NMFM301 – Statistika pro finanční matematiky

Pořádkové statistiky. Nestrannost a konsistence odhadů

Podrobné riešenie príkladov (A) a výsledky (B) z 2. cvičenia

A Vzorové príklady s podrobným řešením

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

A1. Nechť $n = 2k+1$.

- Najděte hustotu prostředního pozorování $X_{(k+1)}$. [Tato statistika se nazývá výběrový medián.]

  Řešení:
  Obecně (podľa Vety 1.9 z prednášky) platí, že hustota k-tej pořadovej statistiky vzhľadom k Lebesgueovej miere je

  $$f_{(k)}(x) = n\binom{n-1}{k-1} f(x)F^{k-1}(x)[1-F(x)]^{n-k}, \qquad \text{pre } x \in \mathbb{R}.$$

  Pre nepárny (lichý) rozsah náhodného výberu (keďže $n = 2k+1$) je výberový médian $(k+1)$ poradová štatistika s príslušnou hustotou

  $$f_{(k+1)}(x) = n\binom{2k}{k} f(x)F^{k}(x)[1-F(x)]^{k},$$

  opäť pre všetky $x \in \mathbb{R}$.

- Nechť X_i má rovnoměrné rozdělení na intervalu $(0,1)$. Spočítejte $\mathsf{E}\, X_{(k+1)}$ a $\mathsf{var}\, X_{(k+1)}$.

  Řešení:
  Keďže náhodné veličiny $X_1, \dots, X_n$ pochádzajú z rovnomerného rozdelenia na intervale $(0,1)$, t.j., $X_i \sim R(0,1)$, tak hustota výberového mediánu, t.j., náhodnej veličiny $X_{(k+1)}$ je

  $$f_{(k+1)} = n\binom{2k}{k} x^k(1-x)^k, \qquad x \in (0,1)$$

  a $f_{(k+1)}(x) = 0$ pre $x \notin (0,1)$. To je hustota náhodnej veličiny s Beta rozdelením s parametrami $\alpha = k+1$ a $\beta = k+1$. To znamená, že výberový medián náhodných veličín $X_1, \dots, X_n$ z rovnomerného rozdelenia na intervale $(0,1)$ (pre $n = 2k+1$) má Beta rozdelenie, t.j., $X_{(k+1)} \sim Beta(k+1, k+1)$. Pre strednú hodnotu a rozptyl zároveň platí, že

  $$\begin{aligned} EX_{(k+1)} &= \frac{\alpha}{\alpha+\beta} = \frac{k+1}{2(k+1)} = \frac{1}{2}, \\ \mathsf{var}\, X_{(k+1)} &= \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} = \frac{(k+1)^2}{4(\alpha+\beta)^2(2k+3)} = \frac{1}{4(2k+3)} = \frac{1}{4(n+2)}. \end{aligned}$$

- Nechť $X_i \sim \mathsf{R}(0,\theta)$. Je $X_{(k+1)}$ nestranným a/nebo konsistentním odhadem mediánu rozdělení $\mathsf{R}(0,\theta)$? [Použijte tvrzení P.7.5]

  Řešení:
  V prvom rade je nutné si uvedomiť jednoduchý fakt, že ak $X_i \sim R(0,1)$, tak potom platí, že $\theta X_i \sim R(0,\theta)$, pro libovolné $\theta > 0$. Pokud teda $X_i \sim R(0,\theta)$, pre nejaké $\theta > 0$, tak potom

  $$EX_{(k+1)} = \frac{\theta}{2} \qquad \text{a tiež} \qquad \mathsf{var}\, X_{(k+1)} = \frac{\theta^2}{4(n+2)}.$$

  Keďže medián náhodnej veličiny s rovnomerným rozdelením na intervale $(0,\theta) \subset \mathbb{R}$ je $\theta/2$, tak náhodná veličina $X_{(k+1)}$ je nestranným odhadom teoretického mediánu. Keďže naviac platí aj

  $$\mathsf{var}\, X_{(k+1)} = \frac{\theta^2}{4(n+2)} \to 0, \qquad \text{pre } n \to \infty,$$

  tak $X_{(k+1)}$ je zároveň aj konzistentným odhadom teoretického mediánu, t.j., parametru $\widetilde{\theta} = \frac{\theta}{2}$.

