

PŘÍKLADY K PROCVIČENÍ Z MATEMATICKÉ STATISTIKY 1

Poslední úprava 18. října 2019.

OBSAH

1	Mnohorozměrné normální rozdělení	1
2	Konvergence posloupnosti náhodných veličin	1
3	Vlastnosti a asymptotické rozdělení odhadů	3
4	Intervalové odhady	6
5	Empirické odhady	7
6	Výsledky	9

1 MNOHORozměRNÉ NORMÁLNÍ RozDĚLENÍ

1. Nechť $\mathbf{X} \sim N_n(\mathbf{0}, \Sigma)$.

(a) Vyjádřete $E\|\mathbf{X}\|^2$ pomocí prvků matice Σ .

(b) Nechť Σ má vlastní čísla pouze 0 nebo 1. Jaké je rozdělení $\|\mathbf{X}\|^2$?

Návod: Využijte spektrální rozklad Σ .

2. Nechť náhodný vektor $\begin{pmatrix} X \\ Y \end{pmatrix}$ má 2-rozměrné normální rozdělení s nulovou střední hodnotou a rozptylovou maticí $\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$. Nalezněte lineární transformaci $\begin{pmatrix} U \\ V \end{pmatrix} = \mathbf{B} \begin{pmatrix} X \\ Y \end{pmatrix}$ takovou, že U a V jsou nezávislé.

3. Nechť $X \sim N(0, 1)$ a definujeme $Y = X \cdot I[|X| > A] - X \cdot I[|X| \leq A] = \begin{cases} X & |X| > A \\ -X & |X| \leq A \end{cases}$, kde $A > 0$.

(a) Ukažte, že obě veličiny X i Y mají normální rozdělení, ale vektor $(X, Y)^T$ nemá sdružené normální rozdělení.

(b) Pro jaké A jsou X a Y nekorelované?

2 KONVERGENCE POSLOUPNOSTI NÁHODNÝCH VELIČIN

1. Nechť X_1, X_2, \dots jsou nezávislé veličiny s rovnoměrným rozdělením na $[0, 1]$. Definujme $Y_n = \min\{X_1, \dots, X_n\}$. Ukažte, že pro $n \rightarrow \infty$ platí (a) $Y_n \xrightarrow{D} 0$, (b) $Y_n \xrightarrow{P} 0$.

2. Nechť X je nějaká náhodná veličina a $X_n = X(1 + 1/n)$. Ukažte, že $X_n \xrightarrow{P} X$.
3. Nechť $X \sim \mathbf{N}(0, 1)$ a $X_n = (-1)^n X$. Ukažte, že $X_n \xrightarrow{D} X$, ale neplatí $X_n \xrightarrow{P} X$.
4. Nechť pro každé $n \in \mathbb{N}$ je X_n náhodná veličina s normálním rozdělením $\mathbf{N}(\mu, \sigma_n^2)$, kde $\sigma_n > 0$. Dále předpokládejte, že $\sigma_n \rightarrow \sigma$.
 - (a) Dokažte, že pokud $\sigma > 0$, potom $X_n \xrightarrow{D} \mathbf{N}(\mu, \sigma^2)$.
 - (b) Dokažte, že pokud $\sigma = 0$, potom $X_n \xrightarrow{D} \delta_\mu$, kde δ_μ je Diracova míra v bodě μ .
5. Nechť X_n jsou nezávislé náhodné veličiny a nechť X_n má alternativní rozdělení $\text{Alt}(1/n)$, tj. $\mathbf{P}(X_n = 1) = 1/n$ a $\mathbf{P}(X_n = 0) = 1 - 1/n$. Ukažte, že $X_n \xrightarrow{P} 0$, ale neplatí $X_n \xrightarrow{s.j.} 0$.
6. Nechť X_n má exponenciální rozdělení se střední hodnotou $1/n$. Ukažte, že $X_n \xrightarrow{P} 0$.
7. Nechť X_n má diskrétní rovnoměrné rozdělení na $1, 2, \dots, n$. Nechť $Y_n = X_n/n$. Ukažte, že $\{Y_n\}$ konverguje v distribuci k rovnoměrnému rozdělení na $[0, 1]$.
8. Nechť $\{X_n\}$ je posloupnost náhodných veličin takových, že X_n má geometrické rozdělení s parametrem λ/n . Definujme $Y_n = 1/n \cdot X_n$. Vyšetřete konvergenci v distribuci veličin $\{Y_n\}$ pro $n \rightarrow \infty$.
9. Uvažujte posloupnost náhodných veličin $\{X_n\}$, kde X_n má distribuční funkci

$$F_n(x) = \begin{cases} 1 - \left(1 - \frac{1}{n}\right)^{nx}, & x \geq 0, \\ 0 & x < 0. \end{cases}$$

Vyšetřete konvergenci v distribuci veličin $\{X_n\}$.

10. Uvažujte posloupnost náhodných veličin $X_n \sim \text{Bi}(n, \lambda/n)$. Vyšetřete konvergenci v distribuci posloupnosti $\{X_n\}$ pro $n \rightarrow \infty$.
11. Nechť X_n má distribuční funkci

$$F_n(x) = \frac{e^{n(x-1)}}{1 + e^{n(x-1)}}, \quad x \geq 1,$$

a $F_n(x) = 0$ pro $x < 1$.

