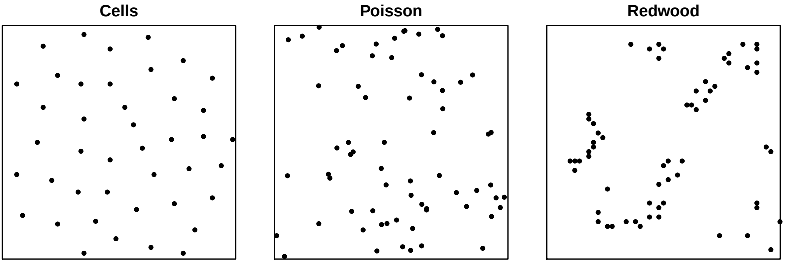
Časoprostorové shot-noise Coxovy bodové procesy – odhady parametrů a jejich asymptotické vlastnosti

Jiří Dvořák

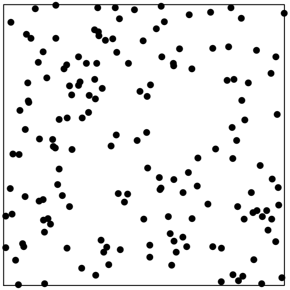
Katedra pravděpodobnosti a matematické statistiky, MFF UK, Praha,
Ústav teorie informace a automatizace AV ČR, v.v.i., Praha

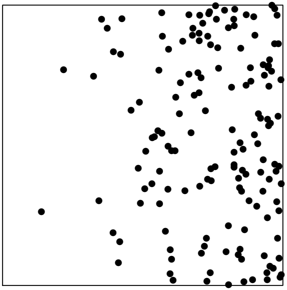
ROBUST 2014
Jetřichovice, 20. 1. 2014

Bodové procesy – regularita vs. shlukování

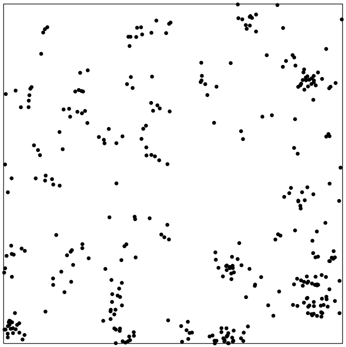


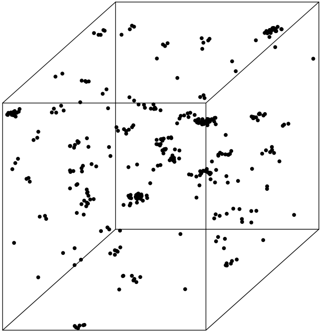
Bodové procesy – stacionární vs. nestacionární





Bodové procesy – prostorové vs. časoprostorové

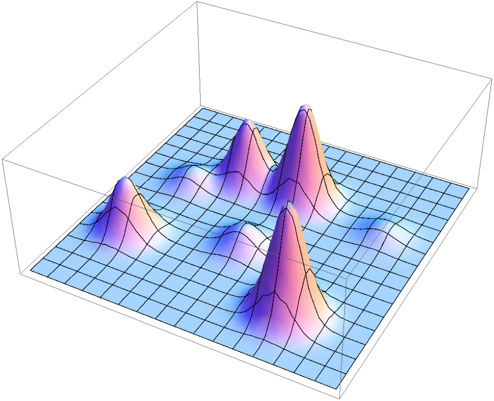




Stacionární shot-noise Coxův b.p.

$$\Lambda(u,t) = \sum_{(r,v,s)\in\Phi} r\,k(v-u, s-t), \quad (u,t) \in \mathbb{R}^2 \times \mathbb{R}$$

Φ je poissonovská míra na $\mathbb{R}^+ \times \mathbb{R}^2 \times \mathbb{R}$, k vyhlazovací jádro



Netacionární shot-noise Coxův b.p.