A2. Nechť X_i má exponenciální rozdělení s parametrem 1.

- Definujte
  $$Z_1 = nX_{(1)}, \quad Z_k = (n-k+1)(X_{(k)} - X_{(k-1)}), \quad k = 2,\ldots,n.$$
  Ukažte, že $Z_1,\ldots,Z_n$ jsou nezávislé stejně rozdělené náhodné veličiny s rozdělením $\mathsf{Exp}(1)$.

  Řešení:
  Z prednášky už vieme, že združená hustota náhodného vektoru $(X_{(1)},\ldots,X_{(n)})^\top$ je

  $$f(y_1,\ldots,y_n) = n!f(y_1)\cdot\cdots\cdot f(y_n), \qquad \text{pre } y_1 < \cdots < y_n$$

  a $f(y_1,\ldots,y_n) = 0$ inak (Veta 1.7 z prednášky). Keďže náhodné veličiny $X_1,\ldots,X_n$ majú exponenciálne rozdelenie s parametrom $\lambda = 1$, tak zároveň vieme, že pre hustotu X_i platí

  $$f(x) = e^{-x} \cdot \mathbb{I}_{\{x>0\}}$$

  a $EX_i = \mathsf{var}\, X_i = 1$. Pre združenú hustotu $f(y_1,\ldots,y_n)$ preto platí

  $$f(y_1,\ldots,y_n) = n!\mathrm{e}^{-\sum_{i=1}^n y_i}, \qquad \text{pre } y_1 < \cdots < y_n$$

  a $f(y_1,\ldots,y_n) = 0$ inak. Následne súčet v exponente zapíšeme iným spôsobom:

  $$\sum_{i=1}^n y_i = \sum_{i=1}^n (n-i+1)(y_i - y_{i-1}),$$

  kde jsme dodefinovali počiatočnú hodnotu $y_0 = 0$.

Združenú hustotu teda môžeme teda zapísať ako

$$f(y_1, \dots, y_n) = n! e^{-\sum_{i=1}^n (n-i+1)(y_i - y_{i-1})}, \qquad \text{opäť pre } y_1 < \dots < y_n$$

a $f(y_1, \dots, y_n) = 0$ inak. Použijeme transformáciu $t : (y_1, \dots, y_n) \to (z_1, \dots, z_n)$ definovanú (podľa zadania) následovne:

$$\begin{aligned}
z_1 &= n y_1 \\
z_2 &= (n-1)(y_2 - y_1) \\
&\dots\dots \\
z_i &= (n-i+1)(y_i - y_{i-1}) \\
&\dots\dots \\
z_n &= 1 \cdot (y_n - y_{n-1}).
\end{aligned}$$

Keďže $y_1 < \dots < y_n$ (resp. stačí nam riešiť tento prípad), tak zároveň dostávame, že $z_i > 0$ pre $\forall i \in \{1, \dots, n\}$. Pre inverznú transformáciu $t^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ platí:

$$\begin{aligned}
y_1 &= \frac{1}{n} z_i \\
y_2 &= \frac{z_2}{n-1} + y_1 = \sum_{j=1}^{2} \frac{z_j}{n-j+1} \\
&\dots\dots \\
y_i &= \frac{z_i}{n-i+1} + y_{i-1} = \sum_{j=1}^{i} \frac{z_j}{n-j+1} \qquad (\star) \\
&\dots\dots \\
y_n &= \sum_{j=1}^{n} \frac{z_j}{n-j+1}.
\end{aligned}$$