- (a) Vyšetřete konvergenci v distribuci $\{X_n\}$.
 - (b) Vyšetřete konvergenci v pravděpodobnosti $\{X_n\}$.
12. Nech X_1, X_2, \dots tvoria náhodný výber z rozdelenia $\mathbf{N}(0, 1)$. Vyšetrite konvergenciu postupnosti $\{X_n\}$ v zmysle
 - (a) skoro iste,
 - (b) v pravdepodobnosti,
 - (c) v distribúcii.
 13. Nechť X_1, X_2, \dots je náhodný výběr z exponenciálního rozdělení s parametrem 1.
 - (a) Nalezněte rozdělení $M_n = \min_{1 \leq i \leq n} X_i$. O jaké rozdělení se jedná?

- (b) Vytřete konvergenci $\{M_n\}$ v pravděpodobnosti pro $n \rightarrow \infty$.
 (c) Vyšetřete konvergenci v distribuci $n^\alpha \cdot M_n$ pro $n \rightarrow \infty$ pro $\alpha \in (0, 1]$.
 (d) Nechť $U_n = \max_{1 \leq i \leq n} X_i$. Rozhodněte, jak se chová $\{U_n\}$ pro $n \rightarrow \infty$. Dále vyšetřete konvergenci v distribuci posloupnosti veličin $\{U_n - \log n\}$.

14. Nech $(X_1, X_2)^\top$ je náhodný vektor s rozdělením $\mathbf{N}_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right)$.

- (a) Najděte rozdělení náhodného vektoru $(X_1, -X_1)^\top$.
 (b) Ukážete, že oba marginální rozdělení vektorů $(X_1, X_2)^\top$ a $(X_1, -X_1)^\top$ jsou identická.
 (c) Z výsledků v částech (a) a (b) usudíte, že z konvergence $Y_n \xrightarrow{D} Y$ a $Z_n \xrightarrow{D} Z$ neplynou $(Y_n, Z_n)^\top \xrightarrow{D} (Y, Z)^\top$.
 (d) Můžeme povedat něco více ak v části (c) navíc předpokládáme nezávislost náhodných veličin Y_n a Z_n pro každé n ?

3 VLASTNOSTI A ASYMPTOTICKÉ ROZDĚLENÍ ODHADŮ

1. Nechť X_1, \dots, X_n je náhodný výběr z Poissonova rozdělení $\text{Po}(\lambda)$, $\lambda > 0$. Podrobně dokažte, že

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(u_{\alpha/2} \leq \frac{\sqrt{n}(\bar{X}_n - \lambda)}{\sqrt{\bar{X}_n}} \leq u_{1-\alpha/2} \right) = 1 - \alpha,$$

kde $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ a u_β je β -kvantil normovaného normálního rozdělení.

2. Nechť X_1, \dots, X_n je náhodný výběr z geometrického rozdělení, tj.

$$\mathbf{P}(X_1 = k) = p(1-p)^k, \quad k = 0, 1, 2, \dots$$

- (a) Najděte odhad T_n parametru p metodou maximální věrohodnosti. Provéřte jeho konzistenci a odvoďte asymptotické rozdělení. Umíte rozhodnout i o nestrannosti?
 (b) Rozhodněte, zda $U_n = \mathbb{I}\{X_n = 0\}$ je nestranný a konzistentní odhad parametru p .
 (c) Rozhodněte, zda $V_n = \frac{1}{n} \sum_{i=1}^n \mathbb{I}\{X_i = 0\}$ je nestranný a konzistentní odhad parametru p .
 (d) Jaké je přesné rozdělení nV_n ? Jaké je asymptotické rozdělení V_n ?
 (e) Rozhodněte, zda $W_n = \frac{n+1}{n + \sum_{i=1}^n X_i}$ je konzistentní odhad parametru p .

3. Nechť X_1, \dots, X_n je náhodný výběr z alternativního rozdělení s parametrem $p \in (0, 1)$.

- (a) Uvažujte odhad rozptylu $\text{Var } X_1 = p(1-p)$ ve tvaru $U_n = \bar{X}_n(1 - \bar{X}_n)$. Vyšetřete jeho nestrannost, konzistenci a odvoďte asymptotické rozdělení.
 (b) Rozhodněte, zda je odhad $V_n = X_1(1 - X_n)$ nestranný a konzistentní odhad parametru $p(1-p)$.
 (c) Uvažujte výběrový rozptyl S_n^2 . Rozepište si, jak se S_n^2 liší od U_n pro tento případ alternativního rozdělení.

4. Nechť X je diskrétní náhodná veličina s useknutým Poissonovým rozdělením, tj.

$$P(X_1 = k) = \frac{1}{e^\lambda - 1} \frac{\lambda^k}{k!}, \quad k = 1, 2, 3, \dots$$

- (a) Dokažte, že $T(X) = 1 + (-1)^X$ je nestranný odhad parametru $\theta_X = 1 - e^{-\lambda}$.
- (b) Dokažte, že $T(X) = 1 + (-1)^X$ je jediný nestranný odhad parametru $\theta_X = 1 - e^{-\lambda}$.
- (c) Zamyslete se nad užitečností tohoto odhadu.

5. Nechť X_1, \dots, X_n je náhodný výběr z binomického rozdělení $\text{Bi}(m, p)$, $p \in (0, 1)$.

- (a) Ukažte, že $\hat{p}_n = \frac{1}{nm} \sum_{i=1}^n X_i$ je nestranný odhad parametru p .
- (b) Ukažte, že neexistuje nestranný odhad parametru $\theta_X = \frac{1}{\sqrt{p(1-p)}}$.

6. Nechť X_1, \dots, X_n je náhodný výběr z Poissonova rozdělení $\text{Po}(\lambda)$, $\lambda > 0$. Uvažujte parametr $\theta_X = P(X_1 = 0) = e^{-\lambda}$.