- nehomogenita nezávislým ztenčením stacionárního procesu
- $0 \leq f \leq 1 \ldots$ „funkce nehomogenity“
- každý bod (u, t) stacionárního procesu zachován s pstí $f(u, t)$, nezávisle na ostatních bodech
- f může záviset na vektoru kovariát $z(u, t)$

Značení a parametrizace modelu

- X … ztenčený (nestacionární) časoprostorový proces
- λ … jeho funkce intenzity (1. řádu)
- $W \times T$ … pozorovací okno, $|W| > 0, |T| > 0$
- k … pravděpodobnostní jádro
- $\max_{W \times T} f = 1$

Značení a parametrizace modelu

- X ... ztenčený (nestacionární) časoprostorový proces
- λ ... jeho funkce intenzity (1. řádu)
- $W \times T$ ... pozorovací okno, $|W| > 0, |T| > 0$
- k ... pravděpodobnostní jádro
- $\max_{W \times T} f = 1$

- β ... (vektorový) parametr funkce nehomogenity $f(\cdot; \beta)$
- ψ ... (vektorový) parametr vyhlazovacího jádra $k(\cdot, \cdot; \psi)$
- θ ... (vektorový) parametr poissonovské báze Φ

Vícekrokový odhad parametrů

1. odhad parametrů nehomogenity β maximalizací poissonovské věrohodnostní funkce

Vícekrokový odhad parametrů

1. odhad parametrů nehomogenity β maximalizací poissonovské věrohodnostní funkce
2. odhad interakčních parametrů ψ pomocí minimálního kontrastu na časoprostorovou K-funkci

Vícekrokový odhad parametrů

1. odhad parametrů nehomogenity β maximalizací poissonovské věrohodnostní funkce
2. odhad interakčních parametrů ψ pomocí minimálního kontrastu na časoprostorovou K-funkci
3. odhad θ dopočítáním z předchozích odhadů a celkové intenzity procesu

Druhý krok – metoda minimálního kontrastu

$$\widehat{\psi} = \arg\min \int_{r_{min}}^{r_{max}} \int_{t_{min}}^{t_{max}} (\widehat{K}(r,t)^q - K(r,t;\psi)^q)^2 \,\mathrm{d}t\,\mathrm{d}r$$

$$\widehat{K}(r,t) =$$

$$\frac{1}{|W \times T|} \sum_{(u_i,t_i),(u_j,t_j)\in W\times T}^{\neq} \frac{I(\|u_i - u_j\| \leq r, |t_i - t_j| \leq t)}{w_1(u_i,u_j)w_2(t_i,t_j)\widehat{\lambda}(u_i,t_i;\widehat{\beta})\widehat{\lambda}(u_j,t_j;\widehat{\beta})}$$

q ... exponent stabilizující rozptyl, typicky $q = 1/2$ nebo $1/4$

Druhý krok – metoda minimálního kontrastu

$$\widehat{\psi} = \arg\min \int_{r_{min}}^{r_{max}} \int_{t_{min}}^{t_{max}} (\widehat{K}(r,t)^q - K(r,t;\psi)^q)^2 \,\mathrm{d}t\,\mathrm{d}r$$

$$\widehat{K}(r,t) =$$

$$\frac{1}{|W \times T|} \sum_{(u_i,t_i),(u_j,t_j)\in W\times T}^{\neq} \frac{I(\|u_i - u_j\| \leq r, |t_i - t_j| \leq t)}{w_1(u_i,u_j) w_2(t_i,t_j) \widehat{\lambda}(u_i,t_i;\widehat{\beta}) \widehat{\lambda}(u_j,t_j;\widehat{\beta})}$$

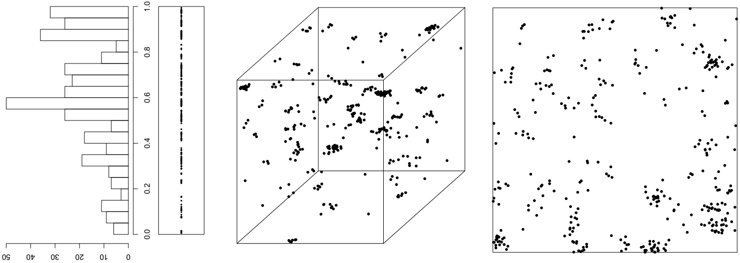
q . . . exponent stabilizující rozptyl, typicky $q = 1/2$ nebo $1/4$

Problémy: stabilita odhadu $\widehat{K}$, náročná optimalizační úloha

Možné řešení – snížení dimenzionality

Prostorová projekce: $X_{space} = \{u : (u,t) \in X, t \in T\}$

Časová projekce: $X_{time} = \{t : (u,t) \in X, u \in W\}$



Dodatečné předpoklady – separabilita

- $f(u, t; \beta) = f_1(u; \beta) f_2(t; \beta), u \in W, t \in T$
- $k(w, \tau; \psi) = k_1(w; \psi_1) k_2(\tau; \psi_2), w \in \mathbb{R}^2, \tau \in \mathbb{R}$

Pak můžeme zavést K_{space}, K_{time} a použít je k odhadu.

