Pravděpodobnost a statistika

10. cvičení

(vytvořeno 1. prosince 2020)

1 Odhady parametrů rozdělení

Úloha 1. Náhodná veličina X může nabývat hodnot 0, 1 a 2. Její parametrické pravděpodobnostní rozdělení a naměřené četnosti hodnot udává následující tabulka.

hodnota	0	1	2
teoretická pravděpodobnost	a	b	$a+b$
naměřená četnost	13	16	21

Odhadněte parametry $a, b \in \mathbb{R}$ pomocí:

1. metody momentů,
2. metody maximální věrohodnosti.

Úloha 2. Náhodná veličina X má geometrické rozdělení. Následující tabulka udává naměřené četnosti hodnot.

hodnota	0	1	2	3
naměřená četnost	29	16	4	1

Odhadněte parametry tohoto geometrického rozdělení pomocí:

1. metody momentů,
2. metody maximální věrohodnosti.

Úloha 3. V osudí máme správné kostky, na kterých padají všechna čísla se stejnou pravděpodobností, a také falešné, na kterých padá šestka s pravděpodobností 0.5 a ostatní hodnoty s pravděpodobností 0.1. Stokrát jsme náhodně vytáhli kostkou, hodili jí a vrátili zpět. Následující tabulka udává naměřené četnosti hodnot.

hodnota	1	2	3	4	5	6
naměřená četnost	11	10	14	10	18	37

Odhadněte, jaké je zastoupení falešných kostek, pomocí:

1. metody momentů,
2. metody maximální věrohodnosti.

Úloha 4. Náhodné veličiny X a Y mají následující pravděpodobnostní rozdělení.

x	1	2	3
$P[X=x]$	0.5	0.2	0.3
$P[Y=x]$	0.2	0.4	0.4

Náhodná veličina Z je dána směsí $Z = \mathrm{Mix}_{(w,1-w)}(X,Y)$. Určete parametr směsi $w \in \langle 0.1 \rangle$ podle následující tabulky, která udává naměřené četnosti hodnot náhodné veličiny Z.

hodnota	1	2	3
naměřená četnost	24	42	34

Úloha 5 (převzato od M. Korbeláře)**.** Náhodná veličina X má rovnoměrné rozdělení na intervalu $\langle -h, h \rangle$, kde $h \in (0, \infty)$. Provedli jsme osm pokusů, při kterých jsme naměřili hodnoty:

-4	-3	-2	-1.5	0.5	1	2.5	3

Odhadněte parametr h tohoto rozdělení pomocí:

1. metody momentů,
2. metody maximální věrohodnosti.

Řešení 5:

- Vzhledem k naměřeným hodnotám musí být hodnota h větší nebo rovna 4.

1. Metoda momentů: Protože je hustota rozdělení sudá (symetrická kolem osy y), nelze první moment (střední hodnotu) použít. U tohoto rozdělení není významné, kde je jeho střed, ale jak moc je kolem středu „rozptýlené". Pro odhad tedy použijeme druhý moment:

$$\begin{aligned}
m_2 &= \frac{1}{n}\sum_{i=1}^{n} x_i^2 = \frac{47.75}{8} = 5.96875 \\
\mathrm{E}\left(X^2\right) &= \int_{\infty}^{\infty} x^2 \cdot f_X(x)\,\mathrm{d}x = \int_{-h}^{h} x^2 \cdot \frac{1}{2h}\,\mathrm{d}x = \frac{1}{6h}\left[x^3\right]_{-h}^{h} = \frac{h^2}{3} \\
\frac{h^2}{3} &= \frac{47.75}{8} \\
h &= \sqrt{\frac{3\cdot 47.75}{8}} = \sqrt{17.90625} = 4.2316
\end{aligned}$$

2. Metoda maximální věrohodnosti:

$$\begin{aligned}
\Lambda(p) &= \prod_{i=1} 5(p+1)(p+2)\cdot x_i\,(1-x_i)^p \\
&= (p+1)^5(p+2)^5\cdot 0.20\cdot 0.80^p\cdot 0.20\cdot 0.80^p\cdot 0.25\cdot 0.75^p\cdot 0.50\cdot 0.50^p\cdot 0.85\cdot 0.15^p \\
&= (p+1)^5(p+2)^5\cdot 0.00425\cdot 0.036^p \\
\lambda(p) &= 5\,\ln(p+1) + 5\,\ln(p+2) + \ln(0.00425) + p\,\ln(0.036) \\
\lambda'(p) &= \frac{5}{p+1} + \frac{5}{p+2} + \ln(0.036) = \frac{5}{p+1} + \frac{5}{p+2} - 3.3242
\end{aligned}$$

$$\begin{aligned}
\frac{5}{p+1} + \frac{5}{p+2} - 3.3242 &= 0 \\
5p + 10 + 5p + 5 - 3.3242\,(p^2+3p+2) &= 0 \\
-3.3242\,p^2 + 0.027291\,p + 8.3515 &= 0
\end{aligned}$$

Řešení kvadratické rovnice: $p \in \{-1.5809, 1.5891\}$; z toho vyhovuje pouze $p \doteq 1.5891$.