- (a) Rozhodněte, zda je odhad $\hat{\theta}_n = e^{-\bar{X}_n}$ nestranný a konzistentní odhad parametru θ_X . Pokud není nestranný, spočítejte jeho vychýlení.
- (b) Rozhodněte, zda je odhad $\tilde{\theta}_n = \frac{1}{n} \sum_{i=1}^n \mathbb{I}\{X_i = 0\}$ nestranný a konzistentní odhad parametru θ_X . Pokud není nestranný, spočítejte jeho vychýlení.
- (c) Porovnejte odhady $\hat{\theta}_n$ a $\tilde{\theta}_n$ na základě jejich středních čtvercových chyb.
- (d) Porovnejte odhady $\hat{\theta}_n$ a $\tilde{\theta}_n$ na základě jejich asymptotických rozptylů.

7. Mějme náhodný výběr X_1, \dots, X_n z exponenciálního rozdělení $\text{Exp}(\lambda)$, $\lambda > 0$.

- (a) Rozhodněte, zda $\hat{\lambda}_n = \frac{n}{\sum_{i=1}^n X_i}$ je nestranný a konzistentní odhad parametru λ .
- (b) Rozhodněte, zda $\hat{\theta}_n = \frac{1}{n} \sum_{i=1}^n X_i$ je nestranný a konzistentní odhad parametru $\theta_X = \frac{1}{\lambda}$.
- (c) Uvažujte odhad $T_n = n \cdot \min_{1 \leq i \leq n} X_i$. Rozhodněte, zda se jedná o nestranný a konzistentní odhad $\theta_X = \frac{1}{\lambda}$.

8. Nechť X_1, \dots, X_n je náhodný výběr z rozdělení s hustotou $f(x) = \frac{a^p}{\Gamma(p)} x^{p-1} e^{-ax} I[x \geq 0]$, pro $a > 0$ a $p > 0$. Uvažujte odhady

$$\hat{a}_n = \frac{\bar{X}_n}{S_n^2}, \quad \hat{p}_n = \frac{\bar{X}_n^2}{S_n^2}.$$

Dokažte konzistenci těchto odhadů, a odvoďte sdružené asymptotické rozdělení vektoru $(\hat{a}_n, \hat{p}_n)^\top$.

Návod: Využijte věty ze skript o sdružené asymptotické normalitě $(\bar{X}_n, S_n^2)^\top$ a dále pak toho, že pro gama rozdělení platí $\mathbf{E}X_1 = \frac{p}{a}$, $\text{Var}(X_1) = \frac{p}{a^2}$, $\gamma_3 = \frac{2}{\sqrt{p}}$ a $\gamma_4 = \frac{6}{p} + 3$.

9. Nechť X_1, \dots, X_n je náhodný výběr z rovnoměrného rozdělení $R(0, \theta)$, kde $\theta > 0$.

- (a) Rozhodněte, zda $U_n = 2\bar{X}_n$ je nestranný a konzistentní odhad parametru θ . Pokud není, spočítejte jeho vychýlení.
- (b) Spočítejte střední čtvercovou chybu odhadu U_n .

- (c) Rozhodněte, zda $T_n = \max_{1 \leq i \leq n} X_i$ je nestranný a konzistentní odhad parametru θ . Pokud není nestranný, spočítejte jeho vychýlení.
- (d) Spočítejte střední čtvercovou chybu odhadu T_n .
- (e) Najděte asymptotické rozdělení $n(T_n - \theta)$.
- (f) Kterému z odhadů $\hat{\theta}_n, \tilde{\theta}_n$ byste dali přednost?

10. Mějme náhodný výběr X_1, \dots, X_n z rozdělení

$$f_X(x) = \begin{cases} \lambda e^{-\lambda(x-\delta)}, & x \in (\delta, \infty), \\ 0, & \text{jinak,} \end{cases}$$

kde $\delta \in \mathbb{R}$ je neznámý parametr a $\lambda > 0$ je známá konstanta.

- (a) Rozhodněte, zda $\hat{\delta}_n = \min_{1 \leq i \leq n} X_i$ je nestranný a konzistentní odhad parametru δ . Pokud není, spočítejte jeho vychýlení.
- (b) Spočítejte střední čtvercovou chybu odhadu $\hat{\delta}_n$.
- (c) Rozhodněte, zda odhad $\tilde{\delta}_n = \max_{1 \leq i \leq n} X_i$ je konzistentní odhad parametru δ .
- (d) Rozhodněte, zda odhad $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ je nestranný a konzistentní odhad parametru δ .

11. Nechť X_1, \dots, X_n je náhodný výběr z rovnoměrného rozdělení $R(a - b/2, a + b/2)$ s hustotou

$$f(x) = \begin{cases} \frac{1}{b}, & a - \frac{b}{2} < x < a + \frac{b}{2}, \\ 0, & \text{jinak,} \end{cases}$$

kde $a < b$. Označme $X_{(1)} = \min_{1 \leq i \leq n} X_i$ a $X_{(n)} = \max_{1 \leq i \leq n} X_i$.

- (a) Rozhodněte, zda $\hat{a}_n = \frac{X_{(1)} + X_{(n)}}{2}$ je nestranný a konzistentní odhad parametru a . Pokud není, spočítejte jeho vychýlení.
- (b) Spočítejte střední čtvercovou chybu odhadu \hat{a}_n .
- (c) Rozhodněte, zda $\hat{b}_n = X_{(n)} - X_{(1)}$ je nestranný a konzistentní odhad parametru b . Pokud není, spočítejte jeho vychýlení.
- (d) Spočítejte střední čtvercovou chybu odhadu \hat{b}_n .