Dodatečné předpoklady – separabilita

- $f(u, t; \beta) = f_1(u; \beta) f_2(t; \beta), u \in W, t \in T$
- $k(w, \tau; \psi) = k_1(w; \psi_1) k_2(\tau; \psi_2), w \in \mathbb{R}^2, \tau \in \mathbb{R}$

Pak můžeme zavést K_{space}, K_{time} a použít je k odhadu.

Problémy: překrývání shluků, těžké odhadnout časový interakční parametr ψ_2

Vylepšená metoda

Cíl: vyhnout se projekci do časové domény.

Vylepšená metoda

Cíl: vyhnout se projekci do časové domény.

Místo odhadu ψ_2 z K_{time}:

- zvolit prostorový dosah $R > 0$
- odhadnout $\widehat{K}(R, t)$ jako funkci t
- použít k odhadu ψ_2 min. kontrast na $K(R, \cdot)$

Vylepšená metoda

Cíl: vyhnout se projekci do časové domény.

Místo odhadu ψ_2 z K_{time}:

- zvolit prostorový dosah $R > 0$
- odhadnout $\widehat{K}(R, t)$ jako funkci t
- použít k odhadu ψ_2 min. kontrast na $K(R, \cdot)$

Odhady ostatních parametrů jako dříve.

Vylepšená metoda

Cíl: vyhnout se projekci do časové domény.

Místo odhadu ψ_2 z K_{time}:

- zvolit prostorový dosah $R > 0$
- odhadnout $\widehat{K}(R, t)$ jako funkci t
- použít k odhadu ψ_2 min. kontrast na $K(R, \cdot)$

Odhady ostatních parametrů jako dříve.

Výhoda: nedochází k překrývání shluků, zlepšení odhadů ψ_2, θ.

Vylepšená metoda – reformulace a asymptotika

Uvažme posloupnost kompaktních pozorovacích oken $W_n \times T_n \nearrow \mathbb{R}^2 \times \mathbb{R}$. Zaměříme se na odhady β, ψ_1, ψ_2.

Vylepšená metoda – reformulace a asymptotika

Uvažme posloupnost kompaktních pozorovacích oken $W_n \times T_n \nearrow \mathbb{R}^2 \times \mathbb{R}$. Zaměříme se na odhady β, ψ_1, ψ_2.

$(\widehat{\beta}_n, \widehat{\psi}_{1,n}, \widehat{\psi}_{2,n})$ jako řešení odhadovací rovnice

$$U_n(\beta, \psi_1, \psi_2) = (U_{n,1}(\beta), U_{n,2}(\beta, \psi_1), U_{n,3}(\beta, \psi_2) = 0,$$

Vylepšená metoda – reformulace a asymptotika

Uvažme posloupnost kompaktních pozorovacích oken $W_n \times T_n \nearrow \mathbb{R}^2 \times \mathbb{R}$. Zaměříme se na odhady β, ψ_1, ψ_2.

$(\widehat{\beta}_n, \widehat{\psi}_{1,n}, \widehat{\psi}_{2,n})$ jako řešení odhadovací rovnice

$$U_n(\beta, \psi_1, \psi_2) = (U_{n,1}(\beta), U_{n,2}(\beta, \psi_1), U_{n,3}(\beta, \psi_2) = 0,$$

kde

- $U_{n,1}(\beta)$ je poissonovská skórová funkce, závisí na $X \cap (W_n \times T_n)$

Vylepšená metoda – reformulace a asymptotika

Uvažme posloupnost kompaktních pozorovacích oken $W_n \times T_n \nearrow \mathbb{R}^2 \times \mathbb{R}$. Zaměříme se na odhady β, ψ_1, ψ_2.

$(\widehat{\beta}_n, \widehat{\psi}_{1,n}, \widehat{\psi}_{2,n})$ jako řešení odhadovací rovnice

$$U_n(\beta, \psi_1, \psi_2) = (U_{n,1}(\beta), U_{n,2}(\beta, \psi_1), U_{n,3}(\beta, \psi_2) = 0,$$

kde

- $U_{n,1}(\beta)$ je poissonovská skórová funkce, závisí na $X \cap (W_n \times T_n)$
- $U_{n,2}(\beta, \psi_1)$ je derivace kritéria min. kontrastu podle ψ_1, závisí na $X_{space} \cap W_n$, tedy $X \cap (W_n \times T)$

Vylepšená metoda – reformulace a asymptotika

Uvažme posloupnost kompaktních pozorovacích oken $W_n \times T_n \nearrow \mathbb{R}^2 \times \mathbb{R}$. Zaměříme se na odhady β, ψ_1, ψ_2.