Pre Jakobián tohto inverzného zorazenia platí, že

$$|J_{t^{-1}}(z_1, \dots, z_n)| = \left| \frac{\partial t^{-1}}{\partial \boldsymbol{z}}(\boldsymbol{z}) \right| = \begin{vmatrix} \frac{1}{n} & 0 & \dots & \dots & 0 \\ \frac{1}{n} & \frac{1}{n-1} & 0 & \dots & 0 \\ \vdots & \vdots & & & \vdots \\ \frac{1}{n} & \frac{1}{n-1} & \frac{1}{n-2} & \dots & 1 \end{vmatrix} = \frac{1}{n!},$$

kde jsme využili značenie $\boldsymbol{z} = (z_1, \dots, z_n)^\top$.
Združené rozdelenie náhodného vektoru $(Z_1, \dots, Z_n)^\top$ (definovaného transformáciou t z náhodného vektoru $(X_{(1)}, \dots, X_{(n)})^\top$) získame pomocou vety o transformácii

$$\begin{aligned}
f_{\boldsymbol{Z}}(z_1, \dots, z_n) &= f_{\boldsymbol{X}_{(\cdot)}}(t^{-1}(z_1, \dots, z_n)) \cdot \left| \frac{\partial t^{-1}}{\partial \boldsymbol{z}}(\boldsymbol{z}) \right| \\
&= e^{-\sum_{i=1}^n z_i}, \qquad \text{pre } z_i > 0,
\end{aligned}$$

kde $f_{\boldsymbol{Z}}$ značí združenú hustotu náhodného vektoru $(Z_1, \dots, Z_n)^\top$ a $f_{\boldsymbol{X}_{(\cdot)}}$ je združená hustota usporiadaného náhodného vektoru $(X_{(1)}, \dots, X_{(n)})^\top$. Zároveň platí, že hustota je nulová pre ľubovolné $z_i \leq 0$. Keďže naviac je možné hustotu faktorizovať v zmysle

$$f_{\boldsymbol{Z}}(z_1, \dots, z_n) = f_{Z_1}(z_1) \cdots f_{Z_n}(z_n) = e^{-z_1} \cdots e^{-z_n},$$

pre všetky $(x_1, \dots, x_n) \in \mathbb{R}^n$, tak náhodné veličiny $Z_1, \dots, Z_n$ sú nezávislé.

- Vyjádřete $X_{(r)}$ pomocí lineární kombinace veličin $Z_1, \dots, Z_n$ a pomocí tohoto vztahu spočítejte $\mathsf{E}\, X_{(r)}$ a $\mathsf{var}\, X_{(r)}$ (pro libovolné $r = 1, \dots, n$).

  Řešení:
  Pro ľubovolné $r \in \{1, \dots, n\}$ už vieme z použitej transformácie – konkrétne z výrazu v ($\star$), že platí
  $$X_{(r)} = g(Z_1, \dots, Z_n) = \sum_{j=1}^{r} \frac{Z_j}{n-j+1}.$$
  Keďže platí, že $Z_1, \dots, Z_n \sim Exp(1)$ tak pre strednú hodnotu $X_{(r)}$ platí
  $$EX_{(r)} = E\Big[\sum_{j=1}^{r} \frac{Z_j}{n-j+1}\Big] = \sum_{j=1}^{r} \frac{EZ_j}{n-j+1} = \sum_{j=1}^{r} \frac{1}{n-j+1}$$
  a pre rozptyl platí
  $$\mathsf{var}\, X_{(r)} = \sum_{j=1}^{r} \frac{\mathsf{var}\, Z_j}{(n-j+1)^2} = \sum_{j=1}^{r} \frac{1}{(n-j+1)^2},$$
  pričom prvá rovnosť v poslednom riadku plynie z nezávislosti náhodných veličín $Z_1, \dots, Z_n$.

- Nechť $X_i \sim \mathsf{Exp}(\lambda)$ a $n = 2k+1$. Je $X_{(k+1)}$ nestranným a/nebo konsistentním odhadem mediánu rozdělení $\mathsf{Exp}(\lambda)$?

