Pravděpodobnost a statistika

9. cvičení

(vytvořeno 23. listopadu 2020)

1 Odhady charakteristik rozdělení

Úloha 1. Karel skáče do dálky. Při deseti skocích dosáhl těchto výsledků.

5.86m	5.78m	5.85m	6.10m	5.42m	5.90m	2.50m	5.80m	5.95m	5.84m

1. Určete odhad střední hodnoty pomocí výběrového průměru.
2. Určete odhad střední hodnoty pomocí „uříznutého” výběrového průměru, kdy uřízneme 10% nejnižších a 10% nejvyšších hodnot.
3. Určete odhad střední hodnoty pomocí výběrového mediánu.
4. Určete odhad rozptylu pomocí výběrového rozptylu.
5. Určete odhad rozptylu pomocí vychýleného výběrového rozptylu.

2 Intervalové odhady charakteristik rozdělení

Úloha 2. Náhodný výběr

8.1	7.4	5.2	8.7	6.5	5.7	9.1	6.8	4.6

pochází z (neznámého) normálního rozdělení $N(\mu, \sigma^2)$.

1. Určete intervalový odhad střední hodnoty μ se spolehlivostí $1 - \alpha = 0.95$.
2. Určete intervalový odhad rozptylu σ^2 se spolehlivostí $1 - \beta = 0.95$.

Řešení

Řešení 1:

i	x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
1.	6.10	0.60	0.3600
2.	5.95	0.45	0.2025
3.	5.90	0.40	0.1600
4.	5.86	0.36	0.1296
5.	5.85	0.35	0.1225
6.	5.84	0.34	0.1156
7.	5.80	0.30	0.0900
8.	5.78	0.28	0.0784
9.	5.42	-0.08	0.0064
10.	2.50	-3.00	9.0000
Σ	55.00		10.2650
$\frac{1}{n}\Sigma$	5.50		1.0265
$\frac{1}{n-1}\Sigma$			1.1406

1. realizace výběrového průměru: $\bar{x} = \frac{55\text{m}}{10} = 5.5\text{m}$
2. realizace uříznutého výběrového průměru: $\frac{46.4\text{m}}{8} = 5.8\text{m}$
3. realizace výběrového mediánu: 5.845m
4. realizace výběrového rozptylu: $s_{\mathbf{x}}^2 = 1.1406\text{m}^2$

- realizace výběrové směrodatné odchylky: $s_{\mathbf{x}} = \sqrt{1.1406}\text{m} = 1.0680\text{m}$

5. realizace vychýleného výběrového rozptylu: $\hat{\sigma}^2_{\mathbf{x}} = 1.0265\text{m}^2$

Řešení 2:

i	x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
1.	8.1	1.20	1.44
2.	7.4	0.50	0.25
3.	5.2	-1.70	2.89
4.	8.7	1.80	3.24
5.	6.5	-0.40	0.16
6.	5.7	-1.20	1.44
7.	9.1	2.20	4.84
8.	6.8	-0.10	0.01
9.	4.6	-2.30	5.29
Σ	62.10		19.56
$\frac{1}{n}\Sigma$	6.90		2.1733
$\frac{1}{n-1}\Sigma$			2.445

- X ... rozdělení $N(\mu, \sigma^2)$
- výběrový průměr: $\bar{x} = 6.9$
- $\mathrm{D}\bar{X}_n = \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{D}X = \frac{1}{n}\mathrm{D}X$
- výběrový rozptyl: $s^2_X = 2.445$
- výběrová směrodatná odchylka: $s_X = 1.5636$

1. Symetrický intervalový odhad střední hodnoty μ se spolehlivostí $1 - \alpha = 0.95$:
   - Hledáme hodnotu $\varepsilon \in \mathbb{R}$ takovou, že:

$$\begin{aligned}
P\left[\mu - \varepsilon \leq \bar{X} \leq \mu + \varepsilon\right] &\geq 1 - \alpha \\
P\left[-\varepsilon \leq \bar{X} - \mu \leq \varepsilon\right] &\geq 1 - \alpha \\
P\left[-\frac{\varepsilon}{\sqrt{\mathrm{D}\bar{X}}} \leq \frac{\bar{X} - \mu}{\sqrt{\mathrm{D}\bar{X}}} \leq \frac{\varepsilon}{\sqrt{\mathrm{D}\bar{X}}}\right] &\geq 1 - \alpha \\
P\left[-\frac{\varepsilon}{\sqrt{\frac{\mathrm{D}X}{n}}} \leq \frac{\bar{X} - \mu}{\sqrt{\frac{\mathrm{D}X}{n}}} \leq \frac{\varepsilon}{\sqrt{\frac{\mathrm{D}X}{n}}}\right] &\geq 1 - \alpha
\end{aligned}$$