$(\widehat{\beta}_n, \widehat{\psi}_{1,n}, \widehat{\psi}_{2,n})$ jako řešení odhadovací rovnice

$$U_n(\beta, \psi_1, \psi_2) = (U_{n,1}(\beta), U_{n,2}(\beta, \psi_1), U_{n,3}(\beta, \psi_2) = 0,$$

kde

- $U_{n,1}(\beta)$ je poissonovská skórová funkce, závisí na $X \cap (W_n \times T_n)$
- $U_{n,2}(\beta, \psi_1)$ je derivace kritéria min. kontrastu podle ψ_1, závisí na $X_{space} \cap W_n$, tedy $X \cap (W_n \times T)$
- $U_{n,3}(\beta, \psi_2)$ je derivace kritéria min. kontrastu podle ψ_2, závisí na $X \cap (W_n \times T_n)$

Asymptotika – poznámky

- pokud pro definici $X_{space,n}$ použijeme T_n, pomocí charakteristik 2. řádu nebude $X_{space,n}$ rozeznatelný od Poissonova procesu $\Rightarrow$ nutno používat X_{space} s pevným T

Asymptotika – poznámky

- pokud pro definici $X_{space,n}$ použijeme T_n, pomocí charakteristik 2. řádu nebude $X_{space,n}$ rozeznatelný od Poissonova procesu $\Rightarrow$ nutno používat X_{space} s pevným T
- množství dat neroste pro všechna $U_{n,i}$ stejně $\Rightarrow$ potřeba normovat pomocí $|W_n \times T_n|$, resp. $|W_n|$

Asymptotika – poznámky

- pokud pro definici $X_{space,n}$ použijeme T_n, pomocí charakteristik 2. řádu nebude $X_{space,n}$ rozeznatelný od Poissonova procesu $\Rightarrow$ nutno používat X_{space} s pevným T
- množství dat neroste pro všechna $U_{n,i}$ stejně $\Rightarrow$ potřeba normovat pomocí $|W_n \times T_n|$, resp. $|W_n|$
- komplikace při formulaci as. vlastností – buď degenerované rozptylové matice nebo (ne)stejnoměrná integrovatelnost

Příklad – konzistenční věta

Při splnění patřičných předpokladů existuje posloupnost $\left\{(\widehat{\beta}_n, \widehat{\psi}_{1,n}, \widehat{\psi}_{2,n})\right\}_{n\geq 1}$ tak, že

$$U_n(\widehat{\beta}_n, \widehat{\psi}_{1,n}, \widehat{\psi}_{2,n}) = 0$$

s pstí jdoucí k 1 a vektor

$$L_n = \begin{pmatrix} |W_n \times T_n|^{1/2}(\widehat{\beta}_n - \beta^*) \\ |W_n|^{1/2}(\widehat{\psi}_{1,n} - \psi_1^*) \\ |W_n \times T_n|^{1/2}(\widehat{\psi}_{2,n} - \psi_2^*) \end{pmatrix}$$

je omezený v psti, tedy

$\forall \varepsilon > 0 \; \exists \delta > 0 : \mathbb{P}(\|L_n\| > \delta) \leq \epsilon$ pro dostatečně velké n.

Reference

1. Møller, J., Ghorbani, M. (2012). *Aspects of second-order analysis of structured inhomogeneous spatio-temporal point processes.* Stat. Neerlandica **66**, 472 – 491.
2. Prokešová, M., Dvořák, J. (2014). *Statistics for inhomogeneous space-time shot-noise Cox processes.* Vyjde v Methodology and Computing in Applied Probability.
3. Waagepetersen, R. P. (2007). *An estimating function approach to inference for inhomogeneous Neyman-Scott processes.* Biometrics **63**, 252 – 258.
4. Waagepetersen, R. P., Guan, Y. (2009). *Two-step estimation for inhomogeneous spatial point processes.* J. R. Stat. Soc. B **71**, 685 – 702.