  Řešení:
  Nechť $X_i \sim Exp(\lambda)$, tak, že $EX_i = \frac{1}{\lambda}$, pre $\lambda > 0$. Potom nutne $\lambda X_i \sim Exp(1)$ a tudíž $E[\lambda X_i] = \lambda \cdot \frac{1}{\lambda} = 1$. Tym pádom dostaneme aj
  $$EX_{(k+1)} = \frac{1}{\lambda} E[\lambda X_{(k+1)}] = \frac{1}{\lambda} \sum_{j=1}^{k+1} \frac{E[\lambda X_i]}{n-j+1} = \frac{1}{\lambda} \sum_{j=1}^{k+1} \frac{1}{n-j+1}$$
  a tiež
  $$\mathsf{var}\, X_{(k+1)} = \frac{1}{\lambda^2} \mathsf{var}\, [\lambda X_{(k+1)}] = \frac{1}{\lambda^2} \sum_{j=1}^{k+1} \frac{\mathsf{var}\, [\lambda X_i]}{(n-j+1)^2} = \frac{1}{\lambda^2} \sum_{j=1}^{k+1} \frac{1}{(n-j+1)^2},$$
  kde druhá rovnosť predchádzajúceho výrazu plynie (opäť) z nezávislosti náhodných veličín $X_1, \dots, X_n$. Pre medián náhodných veličín $X_i \sim Exp\lambda$ platí, že je to hodnota $x_m \in \mathbb{R}$ taková, že
  $$\begin{aligned} \frac{1}{2} &= 1 - e^{-\lambda x_m} \\ \frac{1}{2} &= e^{-\lambda x_m} \\ x_m &= \frac{1}{\lambda} \log 2. \end{aligned}$$
  Je zrejmé, že zároveň dostávame
  $$EX_{(k+1)} = \frac{1}{\lambda}\Big(\frac{1}{n} + \frac{1}{n-1} + \dots + \frac{1}{n-k}\Big) \neq \frac{1}{\lambda} \log 2,$$
  takže náhodná veličina $X_{(k+1)}$ nie je nestranným odhadom pre teoretický medián $x_m = \frac{1}{\lambda} \log 2$. Ukážeme ale, že sa jedná o asymptoticky nestranný odhad (t.j., $EX_{(k+1)} \to \frac{1}{\lambda} \log 2$, pre $n \to \infty$, kde $k = \frac{n-1}{2}$).

Využijeme k tomu tzv. Euler–Mascheroni konštantu $C = 0.5772...$, ktorá je definovaná ako limita

$$C = \lim_{n\to\infty} \Big(\sum_{k=1}^{n} \frac{1}{k} - \log n \Big).$$

Porovnáme dva čiastočné súčty pre ľubovolné nepárne (liché) $n \in \mathbb{N}$ a $\frac{n-1}{2}$:

$$\begin{array}{llllllll} 1+\frac{1}{2} & +\dots & +\frac{1}{\frac{n-1}{2}} & +\dots & +\frac{1}{n} & -\log n & \to C & \text{pre } n\to\infty; \\ 1+\frac{1}{2} & +\dots & +\frac{1}{\frac{n-1}{2}} & & & -\log\frac{n-1}{2} & \to C & \text{tiež pre } n\to\infty. \end{array}$$

Odčítaním hodnôt po stĺpcoch a uvedomením si faktu, že $\frac{1}{\frac{n-1}{2}+1} = \frac{1}{\frac{n+1}{2}} = \frac{1}{n-k}$, dostaneme

$$\Big(\frac{1}{n-k} + \dots + \frac{1}{n}\Big) - \Big(\log n - \log\frac{n-1}{2}\Big) \to (C - C) = 0 \qquad \text{pre } n \to \infty$$

a preto tiež platí, že

$$lim_{n\to\infty} EX_{(k+1)} = \lim_{n\to\infty}\Big[\frac{1}{\lambda}\Big(\frac{1}{n-k} + \dots + \frac{1}{n}\Big)\Big] = \frac{1}{\lambda}\lim_{n\to\infty}\Big(\log n - \log\frac{n-1}{2}\Big),$$

pričom platí, že

$$\lim_{n\to\infty}\Big(\log n - \log\frac{n-1}{2}\Big) = \lim_{n\to\infty}\Big(\log 2 - \log\frac{n}{n-1}\Big) = \log 2.$$