   - V případě, že přímo známe hodnotu $\mathrm{D}X$ (a můžeme ji tedy považovat za reálnou konstantu), má náhodná veličina $\frac{\bar{X}-\mu}{\sqrt{\frac{\mathrm{D}X}{n}}}$ normované normální rozdělení.
   - V případě, že hodnotu $\mathrm{D}X$ neznáme (což je náš případ), odhadneme ji pomocí výběrového rozptylu S^2_X, který má (po vynásobení patřičnými hodnotami) rozdělení χ^2 s $(n-1)$ stupni volnosti. V takovém případě má náhodná veličina $\frac{\bar{X}-\mu}{\sqrt{\frac{S^2_X}{n}}}$ Studentovo rozdělení s $(n-1)$ stupni volnosti ... $t(n-1)$.

$$\begin{aligned}
P\left[-\frac{\varepsilon}{\frac{s_X}{\sqrt{n}}} \leq \frac{\bar{X} - \mu}{\sqrt{\frac{S^2_X}{n}}} \leq \frac{\varepsilon}{\frac{s_X}{\sqrt{n}}}\right] &\geq 1 - \alpha \\
F_{t(n-1)}\left(\frac{\varepsilon\sqrt{n}}{s_X}\right) - F_{t(n-1)}\left(-\frac{\varepsilon\sqrt{n}}{s_X}\right) &\geq 1 - \alpha \\
F_{t(n-1)}\left(\frac{\varepsilon\sqrt{n}}{s_X}\right) - \left[1 - F_{t(n-1)}\left(\frac{\varepsilon\sqrt{n}}{s_X}\right)\right] &\geq 1 - \alpha \\
2 \cdot F_{t(n-1)}\left(\frac{\varepsilon\sqrt{n}}{s_X}\right) - 1 &\geq 1 - \alpha
\end{aligned}$$

$$\begin{aligned}
F_{t(n-1)}\left(\frac{\varepsilon\sqrt{n}}{s_X}\right) &\geq 1-\frac{\alpha}{2} \\
\frac{\varepsilon\sqrt{n}}{s_X} &\geq q_{t(n-1)}\left(1-\frac{\alpha}{2}\right) \\
\varepsilon &\geq \frac{s_X}{\sqrt{n}}\cdot q_{t(n-1)}\left(1-\frac{\alpha}{2}\right) \\
\varepsilon &\geq \frac{1.5636}{\sqrt{9}}\cdot q_{t(8)}\left(0.975\right) \\
\varepsilon &\geq \frac{1.5636}{3}\cdot 2.31 \\
\varepsilon &\geq 3.7259
\end{aligned}$$

- Symetrický intervalový odhad μ je:

$$\mu \in \langle 6.9-3.7259, 6.9+3.7259\rangle = \langle 3.1741, 10.626\rangle .$$

- Případně můžeme ze znalosti, že náhodná veličina $\frac{\bar{X}-\mu}{\sqrt{\frac{S_X^2}{n}}}$ má Studentovo rozdělení s $(n-1)$ stupni volnosti, výsledek určit i takto:

$$\begin{aligned}
P\left[q_{t(n-1)}\left(\frac{\alpha}{2}\right)\leq\frac{\bar{X}-\mu}{\sqrt{\frac{S_X^2}{n}}}\leq q_{t(n-1)}\left(1-\frac{\alpha}{2}\right)\right] &\geq 1-\alpha \\
P\left[-q_{t(n-1)}\left(1-\frac{\alpha}{2}\right)\leq\frac{\bar{X}-\mu}{\frac{S_X}{\sqrt{n}}}\leq q_{t(n-1)}\left(1-\frac{\alpha}{2}\right)\right] &\geq 1-\alpha \\
P\left[-\frac{S_X}{\sqrt{n}}\,q_{t(n-1)}\left(1-\frac{\alpha}{2}\right)\leq\bar{X}-\mu\leq\frac{S_X}{\sqrt{n}}\,q_{t(n-1)}\left(1-\frac{\alpha}{2}\right)\right] &\geq 1-\alpha \\
P\left[\bar{X}-\frac{S_X}{\sqrt{n}}\,q_{t(n-1)}\left(1-\frac{\alpha}{2}\right)\leq\mu\leq\bar{X}+\frac{S_X}{\sqrt{n}}\,q_{t(n-1)}\left(1-\frac{\alpha}{2}\right)\right] &\geq 1-\alpha
\end{aligned}$$

