Pál Somogyi

Optimálna augmentácia experimentu

Fakulta matematiky, fyziky a informatiky Univerzity Komenského, Mlynská dolina F1, SK – 842 48 Bratislava

pal.somogyi@fmph.uniba.sk

Ciel'om výskumu je vypracovať výpočtové nástroje pre problém optimálnej augmentácie experimentu (viď [1] a [3]), ktorého pozorovania zodpovedajú lineárному alebo nelineárному regresnému modelu. Ako východzíu metódu použijeme znáhodnený výmenný algoritmus REX z článku [2] vyvinutý pre situáciu, v ktorej majú elementárne informačné matice hodnosť jedna, čo pre naplnenie vytýčeného cieľa nepostačuje. Okrem mnohých iných aplikácií, augmentácia experimentu je dôležitá v oblasti takzvaného aktívneho učenia (active learning), ktoré sa používa na optimálne sekvenčné označovanie dosiaľ neoznačených dát (unlabeled data) v oblasti strojového učenia.

Literatúra

- [1] Atkinson, A.C., Donev, A.N., Tobias, R.D. (2007) *Optimum Experimental Designs, with SAS*. Oxford University Press, Oxford.
- [2] Harman, R., Filová, L., Richtá, P. (2020) *A Randomized Exchange Algorithm for Computing Optimal Approximate Designs of Experiments*, Journal of the American Statistical Association 115, 348–361.
- [3] Harman, R., Trnovská, M. (2009) *Approximate D-optimal Designs of Experiments on the Convex Hull of a Finite Set of Information Matrices*, Mathematica Slovaca 59 (5), 693–704.

Jan Vávra

Shluková analýza longitudinálních dat smíšeného typu

MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8

vavraj@karlin.mff.cuni.cz

Pro longitudinální data vybudujeme regresní model sestávající se ze zobecněných lineárních smíšených modelů (GLMM) pro různé typy proměnných. Tyto zdánlivě nezávislé struktury propojíme sdruženým normálním rozdělením pro vektor všech náhodných efektů příslušících jednomu subjektu. Následováním metod shlukové analýzy na bázi modelu (model-based clustering) vytvoříme směs těchto modelů, kde jednotlivé skupiny se od sebe budou lišit v předem vybraných parametrech. Zavedením vhodného apriorního rozdělení jsme dokonce schopni se vypořádat s předem neznámým počtem skupin. Celý model včetně klasifikace subjektů je odhadován pomocí MCMC metod, které umožňují se elegantně vyhnout nepříjemným integracím latentních proměnných.

Petr Vejmělka

Přesnost predikcí rezerv v neživotním pojištění u vybraných metod

MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8

vejmelp@karlin.mff.cuni.cz

Jednou ze základních činností pojišťoven je výpočet a tvorba rezerv. V případě neživotního pojištění bylo v literatuře již navrženo mnoho přístupů a modelů, jak rezervy spočítat. Často je u navržených metod pro zvolená data a pro ně predikovanou rezervu spočítána nějaká míra kvality předpovědí rezerv, jako je např. střední čtvercová chyba predikce, či jsou s využitím bootstrapu odhadovány kvantily rozdělení dané rezervy. Velmi zřídka jsou však tyto metody podrobeny ověření, jak přesné jejich predikce ve srovnání se skutečnými hodnotami jsou.

Příspěvek se bude zabývat porovnáním vybraných metod sloužících pro výpočet rezerv v neživotním pojištění z hlediska přesnosti predikcí s využitím dat Garančního fondu České kanceláře pojistitelů.

Ondřej Vencálek

Statistika v době covidové

Přírodovědecká fakulta Univerzity Palackého, KMAAM, 17. listopadu 12 771 46 Olomouc

ondrej.vencalek@upol.cz

Epidemie onemocnění COVID-19 již více než dva roky výrazně ovlivňuje naše životy. Na jednu stranu se ukázala potřeba poznání založeného na datech a nezastupitelná role statistiky v procesu získávání nových poznatků, na druhou stranu se však také ukázalo, že se statistikou je stále ještě možné „lhát“ (manipulovat), viz Huff (1954), a opět se potvrdilo známé rčení, že statistika je nástroj podobný noži (dá se s ním ukrojit krajíc chleba, ovšem také bodnout) a že záleží na tom, kdo a jak s ním zachází, viz Hanousek a Charamza (1992).

Ukážu několik případů špatného použití statistiky, a to zejména v souvislosti s vyhodnocováním efektivity vakcín a efektu nefarmakologických opatření. Tyto chyby mě vedou k otevřené otázce, zda někdy není lepší nemít analýzy žádné, než mít analýzy, jejichž relevance je pochybná?

Literatúra

- [1] Huff, D. *How to lie with Statistics*. 1954.
- [2] Hanousek, J., Charamza, P. (1992) *Moderní metody zpracování dat: matematická statistika pro každého*. Praha: Grada.