Jedná sa teda o asymptotický nestranný odhad teoretického mediánu $x_m \in \mathbb{R}$. Pro konzistenci ešte ukážeme, že $\mathsf{var}\, X_{(k+1)} \to 0$. Platí, že

$$\mathsf{var}\, X_{(k+1)} = \frac{1}{\lambda^2}\Big(\frac{1}{n^2} + \dots + \frac{1}{(n-k)^2}\Big),$$

a tiež, že $\sum_{i=1}^{n} \frac{1}{i^2} \to \frac{\pi^2}{6}$ pre $n \to \infty$. Túto konvergentnú sumu ale môžeme rozpísať na dve časti

$$\sum_{i=1}^{n} \frac{1}{i^2} = \sum_{i-1}^{\frac{n-1}{2}} \frac{1}{i^2} + \sum_{i=\frac{n+1}{2}}^{n} \frac{1}{i^2} \underset{n\to\infty}{\longrightarrow} \frac{\pi^2}{6},$$

pričom platí, že

$$\sum_{i-1}^{\frac{n-1}{2}} \frac{1}{i^2} \underset{n\to\infty}{\longrightarrow} \frac{\pi^2}{6},$$

pre n idúce do nekonečna a tým pádom druhá suma (zbytok) konverguje k nule, t.j.

$$\sum_{i=\frac{n+1}{2}}^{n} \frac{1}{i^2} \underset{n\to\infty}{\longrightarrow} 0.$$

Konečne teda dostávame, že $\mathsf{var}\, X_{(k+1)} \to 0$ a náhodná veličina $X_{(k+1)}$ (resp. $(k+1)$-vá poradová štatistika) je konzistentným odhadom teoretického mediánu $x_m \in \mathbb{R}$.

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

Řešení:
Je zrejmé, že transformované náhodné veličiny $X_i - \theta_1$ pre $i = 1, \dots, n$ majú rovnomerné rozdelenie na intervale $(0, \theta_2 - \theta_1)$, t.j., $X_1 - \theta_1, \dots, X_n - \theta_1 \sim R(0, \theta_2 - \theta_1)$. Zároveň ale náhodné veličiny $Z_i \equiv (X_i - \theta_1)/(\theta_2 - \theta_1)$ pre $i = 1, \dots, n$ majú rovnomerné rozdelenie na intervale $(0, 1)$.

Pre $Z_1, \dots Z_n \sim R(0, 1)$ už vieme, že (opakovanie)

$Z_{(n)}$ má hustotu $f_{(n)}(x) = nx^{n-1}$ pre $x \in (0, 1)$ a nula inak;
$Z_{(1)}$ má hustotu $f_{(1)}(x) = n(1-x)^{n-1}$ pre $x \in (0, 1)$ a nula inak.

To znamená, že pre strednú hodnotu dostaneme

$$EZ_{(n)} = E\Big[\frac{X_{(n)} - \theta_1}{\theta_2 - \theta_1}\Big] = \int_{-\infty}^{\infty} x f_{(n)}(x) \mathrm{d}x = \int_0^1 nx^n \mathrm{d}x = \frac{n}{n+1};$$

a podobne aj

$$EZ_{(1)} = E\Big[\frac{X_{(1)} - \theta_1}{\theta_2 - \theta_1}\Big] = \int_{-\infty}^{\infty} x f_{(1)}(x) \mathrm{d}x = \int_0^1 nx(1-x)^{n-1} \mathrm{d}x = \frac{1}{n+1}.$$