- Horní intervalový odhad střední hodnoty μ se spolehlivostí $1-\alpha=0.95$:
  - Hledáme hodnotu $\varepsilon\in\mathbb{R}$ takovou, že:

$$\begin{aligned}
P\left[\bar{X}\leq\mu+\varepsilon\right] &\geq 1-\alpha \\
P\left[\bar{X}-\mu\leq\varepsilon\right] &\geq 1-\alpha \\
P\left[\frac{\bar{X}-\mu}{\sqrt{\mathrm{D}\bar{X}}}\leq\frac{\varepsilon}{\sqrt{\mathrm{D}\bar{X}}}\right] &\geq 1-\alpha \\
P\left[\frac{\bar{X}-\mu}{\sqrt{\frac{\mathrm{D}X}{n}}}\leq\frac{\varepsilon}{\sqrt{\frac{\mathrm{D}X}{n}}}\right] &\geq 1-\alpha \\
P\left[\frac{\bar{X}-\mu}{\sqrt{\frac{S_X^2}{n}}}\leq\frac{\varepsilon}{\frac{s_X}{\sqrt{n}}}\right] &\geq 1-\alpha \\
F_{t(n-1)}\left(\frac{\varepsilon\sqrt{n}}{s_X}\right) &\geq 1-\alpha \\
\frac{\varepsilon\sqrt{n}}{s_X} &\geq q_{t(n-1)}\left(1-\alpha\right) \\
\varepsilon &\geq \frac{s_X}{\sqrt{n}}\cdot q_{t(n-1)}\left(1-\alpha\right) \\
\varepsilon &\geq \frac{1.5636}{\sqrt{9}}\cdot q_{t(8)}\left(0.95\right) \\
\varepsilon &\geq \frac{1.5636}{3}\cdot 1.86 \\
\varepsilon &\geq 0.9694
\end{aligned}$$

– Horní intervalový odhad μ je:

$$\mu \in (-\infty, 6.9 + 0.9694\rangle = (-\infty, 7.8694\rangle \, .$$

– Podobně bychom určili dolní intervalový odhad μ, který by vyšel:

$$\mu \in \langle 6.9 - 0.9694, \infty) = \langle 5.9306, \infty) \, .$$

- V případě, že známe hodnotu $\mathrm{D}X$:

  – Symetrický intervalový odhad:

$$P\left[-\frac{\varepsilon}{\sqrt{\frac{\mathrm{D}X}{n}}} \leq \frac{\bar{X}-\mu}{\sqrt{\frac{\mathrm{D}X}{n}}} \leq \frac{\varepsilon}{\sqrt{\frac{\mathrm{D}X}{n}}}\right] = 2\Phi\left(\frac{\varepsilon}{\sqrt{\frac{\mathrm{D}X}{n}}}\right) - 1 \geq 1-\alpha$$

$$\varepsilon \geq \sqrt{\frac{\mathrm{D}X}{n}}\,\Phi^{-1}(1-\frac{\alpha}{2})$$

$$\mu \in \left\langle \bar{x} - \sqrt{\frac{\mathrm{D}X}{n}}\,\Phi^{-1}(1-\frac{\alpha}{2}, \bar{x} + \sqrt{\frac{\mathrm{D}X}{n}}\,\Phi^{-1}(1-\frac{\alpha}{2}\right\rangle$$

  – Horní intervalový odhad:

$$P\left[\frac{\bar{X}-\mu}{\sqrt{\frac{\mathrm{D}X}{n}}} \leq \frac{\varepsilon}{\sqrt{\frac{\mathrm{D}X}{n}}}\right] = \Phi\left(\frac{\varepsilon}{\sqrt{\frac{\mathrm{D}X}{n}}}\right) \geq 1-\alpha$$

$$\varepsilon \geq \sqrt{\frac{\mathrm{D}X}{n}}\,\Phi^{-1}(1-\alpha)$$

$$\mu \in \left(-\infty, \bar{x} + \sqrt{\frac{\mathrm{D}X}{n}}\,\Phi^{-1}(1-\alpha)\right\rangle$$