Ján Veselý a Jozef Kováč

Náhodné dynamické systémy pre predpísané stacionárne rozdelenia

Fakulta matematiky, fyziky a informatiky, Univerzita Komenského, Mlynská dolina F1, 842 48 Bratislava

jan.vesely@fmph.uniba.sk

jozef.kovac@fmph.uniba.sk

Príspevok sa bude venovať náhodným dynamickým systémom generovaným konečným počtom funkcií, ktoré v každom čase volíme náhodne s predpísanou pravdepodobnosťou. Špeciálne sa budeme zaoberať tým, či existuje taký konečný počet funkcií, ktoré budú generovať náhodný dynamický systém s požadovaným jediným stacionárnym rozdelením. V článku [1] sa autor zaoberá tým ako nájsť náhodný dynamický systém k predpísanému spojitému stacionárному rozdeleniu. V diplomovej práci [2] sme sa venovali tomu, či vieme nájsť náhodný dynamický systém k predpísanému konečnému diskrétnemu stacionárному rozdeleniu. Pre stacionárne rozdelenie dĺžky n vieme triviálne nájsť n funkcií, ktoré budú generovať náhodný dynamický systém s požadovaným stacionárny rozdelením. Podarilo sa nám však dokázať, že existuje rozdelenie pravdepodobnosti dĺžky tri, ktoré nie je jediným stacionárnym rozdelením žiadneho náhodného dynamického systému, ktorý by bol generovaný najviac dvomi funkiami. Zovšeobecnenie tohto tvrdenia sa nám zatiaľ dokázať nepodarilo.

Literatúra

- [1] Stenflo, Ö. (2012) Iterated function systems with a given continuous stationary distribution, *Fractals* 20, 197–202
- [2] Veselý, J. (2021) Stacionárne rozdelenia náhodných dynamických systémov. Diplomová práca, Bratislava.

Petr Wolf**A model of random walk with transition probabilities depending on the walk history**

ÚTIA AV ČR, Pod Vodárenskou Věží 4, 18208 Praha 8

volf@utia.cas.cz

The contribution presents a model of discrete time Bernoulli-like random walk with transition probabilities depending on past walk steps. Moreover, it is assumed that these probabilities depend on covariates, which is described via the logistic regression model. One inspiration for such a model may be found in modeling certain sports matches development, for instance a tennis match and the sequence of its games, a random walk with steps ± 1 . A variant with steps 1, 0, can model a series of recurrent events, the “step” 1 denoting an event occurrence, “step” 0 then means no event in time interval t . It is assumed that both event occurrence and absence changes future event probability. Further, typically, in such cases the impact of covariates may change with time, therefore time-dependent parameters are considered. The model and its evaluation will be demonstrated on a real “benchmark” example of such recurrent events data.

Ondřej Vozár a Jaromír Antoch**Znáhodněné dotazování pro odhad populačního průměru kvantitativní proměnné**

KS a KE FS VŠE, Nám. W. Churchilla 4, 130 67 Praha 3 a MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8

vozo01@vse.cz, antoch@karlin.mff.cuni.cz

V přednášce představíme nový přístup k znáhodněnému dotazování pro odhad populačního úhrnu či průměru kvantitativní proměnné. Tato metoda poskytuje mnohem větší ochranu respondentova soukromí, jelikož nikdy ne-poskytne přesnou hodnotu důvěrného údaje na rozdíl od ostatních technik. Namísto poskytnutí hodnoty citlivého údaje respondent pouze odpoví, zda hodnota přesahuje jistou hodnotu, kterou si sám vygeneruje v náhodném pokusu a která není známa tazateli. Pro každého respondenta představuje jakousi individuální prahovou hodnotu, která je generována jako pseudonáhodné číslo. Též jsou představeny dvě modifikace této techniky. První modifikace předpokládá, že tazatel zná též pseudonáhodná čísla každého respondenta. Druhá modifikace je určena pro případ, kdy u některých proměnných jako výše příjmu, je citlivé udat jak příliš vysoké, tak příliš nízké hodnoty. V závislosti na tom, zda je respondentovo pseudonáhodné číslo vyšší než pevně stanovená prahová hodnota, je respondent dotazován tak, aby v závislosti na hodnotě citlivé proměnné se vyhnul nevhodné otázce. Navržené techniky jsou detailně studovány pro prostý náhodný výběr bez vracení, ale mohou být modifikovány též pro složitější výběrová schémata jako skupinkový výběr, dvoustupňový výběr anebo stratifikovaný výběr. Vlastnosti navržených technik jsou studovány pomocí simulací.

Poděkování: Výzkum je podpořen grantem GA ČR P403/22/1935S.