(resp. je možné priamo využiť Beta rozdelenie—viď riešenie príkladu A1(b))

Vďaka linearite strednej hodnoty ale z tohto dokážeme vyjadriť aj stredné hodnoty $EX_{(n)}$ a $EX_{(1)}$ ako funkcie neznámych parametrov $\theta_1, \theta_2 \in \mathbb{R}$, takých, že $\theta_1 < \theta_2$. Dostávame

$$EX_{(n)} = \frac{n}{n+1}\theta_2 + \frac{1}{n+1}\theta_1$$

a tiež

$$EX_{(1)} = \frac{n}{n+1}\theta_1 + \frac{1}{n+1}\theta_2.$$

To je ale sústava dvoch rovníc pre dve nezáme (θ_1 a θ_2). Riešením sústavy rovníc dostaneme riešenie:

$$\theta_1 = \frac{n+1}{n^2+1}(nEX_{(1)} - EX_{(n)}) \qquad \text{a} \qquad \theta_2 = \frac{n+1}{n^2+1}(nEX_{(n)} - EX_{(1)}).$$

Príslušné nestranné odhady pre parametre $\theta_1, \theta_2 \in \mathbb{R}$ sú

$$\widehat{\theta_1} = \frac{n+1}{n^2+1}(nX_{(1)} - X_{(n)}) \qquad \text{a} \qquad \widehat{\theta_2} = \frac{n+1}{n^2+1}(nX_{(n)} - X_{(1)}).$$

(pretože $E\widehat{\theta_1} = \frac{n+1}{n^2+1}(nEX_{(1)} - EX_{(n)}) = \theta_1$ a $E\widehat{\theta_2} = \frac{n+1}{n^2+1}(nEX_{(n)} - EX_{(1)}) = \theta_2$)

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

Řešení:
V prvom rade je nutné si uvedomiť tzv. generickú vlastnosť Poissonovho rozdelenia. To znamená, že pre X_1 a X_2 nezávislé náhodné veličiny s Poissonovým rozdelením s parametrom $\lambda > 0$ platí,

že $(X_1 + X_2)$ je náhodná veličina, ktorá má opäť Poissonové rozdelenie, tentokrát s parametrom 2λ (samostatne odvodiť pomocou vety o konvolúcii). Rekurzívne teda dostaneme, že

$$\sum_{i=1}^{n} X_i \sim Poiss(n\lambda).$$

Túto vlastnosť využijeme pre dôkaz konzistencie, t.j. výpočet strednej hodnoty odhadu $\widehat{p}_0$.

$$\begin{aligned}
E\widehat{p}_0 &= E\left[\left(1 - \frac{1}{n}\right)^{n\overline{X}_n}\right] = E\left[\left(1 - \frac{1}{n}\right)^{\sum_{i=1}^{n} X_i}\right] \\
&= \sum_{k=1}^{\infty}\left(1 - \frac{1}{n}\right)^{k} \cdot P\Big[\sum_{i=1}^{\infty} X_i = k\Big] = \sum_{k=0}^{\infty}\left(1 - \frac{1}{n}\right)^{k} \cdot e^{-n\lambda}\frac{(n\lambda)^k}{k!} \\
&= e^{-n\lambda} \cdot \sum_{i=1}^{\infty}\left[\left(1 - \frac{1}{n}\right)n\lambda\right]^{k} \cdot \frac{1}{k!} = e^{-n\lambda} \cdot e^{(1-1/n)\lambda n} = e^{-n\lambda} \cdot e^{n\lambda-\lambda} \\
&= e^{-\lambda} = e^{-\lambda} \cdot \frac{\lambda^0}{0!} = P[X_1 = 0] = p_0.
\end{aligned}$$

Tretia rovnosť plynie z definície strednej hodnoty diskrétnej náhodnej veličiny, štvrtá rovnosť z faktu, že $\sum X_i \sim Poiss(n\lambda)$. Odhad $\widehat{p}_0$ pro neznámy parameter $p_0 = P[X_1 = 0]$ je teda nestranným odhadom.