  – Dolní intervalový odhad:

$$\mu \in \left\langle \bar{x} - \sqrt{\frac{\mathrm{D}X}{n}}\,\Phi^{-1}(1-\alpha), \infty\right)$$

2. Pro intervalový odhad rozptylu σ^2 využijeme znalost, že náhodná veličina $\frac{n-1}{\sigma^2}S_X^2$ má rozdělení χ^2 s $(n-1)$ stupni volnosti, protože:

$$\begin{aligned}
S_X^2 &= \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 \\
S_X^2 &= \frac{\sigma^2}{n-1}\sum_{i=1}^{n}\left(\frac{X_i - \bar{X}}{\sigma}\right)^2 \\
\frac{n-1}{\sigma^2}S_X^2 &= \underbrace{\sum_{i=1}^{n}\left(\mathrm{norm}X_i\right)^2}_{\chi^2(n-1)}
\end{aligned}$$

- Symetrický intervalový odhad rozptylu se spolehlivostí $1-\beta = 0.95$:

$$\begin{aligned}
P\left[q_{\chi^2(n-1)}\left(\frac{\beta}{2}\right) \leq \frac{n-1}{\sigma^2}S_X^2 \leq q_{\chi^2(n-1)}\left(1-\frac{\beta}{2}\right)\right] &= 1-\beta \\
P\left[\frac{(n-1)S_X^2}{q_{\chi^2(n-1)}\left(\frac{\beta}{2}\right)} \geq \sigma^2 \geq \frac{(n-1)S_X^2}{q_{\chi^2(n-1)}\left(1-\frac{\beta}{2}\right)}\right] &= 1-\beta \\
P\left[\frac{8\cdot 2.445}{q_{\chi^2(8)}(0.025)} \geq \sigma^2 \geq \frac{8\cdot 2.445}{q_{\chi^2(8)}(0.975)}\right] &= 0.95 \\
P\left[\frac{8\cdot 2.445}{2.18} \geq \sigma^2 \geq \frac{8\cdot 2.445}{17.53}\right] &= 0.95 \\
P\left[8.9725 \geq \sigma^2 \geq 1.1158\right] &= 0.95
\end{aligned}$$

$$\sigma^2 \in \langle 1.1158, 8.9725\rangle$$

- Horní intervalový odhad rozptylu se spolehlivostí $1-\beta = 0.95$:

$$\begin{aligned}
P\left[q_{\chi^2(n-1)}(\beta) \leq \frac{n-1}{\sigma^2}S_X^2\right] &= 1-\beta \\
P\left[\frac{(n-1)S_X^2}{q_{\chi^2(n-1)}(\beta)} \geq \sigma^2\right] &= 1-\beta \\
P\left[\frac{8\cdot 2.445}{q_{\chi^2(8)}(0.05)} \geq \sigma^2\right] &= 0.95 \\
P\left[\frac{8\cdot 2.445}{2.73} \geq \sigma^2\right] &= 0.95 \\
P\left[7.1648 \geq \sigma^2\right] &= 0.95
\end{aligned}$$

$$\sigma^2 \in (-\infty, 7.1648\rangle$$

- Dolní intervalový odhad rozptylu se spolehlivostí $1-\beta = 0.95$:

$$\begin{aligned}
P\left[\frac{n-1}{\sigma^2}S_X^2 \leq q_{\chi^2(n-1)}(1-\beta)\right] &= 1-\beta \\
P\left[\sigma^2 \geq \frac{(n-1)S_X^2}{q_{\chi^2(n-1)}(1-\beta)}\right] &= 1-\beta \\
P\left[\sigma^2 \geq \frac{8\cdot 2.445}{q_{\chi^2(8)}(0.95)}\right] &= 0.95 \\
P\left[\sigma^2 \geq \frac{8\cdot 2.445}{15.51}\right] &= 0.95 \\
P\left[\sigma^2 \geq 1.2611\right] &= 0.95
\end{aligned}$$

$$\sigma^2 \in \langle 1.2611, \infty)$$