
NMFM301 – Statistika pro finanční matematiky

Pořádkové statistiky. Nestrannost a konsistence odhadů

Teoretické cvičenie #2 | 10.10.2022 & 12.10.2022

Pro náhodný výběr $X_1, \dots, X_n$ z nějakého rozdělení F definujeme uspořádaný výběr $X_{(1)}, \dots, X_{(n)}$, kde $X_{(1)}$ je tzv. první pořadová statistika náhodného výběru (t.j. minimum $X_1, \dots, X_n$) a $X_{(n)}$ je n -tá pořadová statistika náhodného výběru (t.j. maximum $X_1, \dots, X_n$).

V případě, že $n = 2k + 1$ pro nějaké $k \in \mathbb{N}$, pak definujeme výběrový medián náhodného výběru $X_1, \dots, X_n$ jako $X_m = X_{(k+1)}$. V případě, že rozsah náhodného výběru je sudý (t.j. $n = 2k$ pro nějaké $k \in \mathbb{N}$), pak je výběrový medián definován jako $X_m = (X_{(k)} + X_{(k+1)})/2$.

A Příklady na cvičení

Nechť $X_1, \dots, X_n$ je náhodný výběr ze spojitého rozdělení s distribuční funkcí F a hustotou f .

A1. [Instruktážní] Nechť $n = 2k + 1$.

- Najděte hustotu prostředního pozorování, t.j. výběrového mediánu $X_{(k+1)}$.
- Nechť X_i má rovnoměrné rozdělení na intervalu $(0, 1)$. Spočítejte $E X_{(k+1)}$ a $\text{var } X_{(k+1)}$.
- Nechť $X_i \sim R(0, \theta)$. Je $X_{(k+1)}$ nestranným a/nebo konsistentním odhadem mediánu rozdělení $R(0, \theta)$? [Použijte tvrzení P.7.5]

A2. [Instruktážní] Nechť X_i má exponenciální rozdělení s parametrem 1.

- Definujte

$$Z_1 = nX_{(1)}, \quad Z_k = (n - k + 1)(X_{(k)} - X_{(k-1)}), \quad k = 2, \dots, n.$$

Ukažte, že $Z_1, \dots, Z_n$ jsou nezávislé stejně rozdělené náhodné veličiny s rozdělením $\text{Exp}(1)$.

- Vyjádřete $X_{(r)}$ pomocí lineární kombinace veličin $Z_1, \dots, Z_n$ a pomocí tohoto vztahu spočítejte $E X_{(r)}$ a $\text{var } X_{(r)}$ (pro libovolné $r = 1, \dots, n$).
- Nechť $X_i \sim \text{Exp}(\lambda)$ a $n = 2k + 1$. Je $X_{(k+1)}$ nestranným a/nebo konsistentním odhadem mediánu rozdělení $\text{Exp}(\lambda)$?

A3. [Procvičovací] Nechť X_i má rozdělení $R(\theta_1, \theta_2)$. Najděte nestranné odhady parametrů θ_1 a θ_2 založené na maximu $X_{(n)}$ a minimu $X_{(1)}$.

A4. [Procvičovací] Pro náhodný výběr $X_1, \dots, X_n$ z Poissonova rozdělení s parametrem $\lambda > 0$ ukažte, že $\hat{p}_0 = (1 - \frac{1}{n})^{n\bar{X}_n}$ je nestranným a konzistentním odhadem $p_0 = \mathbb{P}[X_1 = 0]$.

B Doplnující příklady (nahrazování, samostatné procvičování)

Z těchto příkladů je potřebné samostatně spočítat aspoň dva příklady a řešení zaslat emailom na adresu `maciak[AT]karlin.mff.cuni.cz`, případně doručit v papírové verzi—v oboch případech nejpozději před začátkem třetího cvičení.

B1. Uvažujme náhodný výběr $X_1, \dots, X_n$ z rozdělení s hustotou

$$f(x; \theta) = \theta x^{\theta-1} 1_{[0,1]}(x), \quad x \in \mathbb{R};$$

kde $\theta > 0$ je neznámý parametr.

- (a) Skonstruujte přesný interval spolehlivosti pro parametr θ s hladinou spolehlivosti $1 - \alpha$.
- (b) Skonstruujte přibližný $100(1 - \alpha)\%$ interval spolehlivosti pro parametr θ pomocí CLV.

[Návod: Uvažujte transformaci $Y_i = -\log X_i$]

B2. Nechť $X_1, \dots, X_n$ je náhodný výběr z rozdělení $R(0, \theta)$. Zjistěte, zdali $X_{(n)}$ je nestranným a/nebo konsistentním odhadem parametru θ .