Pre vyšetrenie konzistencie postačí ukázať, že $\widehat{p}_0 \longrightarrow p_0$ v pravdepodobnosti pre $n \to \infty$, resp. platí

$$\left(1 - \frac{1}{n}\right)^{n\overline{X}_n} \xrightarrow[n\to\infty]{P} p_0.$$

Ľavú stranu rovnosti môžeme upraviť následovne:

$$\left(1 - \frac{1}{n}\right)^{n\overline{X}_n} = \left(1 - \frac{1}{n}\right)^{\sum_{i=1}^{n} X_i} = \left(1 - \frac{1}{n} \cdot \frac{\sum_{i=1}^{n} X_i}{\sum_{i=1}^{n} X_i}\right)^{\sum_{i=1}^{n} X_i} = \left(1 - \frac{\overline{X}_n}{\sum_{i=1}^{n} X_i}\right)^{\sum_{i=1}^{n} X_i},$$

kde sme využili rozšírenie zlomku (vynásobenie) hodnotou jedna vo vhodnom tvare. Následne si stačí uvedomiť že

$$\sum_{i=1}^{n} X_i \xrightarrow[n\to\infty]{P} \infty, \quad \text{zároveň} \quad \overline{X}_n \xrightarrow[n\to\infty]{P} \lambda,$$

a tiež

$$\lim_{m\to\infty}\left(1 - \frac{a}{m}\right)^{m} = e^{-a}.$$

Prvé dve konvergencie sú pritom konvergencie v pravdepodobnosti (t.j. celočíselná nahodná veličina $\sum X_i$ diverguje v pravdepodobnosti do nekonečna a analogicky náhodná veličina $\overline{X}_n$ konverguje (slabě/silně)—použitím silného/slabého zákona veľkých čísel—ku svojej strednej hodnote) a posledná konvergencia je deterministická (t.j. reálna postupnosť konverguje ku svojej limite). Použitím vety o zloženej limite (resp. vety o limite zloženej funkcie) dostávame, že

$$\left[\left(1 - \frac{\overline{X}_n}{\sum_{i=1}^{n} X_i}\right)^{\sum_{i=1}^{n} X_i}\right] \xrightarrow[\sum X_i \longrightarrow\infty]{P} e^{-\lambda}, \qquad \text{pre } n \to \infty \text{ a } \sum X_i \to \infty \text{ v pravdepodobnosti.}$$

Odhad $\widehat{p}_0$ je teda silně/slabě konzistentným odhadem neznámeho parametru $p_0 = P[X_1 = 0]$.

B Výsledky

B1. Není nestranný, protože $EX_{(n)} = \frac{n}{n+1}\theta$;
Je konzistentný, protože $EX_{(n)} \to \theta$ a $VarX_{(n)} = (n/(n+2) - n^2/(n+1)^2)\theta \to 0$, pro $n \to \infty$;

B2. (a) $E\widehat{\theta}_n = \frac{4}{3n}\sum_{i=1}^n \int_0^\theta \frac{3}{\theta^3}x^3 \mathrm{d}x = \theta$
(b) $E\widetilde{\theta}_n = \frac{3n+1}{3n}\int_0^\theta 3n\frac{1}{\theta^3}x^{3n}\mathrm{d}x = \theta$
(c) $Var\widehat{\theta}_n = \frac{1}{15n}\theta^2 \qquad Var\widetilde{\theta}_n = \frac{1}{9n^2+6n}\theta^2$

B3. Napr. $\widehat{\theta}_n = \frac{n}{n-1}(\overline{X}_n - (\overline{X}_n)^2)$

B4. Náhodný součet $\sum_{i=1}^n X_i$ má Erlangovo (Gamma) rozdělení s parametrami $\lambda > 0$ a $n \in \mathbb{N}$. Příslušná hustota je $f(x) = \frac{1}{(n-1)!}\lambda^n x^{n-1}e^{-\lambda x} \cdot \mathbb{I}_{\{x \geq 0\}}$. Zároveň $\theta \in (0,1)$ a tudíž i $\widehat{\theta}_n \in (0,1)$ s.j. Výsledek proto vede na výpočet integrálu $\int_u^\infty (1 - \frac{u}{x})^{n-1} f(x)\mathrm{d}x$.