Gejza Wimmer* a Viktor Witkovský****Lineárny model v prípade nie normálne rozdelených meraní**

*MÚ SAV, Štefániková 49, 814 73 Bratislava

**ÚM SAV, Dúbravská cesta 9, 84104 Bratislava

wimmer@mat.savba.sk, witkovsky@savba.sk

Budeme uvažovať špeciálny regulárny lineárny model $\mathbf{Y}^* = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}^*$, pričom predpokladajme, že $\varepsilon_i^*, i = 1, 2, \dots, n$ je známa lineárna kombinácia nezávislých náhodných premenných $\varepsilon_j^{**}, j = 1, 2, \dots, m, m \geq n$ a rozdelenie každej $\varepsilon_j^{**}, j = 1, 2, \dots, m$ poznáme. Uvažovaný model patrí v metrológii medzi základné modely merania. Uvažujme lineárnu funkciu parametra $\boldsymbol{\beta}$, teda $\theta = \mathbf{d}'\boldsymbol{\beta}$. Jej NNLO označme $\widehat{\theta}$. Hustotu a distribučnú funkciu náhodnej veličiny $\widetilde{\theta} = \widehat{\theta} - \theta$ stanovíme metódou CFA (Characteristic Function Approach) ako numerickú inverziu charakteristickej funkcie náhodnej veličiny $\widetilde{\theta}$. Na určenie hľadanej hustoty a/alebo distribučnej funkcie $\widetilde{\theta}$ možno využiť softvér CharFunTool. Celý postup bude ilustrovaný na modeli priameho merania jednej veličiny.

Práca bola podporená projektami APVV-21-0216, VEGA 2/0096/21 a VEGA 2/0023/22.

Viktor Witkovský* and Gejza Wimmer**

A note on computing the exact distribution of the bootstrap mean based on the numerical inversion of the empirical characteristic function

*Institute of Measurement Science, Slovak Academy of Sciences, Bratislava

**Mathematical Institute, Slovak Academy of Sciences, Bratislava

witkovsky@savba.sk, wimmer@mat.savba.sk

This lecture deals with numerical methods for computing exact probability distributions by numerical inversion of the characteristic function. We briefly introduce the Characteristics Functions Toolbox (CharFunTool) - the MATLAB algorithms and methods for evaluating characteristic functions and tools for their combinations and numerical inversion, [4]. For illustrative purposes, we focus here on possible applications based on the use of the empirical characteristic functions. The method is particularly suitable for the fast calculation of the bootstrap distribution of the sample mean as well as for other linear functions, e.g. the sample moments. For standard results based on the limit theorems for the bootstrap mean, see e.g. [1, 2]. Note that the proposed method is exact when applied to lattice distributions (i.e. distributions where every possible value can be represented in the form $a + bn$, where $b \neq 0$ and n is an integer). For more details on suitable fast numerical methods, see [3]. Other possibilities and applications are open for discussion.

Literatúra

- [1] Yang, S.S. (1988) A central limit theorem for the bootstrap mean. *The American Statistician*, 42(3), 202–203.
- [2] Angus, J.E. (1989) A note on the central limit theorem for the bootstrap mean. *Communications in Statistics- Theory and Methods* 18(5), 1979–1982.
- [3] Warr, R.L. (2014) Numerical approximation of probability mass functions via the inverse discrete Fourier transform. *Methodology and Computing in Applied Probability*, 16(4), 1025–1038.
- [4] Witkovský, V. (2022) CharFunTool: The characteristic functions toolbox (MATLAB). Available at GitHub: <https://github.com/witkovsky/CharFunTool>.

Acknowledgement: The work was supported by the Slovak Research and Development Agency, projects APVV-18-0066 and APVV-21-0216, and VEGA 2/0096/21 and VEGA 2/0023/22.

Markéta Zoubková

Modeling Bitcoin price using the attention factor

Ústav matematiky a statistiky, Přírodovědecká fakulta, Masarykova univerzita, Kotlářská 2, 611 37 Brno
460769@mail.muni.cz

Cryptocurrencies are the new phenomenon in the world of finance over the last decade. In this paper we look into the way of using the attention factor to model price of the first cryptocurrency – Bitcoin. We focus on modeling through stochastic processes and stochastic differential equations.

Ivan Žežula a Daniel Klein

Vícenásobné testování středních hodnot u vícerozměrných dat s dvojitou blokově rovnoměrnou varianční strukturou

PF UPJŠ, Jesenná 5, 040 01 Košice, SR

ivan.zezula@upjs.sk, daniel.klein@upjs.sk

U mnohorozměrných dat bývá problém, že počet parametrů modelu může značně převyšovat počet pozorování. To pak brání možnosti rozumné statistické inference. Proto se v poslední době těší velké pozornosti modely se speciálními variančními strukturami, u kterých je počet parametrů druhého rádu podstatně redukován. Jednou z těchto struktur, která může přirozeně vyplývat z povahy věci nebo plánu experimentu, je blokově rovnoměrná varianční struktura. V našem příspěvku se budeme zabývat modelem MANOVA pro takováto data, vycházející z modifikovaného Hotellingova testu.

Poděkování: Tato práce byla podpořena Agentúrou na podporu výskumu a vývoja na základě smlouvy č. APVV-17-0568.