B3. Uvažujme náhodný výběr $X_1, \dots, X_n$ s hustotou

$$f(x) = 3\theta^{-3}x^2 1_{(0,\theta)}(x), \quad x \in \mathbb{R}, \theta > 0.$$

- (a) Ověřte, že $\hat{\theta}_n = \frac{4}{3}\bar{X}_n$ je nestranný odhad parametru θ .
- (b) Ověřte, že $\tilde{\theta}_n = \frac{3n+1}{3n}X_{(n)}$ je nestranný odhad parametru θ .
- (c) [Pro nahrazování nepovinný] Najděte rozptyl $\hat{\theta}_n$ a $\tilde{\theta}_n$ a porovnejte rychlost konvergence rozptylů k 0 při $n \rightarrow \infty$.

B4. Nechť X_i má rozdělení $\text{Alt}(p)$. Najděte nestranný odhad parametru $\theta = p(1 - p)$ založený na $\bar{X}_n$.

B5. Nechť X_i má rozdělení $\text{Exp}(\lambda)$. Ukažte, že

$$\hat{\theta}_n = 1 - \left(1 - \frac{u}{n\bar{X}_n}\right)^{n-1}$$

je nestranným odhadem parametru $\theta = 1 - e^{-\lambda u} = F_X(u)$.

B6. Uvažujte nezávislé náhodné veličiny $X_1, X_2, \dots$ s rozdělením $\text{Alt}(p)$.

- (a) Ukažte (sporem), že při pevném rozsahu výběru n neexistuje nestranný odhad parametru $\theta = 1/p$ založený na $X_1, \dots, X_n$.
- (b) Nechť Z značí počet nul předcházejících první jedničce v posloupnosti $X_1, X_2, \dots$. Víme, že Z má rozdělení $\text{Geo}(p)$. Ukažte, že $Z + 1$ je nestranným odhadem parametru $\theta = 1/p$.

B7. Nechť $X_1, \dots, X_n$ je náhodný výběr z rozdělení $\text{Exp}(1)$ a $U_1, \dots, U_n$ je náhodný výběr z rozdělení $R(0, 1)$. Ukažte, že $-\log U_{(n-r+1)}$ má stejné rozdělení jako $X_{(r)}$. Pomocí příkladu **A2** ukažte, že

$$Q_r = \left(\frac{U_{(r)}}{U_{(r+1)}}\right)^r$$

jsou nezávislé stejně rozdělené náhodné veličiny s rozdělením $R(0, 1)$.

B8. Máme-li dva nezávislé náhodné výběry $X_1, \dots, X_n \sim \text{Exp}(\lambda)$ a $Y_1, \dots, Y_m \sim \text{Exp}(\nu)$, odvozte konfidenční interval pro parametr $\varrho = \lambda/\nu$. [Návod: Uvědomte si vztahy mezi exponenciálním, Gamma a χ^2 -rozdělením.]

B9. Necht náhodný vektor $\mathbf{X} = [X_1, X_2, X_3, X_4]^\top$ má multinomické rozdělení s parametry $n \in \mathbb{N}$ a $\mathbf{p} = [p_1, p_2, p_3, p_4]^\top \in (0, 1)^4$.

- (a) Odvoďte asymptotické rozdělení vektoru $\mathbf{X}/n$ vzhledem k rostoucímu n na základě CLV.
- (b) Ukažte asymptotickou normalitu odhadu

$$\hat{\theta} = \frac{\frac{X_1}{n} \frac{X_4}{n}}{\frac{X_2}{n} \frac{X_3}{n}}$$

parametru $\theta = \frac{p_1 p_4}{p_2 p_3}$ pomocí Delta metody.

- (c) Použijte Delta metodu na odhad parametru $\vartheta = \log \theta$.
- (d) Na základě asymptotické normality odhadu parametru ϑ konstruujte přibližný $100(1 - \alpha)\%$ interval spolehlivosti pro parametr θ .
- (e) Yuleovo $Q = \frac{\theta - 1}{\theta + 1} = \frac{p_1 p_4 - p_2 p_3}{p_1 p_4 + p_2 p_3}$ odhadujeme pomocí $\hat{Q} = \frac{\hat{\theta} - 1}{\hat{\theta} + 1}$. Ukažte, že $1 - \hat{Q}^2 = \frac{4\hat{\theta}}{(\hat{\theta} + 1)^2}$.
- (f) Skonstruujte přibližný $100(1 - \alpha)\%$ interval spolehlivosti pro Yuleovo Q .