Návod: Uvědomte si, že náhodný výběr X_1, \dots, X_n můžeme „vyrobit“ pomocí lineární transformace jako $X_i = a + b(Y_i - \frac{1}{2})$, $i = 1, \dots, n$, kde Y_1, \dots, Y_n je náhodný výběr z rovnoměrného rozdělení $R(0, 1)$. Sdružená hustota náhodného vektoru $(Y_{(1)}, Y_{(n)})$ pak je

$$f_{(Y_{(1)}, Y_{(n)})}(y_1, y_2) = n(n-1)(y_2 - y_1)^{n-2} \mathbb{I}\{0 < y_1 < y_2 < 1\}.$$

12. Nechť $(X_1, Y_1)^T, \dots, (X_n, Y_n)^T$ je náhodný výběr z dvourozměrného rozdělení s regulární varianční maticí. Dokažte, že odhad

$$\hat{\rho}_n = \frac{\sum_{i=1}^n (X_i - \bar{X}_n)(Y_i - \bar{Y}_n)}{\sqrt{\sum_{i=1}^n (X_i - \bar{X}_n)^2 \sum_{i=1}^n (Y_i - \bar{Y}_n)^2}}$$

je konzistentním odhadem korelačního koeficientu $\rho = \frac{\text{Cov}(X_1, Y_1)}{\sqrt{\text{Var}(X_1) \text{Var}(Y_1)}}$.

4 INTERVALOVÉ ODHADY

1. Nechť X_1, \dots, X_n je náhodný výběr z alternativního rozdělení s parametrem $p \in (0, 1)$.
 - (a) Najděte asymptotické rozdělení \bar{X}_n a pomocí tohoto rozdělení sestavte oboustranný (a pravostranný) interval spolehlivosti pro p .
 - (b) Najděte asymptotické rozdělení $\arcsin(\sqrt{\bar{X}_n})$ a pomocí tohoto rozdělení sestavte oboustranný (a levostranný) interval spolehlivosti pro p .
2. Mějme náhodný výběr X_1, \dots, X_n z exponenciálního rozdělení s hustotou

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \in (0, \infty), \\ 0, & \text{jinak,} \end{cases}$$

kde $\lambda > 0$.

- (a) Najděte asymptotické rozdělení výběrového průměru \bar{X}_n a pomocí něj zkonstruuje intervalový odhad pro parametr λ .
 - (b) Najděte asymptotické rozdělení odhadu parametru λ daného předpisem $\hat{\lambda}_n = \frac{1}{\bar{X}_n}$ a pomocí tohoto rozdělení sestavte intervalový odhad pro parametr λ .
 - (c) Najděte transformaci, která stabilizuje asymptotický rozptyl náhodné veličiny $\frac{1}{\bar{X}_n}$ a pomocí této transformace sestavte intervalový odhad pro parametr λ .
3. Nechť X_1, \dots, X_n je náhodný výběr z geometrického rozdělení, tj.

$$P(X_1 = k) = p(1-p)^k, \quad k = 0, 1, 2, \dots$$

- (a) Najděte asymptotické rozdělení odhadu $\hat{p}_n = \frac{1}{1 + \bar{X}_n}$ a na základě tohoto rozdělení odvoďte klasický asymptotický interval spolehlivosti pro p .
 - (b) Najděte asymptotické rozdělení odhadu $\tilde{p}_n = \frac{1}{n} \sum_{i=1}^n \mathbb{I}\{X_i = 0\}$ a na základě tohoto rozdělení odvoďte klasický asymptotický interval spolehlivosti pro p .
 - (c) Porovnejte intervaly z (a) a (b). Který vám přijde vhodnější?
4. Nechť X_1, \dots, X_n je náhodný výběr z exponenciálního rozdělení s parametrem $\lambda_X > 0$ a Y_1, \dots, Y_n je náhodný výběr z exponenciálního rozdělení s parametrem $\lambda_Y > 0$. Předpokládejte, že oba náhodné výběry jsou na sobě nezávislé.

- (a) S využitím asymptotického rozdělení $\frac{\bar{Y}_n}{\bar{X}_n}$ sestavte oboustranný intervalový odhad pro $\frac{\lambda_X}{\lambda_Y}$.
- (b) Najděte asymptotické rozdělení náhodné veličiny $\log\left(\frac{\bar{Y}_n}{\bar{X}_n}\right)$ a využijte této znalosti ke konstrukci oboustranného intervalového odhadu pro $\frac{\lambda_X}{\lambda_Y}$.

5. Nechť $(X_1, Y_1)^\top, \dots, (X_n, Y_n)^\top$ je náhodný výběr z dvourozměrného normálního rozdělení s regulární varianční maticí a korelačním koeficientem ρ . Označme

$$\hat{\rho}_n = \frac{\sum_{i=1}^n (X_i - \bar{X}_n)(Y_i - \bar{Y}_n)}{\sqrt{\sum_{i=1}^n (X_i - \bar{X}_n)^2 \sum_{i=1}^n (Y_i - \bar{Y}_n)^2}}$$

odhad korelačního koeficientu. Potom se dá pomocí Δ -metody dokázat, že

$$\sqrt{n}(\hat{\rho}_n - \rho) \xrightarrow{D} \mathbf{N}(0, (1 - \rho^2)^2).$$

- Pro parametr ρ sestavte klasický asymptotický intervalový odhad.
- Pro parametr ρ nalezněte spolehlivostní množinu B_n .
- Najděte transformaci, která stabilizuje asymptotický rozptyl odhadu korelačního koeficientu $\hat{\rho}_n$ a pomocí této transformace sestavte intervalový odhad pro parametr ρ .