Úloha 6. Náhodná veličina X má exponenciální rozdělení. Provedli jsme pět pokusů, při kterých jsme naměřili hodnoty:

0.87	1.28	0.9	0.37	0.25

Odhadněte parametry tohoto exponenciálního rozdělení pomocí:

1. metody momentů,
2. metody maximální věrohodnosti.

Úloha 7 (převzato od M. Korbeláře)**.** Náhodná veličina X má pro parametr $a \in \mathbb{R}$ hustotu:

$$f_X(x) = \begin{cases} 0 & \text{pokud } x \in (-\infty, a)\,, \\ \mathrm{e}^{a-x} & \text{pokud } x \in \langle a, \infty)\,. \end{cases}$$

Provedli jsme sedm pokusů, při kterých jsme naměřili hodnoty:

1	2	2	2	3	3	4

Odhadněte parametr a tohoto rozdělení pomocí:

1. metody momentů,
2. metody maximální věrohodnosti.

Úloha 8. Náhodná veličina X má pro parametr $p \in (0, \infty)$ hustotu:

$$f_X(x) \;=\; \begin{cases} (p+1)(p+2)\cdot x\,(1-x)^p & \text{pokud } \; x \in \langle 0,1\rangle\,, \\ 0 & \text{jinak.} \end{cases}$$

Provedli jsme pět pokusů, při kterých jsme naměřili hodnoty:

0.20	0.20	0.25	0.50	0.85

Odhadněte parametr p tohoto rozdělení pomocí:

1. metody momentů,
2. metody maximální věrohodnosti.

Řešení

Řešení 1: Z rovnice $a+b+(a+b) = 1$ dostaneme $b = \frac{1}{2} - a$ a $(a+b) = \frac{1}{2}$. Stačí tedy odhadnout pouze parametr a.

hodnota	0	1	2
teoretická pravděpodobnost	a	$\frac{1}{2} - a$	$\frac{1}{2}$
naměřená četnost	13	16	21

1. Metoda momentů:

$$\begin{aligned} \mathrm{E}X &= \frac{3}{2} - a \\ \bar{x} &= \frac{58}{50} \\ \frac{3}{2} - a &= \frac{58}{50} \\ a &= \frac{17}{50} \end{aligned}$$

2. Metoda maximální věrohodnosti:

$$\begin{aligned} L(a) &= a^{13} \cdot \left(\frac{1}{2} - a\right)^{16} \cdot \left(\frac{1}{2}\right)^{21} \\ l(a) &= 13\ln a + 16 \ln\left(\frac{1}{2} - a\right) + 21 \ln\left(\frac{1}{2}\right) \\ l'(a) &= \frac{13}{a} + \frac{-16}{\frac{1}{2} - a} \;=\; \frac{13 - 58a}{a(1-2a)} \\ \frac{13-58a}{a(1-2a)} &= 0 \\ a &= \frac{13}{58} \end{aligned}$$

Řešení 2: Geometrické rozdělení

- $P\,[X = x] = q^x(1-q)$
- $\mathrm{E}X = \frac{q}{1-q}$
- $\mathrm{D}X = \frac{q}{(1-q)^2}$

1. Metoda momentů:

$$\begin{aligned} \bar{x} &= \frac{27}{50} \\ \frac{q}{1-q} &= \frac{27}{50} \\ q &= \frac{27}{77} \end{aligned}$$

2. Metoda maximální věrohodnosti:

$$\begin{aligned} L(q) &= \Big(q^0(1-q)\Big)^{29} \cdot \Big(q^1(1-q)\Big)^{16} \cdot \Big(q^2(1-q)\Big)^{4} \cdot \Big(q^3(1-q)\Big)^{1} \\ &= q^{27}(1-q)^{50} \\ l(q) &= 27 \ln q + 50 \ln(1-q) \\ l'(q) &= \frac{27}{q} + \frac{-50}{1-q} \quad = \quad \frac{27-77q}{q(1-2q)} \\ \frac{27-77q}{q(1-2q)} &= 0 \\ q &= \frac{27}{77} \end{aligned}$$

Řešení 3:

- S ... správná kostka
- F ... falešná kostka
- $X = \mathrm{Mix}_{(1-w,w)}(S, F)$

x	$1 \ldots 5$	6
$P[S=x]$	$\frac{1}{6}$	$\frac{1}{6}$
$P[F=x]$	$\frac{1}{10}$	$\frac{1}{2}$
$P[X=x]$	$(1-w)\frac{1}{6} + w\frac{1}{10} = \frac{5-2w}{30}$	$(1-w)\frac{1}{6} + w\frac{1}{2} = \frac{1+2w}{6}$

1. Metoda momentů:

$$\begin{aligned} \mathrm{E}S &= \frac{7}{2} \\ \mathrm{E}F &= \frac{9}{2} \\ \mathrm{E}X &= (1-w)\frac{7}{2} + w\frac{9}{2} \quad = \quad w + \frac{7}{2} \\ \bar{x} &= \frac{425}{100} \\ w + \frac{7}{2} &= \frac{425}{100} \\ w &= \frac{75}{100} \end{aligned}$$

V osudí je 75% falešných kostek.