B5. (a) Pro libovolné $n \in \mathbb{N}$ je odhad $T : \{0,1\}^{\otimes n} \to (0,1)$ pouze polynomiální funkci v $p \in (0,1)$. Lze ukázat napr. indukcíí: Pro $n = 1$ je odhad $\widehat{\theta}_1 = T(X_1)$ měritelnou funkci z množiny $\{0,1\}$ do intervalu $(0,1)$. Pro nestrannost musí platit $E\widehat{\theta}_1 = ET(X_1) = \sum_{x\in\{0,1\}} T(x)p^x(1-p)^{1-x} = p(T(1)-T(0))+T(0)$. Jelikož $T(\cdot)$ nezávisí na $p \in (0,1)$, nelze dosáhnout rovnost $E(T(X_1)) = 1/p$. Pro $n = 2$ dostaneme analogicky $ET(X_1,X_2) = \sum_{x,y\in\{0,1\}} T(x,y)p^(x+y)(1-p)^{2-x-y}$ což lze zapsát jako $T(1,1)p^2 + 2T(1,0)p(1-p) + T(0,0)(1-p)^2$ (teda polynom druhého řádu).

(b)$E(Z+1) = EZ + 1 = 1/p = \theta \quad \forall\theta \in (1,\infty)$

B6. Transformovaný náhodný výběr $V_i = -\log U_i$ ma exponenciální rozdělení s paramerem $\lambda = 1$ (t.j., stejné rozdělení, jako náhodný výběr $X_1, \ldots, X_n$). K první části si postačí uvědomit, že transformace přehodí pořadí v uspořádanem výberu, tudíž $V_{(1)} = -\log U_{(n)}, \ldots, V_{(n)} = -\log U_{(1)}$. Obecně teda $V_{(k)} = -\log U_{(n-k+1)}$, pro $k \in \{1,\ldots,n\}$.

K druhé části lze použít analogicky postup: Z transformace $-\log Q_r = r\big[-\log U_{(r)} - (-\log U_{(r+1)})\big]$ a z příkladu A2 vyplýva, že náhodné veličiny $-\log Q_r$, pro $r = 1,\ldots,n$, mají exponenciální rozdělení s parametrem $\lambda = 1$ a jsou nezávislé (a dodefinovali jsme hodnotu $U_{(n+1)}$ jako ∞). Použitím zpětné (inverzní) transformace dostaneme $Q_r \sim R(0,1)$ (a nezávislost se zachová).

B7. Nechť $EX_i = 1/\lambda$ a $EY_i = 1/\nu$. Pak platí, že $2\lambda\sum_{i=1}^n X_i \sim \Gamma(n, 1/2) \equiv \chi^2_{2n}$ (t.j., χ^2 rozdělení s $2n$ stupni volnosti) a zároveň aj $2\nu\sum_{i=1}^m Y_i \sim \Gamma(m,1/2) \equiv \chi^2_{2m}$. Z definice F rozdělení teda platí

$$P\Big[F_{2n,2m}(\alpha/2) \leq \frac{2\lambda\sum_{i=1}^n X_i}{2\nu\sum_{i=1}^m Y_i}\frac{2m}{2n} \leq F_{2n,2m}(1-\alpha/2)\Big] = 1-\alpha,$$

kde $F_{2n,2m}(\cdot)$ je příslušný kvantil Fisherovho F rozdělení (s $2n$ a $2m$ stupni volnosti).

B8. (a)