6. Mějme náhodný výběr X_1, \dots, X_n z rozdělení s hustotou

$$f_X(x) = \begin{cases} \lambda e^{-\lambda(x-\delta)}, & x \in (\delta, \infty), \\ 0, & \text{jinak,} \end{cases}$$

kde $\delta \in \mathbb{R}$ je neznámý parametr a $\lambda > 0$ je známá konstanta.

- Ověřte, že $Z_n = \min_{1 \leq i \leq n} X_i - \delta$ je pivotální statistika.
 - Pomocí Z_n najděte oboustranný intervalový odhad pro parametr δ .
7. Nechť f_1 a f_2 jsou hustoty rozdělení náhodných veličin W_1 a W_2 , pro které platí $\mathbf{E} W_i = \mu_i$, $\mathbf{Var} W_i = \sigma_i^2$, pro $\mu_i \in \mathbb{R}$ známé a $\sigma_i > 0$ známé, $i = 1, 2$. Uvažujte směs těchto hustot

$$f(x; \theta) = \theta f_1(x) + (1 - \theta) f_2(x)$$

pro $\theta \in (0, 1)$ neznámé. Máme k dispozici náhodný výběr X_1, \dots, X_n pozorování z rozdělení daného hustotou f .

- Najděte odhad $\hat{\theta}_n$ parametru θ pomocí metody momentů. Vyšetřete jeho nestrannost (případně vychýlení) a konzistenci.
 - Spočtěte střední čtvercovou chybu $\hat{\theta}_n$.
 - Najděte asymptotické rozdělení $\hat{\theta}_n$ a z něj odvoďte intervalový odhad parametru θ .
8. Nechť X_1, \dots, X_n je náhodný výběr s hustotou $f(x; b) = xb^{-2} \exp(-x^2/2b^2) \mathbb{I}[x \geq 0]$ pro $b > 0$.
- Nalezněte odhad \hat{b}_n parametru b momentovou metodou.
 - Nalezněte asymptotické rozdělení \hat{b}_n a na jeho základě sestrojte klasický asymptotický intervalový odhad b .

5 EMPIRICKÉ ODHADY

1. Nechť X_1, \dots, X_n je náhodný výběr z exponenciálního rozdělení $\text{Exp}(\lambda)$, $\lambda > 0$. Naším cílem je odhadnout hodnotu distribuční funkce v nějakém daném bodě y , tj. $\theta_X = 1 - e^{-\lambda y}$. Uvažujte následující dva odhady

$$\tilde{\theta}_n = \frac{1}{n} \sum_{i=1}^n \mathbb{I}\{X_i \leq y\} \quad \hat{\theta}_n = 1 - e^{-\hat{\lambda}_n y}, \quad \text{kde } \hat{\lambda}_n = \frac{1}{\bar{X}_n}.$$

Který z odhadů se Vám zda vhodnější a proč?

2. V následující tabulce jsou zachyceny porodní hmotnosti chlapců (naroděných v daném roce v daném regionu).

Hmotnost [kg]	(1.5, 2.0]	(2.0, 2.5]	(2.5, 3.0]	(3.0, 3.5]	(3.5, 4.0]	(4.0, 4.5]	(4.5, 5.0]	(5.0, 5.5]	Σ
Počet	4	101	769	1904	1651	369	37	3	4838

- (a) Bodově i intervalově odhadněte o kolik je větší pravděpodobnost, že se narodí chlapec s hmotností do 4 kg (včetně) než pravděpodobnost že se narodí chlapec s hmotností do 3 kg (včetně).
- (b) Bodově i intervalově odhadněte, kolikrát je větší pravděpodobnost, že se narodí chlapec s hmotností do 4 kg (včetně) než pravděpodobnost že se narodí chlapec s hmotností do 3 kg (včetně).
- (c) Bodově i intervalově odhadněte, o kolik je větší pravděpodobnost, že se narodí chlapec s hmotností do 3 kg (včetně) než pravděpodobnost že se narodí chlapec s hmotností větší než 4 kg.
- (d) Bodově i intervalově odhadněte, kolikrát je větší pravděpodobnost, že se narodí chlapec s hmotností do 3 kg (včetně) než pravděpodobnost že se narodí chlapec s hmotností větší než 4 kg.
3. Nechť X_1, \dots, X_n je náhodný výběr.

- (a) Najděte asymptotické rozdělení třetího centrálního momentu $\hat{\mu}_3 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^3$.
- (b) Na základě znalosti z (a) sestavte intervalový odhad pro μ_3 .
- (c) Najděte asymptotické rozdělení třetího centrálního momentu $\hat{\mu}_3$ za předpokladu, že X_i má normální rozdělení $N(\mu, \sigma^2)$.

Návod. Všimněte si, že v (a) můžete bez újmy na obecnosti předpokládat, že $EX_1 = 0$.

4. Nechť X_1, \dots, X_n je náhodný výběr z normálního rozdělení $N(\mu, \sigma^2)$.

- (a) Najděte asymptotické rozdělení empirického odhadu šikmosti $\hat{\gamma}_3 = \frac{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^3}{\left(\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2\right)^{3/2}}$.
- (b) Zkuste promyslet, jak by se informace z (a) dala využít k testování normality (tj. testování, že náhodný výběr pochází z normálního rozdělení).

Návod. Všimněte si, že v (a) můžete bez újmy na obecnosti předpokládat, že $\mu = 0$ a $\sigma^2 = 1$.

5. Nechť X_1, \dots, X_n je náhodný výběr z rozdělení se spojitou distribuční funkcí F_X . Nechť $u_X(\beta)$ je β -kvantilem rozdělení F_X a $\hat{u}_n(\alpha)$ je empirický (výběrový) α -kvantil.

- (a) Pro $\alpha = 0.25$, $\beta = 0.30$ a $n = 100$ spočítejte (přibližně) pravděpodobnost $P(\hat{u}_n(\alpha) > u_X(\beta))$.
- (b) Jaká bude pravděpodobnost z (a) pro $n = 1000$?

6. Nechť X_1, \dots, X_n je náhodný výběr z rozdělení se spojitou distribuční funkcí F_X . Označme \tilde{m} medián rozdělení F_X . Najděte největší možné přirozené číslo k_1 a nejmenší možné přirozené číslo k_2 takové, že

$$P(X_{(k_1)} > \tilde{m}) \leq \frac{\alpha}{2}, \quad P(X_{(k_2)} < \tilde{m}) \leq \frac{\alpha}{2}.$$

Ukažte, že $(X_{(k_1)}, X_{(k_2)})$ je intervalový odhad pro \tilde{m} se spolehlivostí alespoň $1 - \alpha$.

6 VÝSLEDKY

Tato část je neúplná a může obsahovat překlepy. Budeme rádi, když nás na případné chyby upozorníte.

MNOHOROZMĚRNÉ NORMÁLNÍ ROZDĚLENÍ

1. (a) $E\|\mathbf{X}\|^2 = \sum_{i=1}^n \text{Var } X_i = \text{tr}\Sigma$, kde (Σ) je stopa matice Σ .
 (b) Jsou-li $\lambda_1, \dots, \lambda_n$ vlastní čísla Σ , pak dostaneme $\|\mathbf{X}\|^2 = \sum_{i=1}^n \lambda_i Z_i^2$, kde Z_1, \dots, Z_n jsou nezávislé $N(0, 1)$. Jsou-li λ_i pouze 0 nebo 1, pak $\|\mathbf{X}\|^2 \sim \chi_r^2$, kde r je počet nenulových vlastních čísel a tedy $r = h(\Sigma)$.
2. Například $U = X + Y$ a $V = X - Y$.
3. (b) Pro A musí platit, že $\int_0^A x^2 \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = 1/4$.

KONVERGENCE NÁHODNÝCH VELIČIN

1. Využijeme, že pro distribuční funkci minima platí $F_{Y_n}(x) = 1 - (1 - F(x))^n$. Konvergence se pak ověří z definice.
- 2.-6. Ověří se z definice příslušné konvergence.
7. Ověří se z definice. Je potřeba vyjádřit si distribuční funkci diskrétního rovnoměrného rozdělení.
8. Geometrické rozdělení: Y_n konverguje k exponenciálnímu rozdělení s parametrem λ . Ověří se podobně jako v 7.
9. Konverguje k $\text{Exp}(1)$.
10. Binomické rozdělení konverguje k Poissonovu.
11. Konverguje v distribuci i v pravděpodobnosti k 1.
12. Posloupnost konverguje v distribuci k $N(0, 1)$, v pravděpodobnosti ani s.j. nekonverguje.
13. (a) $M_n \sim \text{Exp}(n)$. (b) $M_n \xrightarrow{P} 0$. (c) $n^\alpha M_n \rightarrow 0$ pro $\alpha \in (0, 1)$ a pro $nM_n \sim \text{Exp}(1)$.
 (d) pro $n \rightarrow \infty$ U_n roste nade všechny meze. $U_n - \log n \rightarrow W$, kde W má tzv. Gumbelovo rozdělení s distribuční funkcí $F(x) = e^{-e^{-x}}$. Tudíž pro n velké $U_n \approx \log n + W$.
14. (a) Normální rozdělení s varianční maticí $\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$.
 (b) V obou případech mají marginály $N(0, 1)$.
 (c) Známe-li pouze marginální rozdělení, nejsme schopni říci nic o sdruženém rozdělení.
 (d) Pro nezávislé složky konvergence po složkách implikuje konvergenci sdruženého rozdělení. Je to vidět např. z tvaru sdružené distribuční funkce pro nezávislé složky.

VLASTNOSTI ODHADŮ

1. plyne z CLV a Cramér-Slutského věty

2. geom. rozdělení:

(a) $T_n = 1/(1 + \bar{X}_n)$, je konzistentní,

$$\sqrt{n}(T_n - p) \xrightarrow{D} \mathbf{N}(0, p^2(1 - p)),$$

(b) U_n je nestranný, není konzistentní,

(c) V_n je nestranný i konzistentní,

(d) nV_n má binomické rozdělení $\text{Bi}(n, p)$; asymptoticky

$$\sqrt{n}(V_n - p) \xrightarrow{D} \mathbf{N}(0, p(1 - p))$$

(e) W_n je konzistentní

3. alternativní rozdělení

(a) není nestranný, je konzistentní

$$\sqrt{n}(U_n - p(1 - p)) \xrightarrow{D} \mathbf{N}(0, p(1 - p)(1 - 2p)^2),$$

přičemž pro $p = 1/2$ máme $\sqrt{n}(U_n - p(1 - p)) \xrightarrow{P} 0$,

(b) je nestranný, není konzistentní

4. useknuté Poissonovo: (a) K výpočtu se využijí vztahy $e^\lambda + e^{-\lambda} - 2 = (1 - e^{-\lambda})(e^\lambda - 1) = 2 \sum_{k=1}^{\infty} \frac{\lambda^{2k}}{(2k)!}$.

(b) Je-li T nestranný, pak musí platit $T(x) = 0$ pro x liché a $T(x) = 2$ pro x sudé. Tudíž odhad z (a) je jediný možný

5. $\text{Bi}(m, p)$: (b) má-li být $T(X_1, \dots, X_n)$ nestranný odhad θ , pak musí být $ET(X_1, \dots, X_n) = \frac{1}{\sqrt{p(1-p)}}$ pro všechna $p \in (0, 1)$. To ale nemůže nastat, protože na levé straně je polynom v p stupně $n \cdot m$.

6. (a) konzistentní, není nestranný, (b) konzistentní a nestranný,

7. (a) je konzistentní ale není nestranný, (b) je nestranný a konzistentní, c) je nestranný, ale není konzistentní (pro každé n má $T_n \sim \text{Exp}(\lambda)$).

8. Gama rozdělení:

$$\sqrt{n} \left[\begin{pmatrix} \hat{a}_n \\ \hat{p}_n \end{pmatrix} - \begin{pmatrix} a \\ p \end{pmatrix} \right] \xrightarrow{D} \mathbf{N}_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \mathbb{V} \right),$$

kde $\mathbb{V} = \mathbb{D} \Sigma \mathbb{D}^T$, přičemž

$$\mathbb{D} = \begin{pmatrix} \frac{a^2}{p} & -\frac{a^3}{p} \\ 2a & -a^2 \end{pmatrix} \quad \text{a} \quad \Sigma = \begin{pmatrix} \frac{p}{a^2} & \frac{2p}{a^3} \\ \frac{2p}{a^3} & \frac{p^2}{a^4} \left(\frac{6}{p} + 2 \right) \end{pmatrix}.$$

9. rovn. rozdělení:

- (a) U_n je nestranný i konzistentní,
 (b) $MSE(U_n) = \frac{\theta^2}{3n}$,
 (c) T_n je konzistentní, není nestranný a $ET_n = \theta \frac{n}{n+1}$, tedy vychýlení je $\frac{-1}{n+1}\theta$,
 (d) $MSE(T_n) = \frac{n\theta^2}{(n+1)^2(n+2)} + \frac{\theta^2}{(n+1)^2} = \frac{2\theta^2}{(n+1)(n+2)}$
 (e) konverguje v distribuci k náhodné veličině s distribuční funkcí $F(x) = e^{z/\theta}I[z < 0] + 1 \cdot I[z \geq 0]$, což je veličina, která má stejné rozdělení jako $-Z$, kde $Z \sim \text{Exp}(1/\theta)$.

10. Posunutý Exp rozdělení:

- (a) $\hat{\delta}_n$ je konzistentní, ale není nestranný – jeho vychýlení je $1/(n\lambda)$
 (b) $MSE(\hat{\delta}_n) = \frac{2}{n^2\lambda^2}$
 (c) není konzistentní, zjevně $\max X_i$ roste nade všechny meze
 (d) není ani nestranný ani konzistentní odhad δ (je to nestranný a konzistentní odhad $\delta + 1/\lambda$)

INTERVALOVÉ ODHADY

1. $\text{Alt}(p)$:

- (a) $\sqrt{n}(\bar{X}_n - p) \xrightarrow{D} \text{N}(0, p(1-p))$ a odtud klasický asymptotický interval spolehlivosti je

$$\left(\bar{X}_n - \frac{u_{1-\alpha/2}}{\sqrt{n}} \sqrt{\bar{X}_n(1-\bar{X}_n)}, \bar{X}_n + \frac{u_{1-\alpha/2}}{\sqrt{n}} \sqrt{\bar{X}_n(1-\bar{X}_n)} \right),$$

případně interval založený na spolehlivostní množině je tvaru

$$\left(\frac{\bar{X}_n + \frac{u_{1-\alpha/2}^2}{2n}}{1 + \frac{u_{1-\alpha/2}^2}{n}} \mp \frac{u_{1-\alpha/2}}{1 + \frac{u_{1-\alpha/2}^2}{n}} \sqrt{\frac{\bar{X}_n(1-\bar{X}_n)}{n} + \frac{u_{1-\alpha/2}^2}{4n^2}} \right)$$

Pravostranný interval:

$$\left(-\infty, \bar{X}_n + \frac{u_{1-\alpha}}{\sqrt{n}} \sqrt{\bar{X}_n(1-\bar{X}_n)} \right)$$

- (b) Jde o transformaci stabilizující rozptyl, tedy

$$\sqrt{n} \left(\arcsin \sqrt{\bar{X}_n} - \arcsin \sqrt{p} \right) \xrightarrow{D} \text{N}\left(0, \frac{1}{4}\right)$$

a odtud oboustranný intervalový odhad je

$$\left(\sin^2 \left(\arcsin \sqrt{\bar{X}_n} - \frac{u_{1-\alpha/2}}{2\sqrt{n}} \right), \sin^2 \left(\arcsin \sqrt{\bar{X}_n} + \frac{u_{1-\alpha/2}}{2\sqrt{n}} \right) \right).$$

Levostranný intervalový odhad je

$$\left(\sin^2 \left(\arcsin \sqrt{\bar{X}_n} - \frac{u_{1-\alpha}}{2\sqrt{n}} \right), \infty \right)$$

2. Exp

- (a) Klasický asymptotický interval: $\left(\frac{1}{\bar{X}_n \left(1 + \frac{u_{1-\alpha/2}}{\sqrt{n}}\right)}, \frac{1}{\bar{X}_n \left(1 - \frac{u_{1-\alpha/2}}{\sqrt{n}}\right)} \right)$,
 Interval založený na spolehlivostní množině: $\left(\frac{1 - \frac{u_{1-\alpha/2}}{\sqrt{n}}}{\bar{X}_n}, \frac{1 + \frac{u_{1-\alpha/2}}{\sqrt{n}}}{\bar{X}_n} \right)$,
- (b) Klasický asymptotický interval: $\left(\frac{1 - \frac{u_{1-\alpha/2}}{\sqrt{n}}}{\bar{X}_n}, \frac{1 + \frac{u_{1-\alpha/2}}{\sqrt{n}}}{\bar{X}_n} \right)$,
 Interval založený na spolehlivostní množině: $\left(\frac{1}{\bar{X}_n \left(1 + \frac{u_{1-\alpha/2}}{\sqrt{n}}\right)}, \frac{1}{\bar{X}_n \left(1 - \frac{u_{1-\alpha/2}}{\sqrt{n}}\right)} \right)$,
- (c) $\left(\frac{1}{\bar{X}_n \exp\left(\frac{u_{1-\alpha/2}}{\sqrt{n}}\right)}, \frac{1}{\bar{X}_n \exp\left(-\frac{u_{1-\alpha/2}}{\sqrt{n}}\right)} \right)$

3. geometrické rozdělení

(a)

$$\sqrt{n}(\hat{p}_n - p) \xrightarrow{D} \mathbf{N}(0, p^2(1-p))$$

a odtud intervalový odhad p je

$$\left(\hat{p}_n - \frac{u_{1-\alpha/2}}{\sqrt{n}} \hat{p}_n \sqrt{1 - \hat{p}_n}, \hat{p}_n + \frac{u_{1-\alpha/2}}{\sqrt{n}} \hat{p}_n \sqrt{1 - \hat{p}_n} \right)$$

(b)

$$\sqrt{n}(\tilde{p}_n - p) \xrightarrow{D} \mathbf{N}(0, p(1-p))$$

$$\left(\tilde{p}_n - \frac{u_{1-\alpha/2}}{\sqrt{n}} \sqrt{\tilde{p}_n(1 - \tilde{p}_n)}, \tilde{p}_n + \frac{u_{1-\alpha/2}}{\sqrt{n}} \sqrt{\tilde{p}_n(1 - \tilde{p}_n)} \right)$$

7. směs dvou hustot:

- (a) $\hat{\theta}_n = \frac{\bar{X}_n - \mu_2}{\mu_1 - \mu_2}$ je nestranný a konzistentní
- (b) $MSE\hat{\theta}_n = \text{Var } \hat{\theta}_n = \frac{\text{Var } X_1}{n(\mu_1 - \mu_2)^2}$, kde $\text{Var } X_1 = \theta\sigma_1^2 + (1 - \theta)\sigma_2^2 + \theta(1 - \theta)(\mu_1 - \mu_2)^2$.
- (c) $\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{D} \mathbf{N}(0, \frac{\text{Var } X_1}{(\mu_1 - \mu_2)^2})$, intervalový odhad

$$\hat{\theta}_n \mp \frac{u_{1-\alpha/2}}{\sqrt{n}} \sqrt{\frac{\bar{X}_n - \mu_2}{(\mu_1 - \mu_2)^3} \sigma_1^2 + \frac{(\mu_1 - \bar{X}_n)}{(\mu_1 - \mu_2)^3} \sigma_2^2 + \frac{(\bar{X}_n - \mu_2)(\mu_1 - \bar{X}_n)}{(\mu_1 - \mu_2)^2}}$$

8. (a) $\hat{b}_n = \sqrt{\frac{2}{\pi}} \bar{X}_n$, (b) $\sqrt{n}(\hat{b}_n - b) \xrightarrow{D} \mathbf{N}(0, \frac{4-\pi}{\pi} b^2)$, intervalový odhad b je

$$\hat{b}_n \mp \frac{u_{1-\alpha/2}}{\sqrt{n}} \hat{b}_n \sqrt{\frac{4}{\pi} - 1}$$

EMPIRICKÉ ODHADY

1. Vhodnější je $\hat{\theta}_n$, protože podíl asymptotických rozptylů je

$$\frac{\text{avar}(\tilde{\theta}_n)}{\text{avar}(\hat{\theta}_n)} = \frac{e^{\lambda y} - 1}{(\lambda y)^2} > 1.$$

2. (a) bodový odhad 0.735, intervalový odhad (0.722, 0.747)
(b) bodový odhad 5.07, intervalový odhad (4.767, 5.368)
(c)
3. (a) $\sqrt{n} (\hat{\mu}_3 - \mu_3) \xrightarrow{D} \mathbf{N}(0, v^2)$, kde $v^2 = \mu_6 - \mu_3^2 - 6 \mu_2 \mu_4 + 9 \mu_2^3$
(c) $\sqrt{n} \hat{\mu}_3 \xrightarrow{D} \mathbf{N}(0, 6 \sigma^6)$
4. (a) $\Phi \left(\frac{k_\alpha - n\beta - 1/2}{\sqrt{n\beta(1-\beta)}} \right)$, kde $k_\alpha = n\alpha$, pokud $n\alpha$ je celé číslo, a $k_\alpha = \lfloor n\alpha \rfloor + 1$ jinak. Pro $n = 100$ vychází 0.115. (b) $2 \cdot 10^{-4}$