2. Metoda maximální věrohodnosti:

$$\begin{aligned}
L(q) &= \left(\frac{5-2w}{30}\right)^{11+10+14+10+18} \cdot \left(\frac{1+2w}{6}\right)^{37} \\
&= \left(\frac{5-2w}{30}\right)^{63} \cdot \left(\frac{1+2w}{6}\right)^{37} \\
l(q) &= 63\ln(5-2w) - 63\ln 30 + 37\ln(1+2w) - 37\ln 6 \\
l'(q) &= \frac{-2\cdot 63}{5-2w} + \frac{2\cdot 37}{1+2w} \quad = \quad \frac{244-400w}{(5-2w)(1+2w)} \\
\frac{244-400w}{(5-2w)(1+2w)} &= 0 \\
w &= \frac{244}{400} \quad = \quad 0.61
\end{aligned}$$

V osudí je 61% falešných kostek.

Řešení 4: Určíme parametrické pravděpodobnostní rozdělení náhodné veličiny Z:

x	1	2	3
$P[Z=x]$	$0.5w+0.2(1-w)$ $=0.3w+0.2$	$0.2w+0.4(1-w)$ $=-0.2w+0.4$	$0.3w+0.4(1-w)$ $=-0.3w+0.6$

1. Metoda momentů:

$$\begin{aligned}
\mathrm{E}Z &= 1\cdot(0.3w+0.2) + 2\cdot(-0.2w+0.4) + 3\cdot(-0.3w+0.6) \\
&= -0.4w+2.2 \\
\bar{Z} &= 2.1 \\
-0.4w+2.2 &= 2.1 \\
w &= 0.25 \\
Z &= \mathrm{Mix}_{(0.25;0.75)}(X,Y)
\end{aligned}$$

2. Metoda maximální věrohodnosti:

$$\begin{aligned}
L(w) &= (0.3w+0.2)^{24} \cdot (-0.2w+0.4)^{42} \cdot (-0.3w+0.6)^{34} \\
l(w) &= 24\ln(0.3w+0.2) + 42\ln(-0.2w+0.4) \\
&\quad +34\ln(-0.3w+0.6) \\
l'(w) &= \frac{0.3\cdot 24}{0.3w+0.2} + \frac{0.2\cdot 42}{-0.2w+0.4} + \frac{-0.3\cdot 34}{-0.3w+0.6} \\
&= \frac{300w+8}{(3w+2)(w-2)} \\
\frac{300w+8}{(3w+2)(w-2)} &= 0 \\
w &= -\frac{8}{300}
\end{aligned}$$

Metoda maximální věrohodnosti v tomto případě nedává smysluplný výsledek.

Řešení 6: Exponenciální rozdělení

- hustota pravděpodobnosti: $f_X(x) = \frac{1}{\tau}\mathrm{e}^{\frac{-x}{\tau}}$ pro $x>0$, kde $\tau \in (0,\infty)$
- distribuční funkce: $F_X(x) = 1-\exp\left(-\frac{x}{\tau}\right)$ pro $x>0$
- $\mathrm{E}X = \tau$
- $\mathrm{D}X = \tau^2$

1. Metoda momentů:

$$\bar{x} = \frac{3.67}{5} = 0.734$$
$$\tau = 0.734$$

2. Metoda maximální věrohodnosti:

$$\begin{aligned}
\Lambda(\tau) &= \left(\frac{1}{\tau}\mathrm{e}^{\frac{-0.87}{\tau}}\right)\cdot\left(\frac{1}{\tau}\mathrm{e}^{\frac{-1.28}{\tau}}\right)\cdot\left(\frac{1}{\tau}\mathrm{e}^{\frac{-0.9}{\tau}}\right)\cdot\left(\frac{1}{\tau}\mathrm{e}^{\frac{-0.37}{\tau}}\right)\cdot\left(\frac{1}{\tau}\mathrm{e}^{\frac{-0.25}{\tau}}\right)\\
&= \left(\frac{1}{\tau}\right)^5 \mathrm{e}^{\frac{-3.67}{\tau}}\\
\lambda(\tau) &= -5\ln\tau - \frac{3.67}{\tau}\\
\lambda'(\tau) &= -\frac{5}{\tau} + \frac{3.67}{\tau^2} = \frac{3.67-5\tau}{\tau^2}\\
\tau &= \frac{3.67}{5} = 0.734
\end{aligned}$$

Řešení 7:

- Protože je hustota na $(-\infty, a)$ nulová, musí být parametr a menší nebo roven nejmenší naměřené hodnotě, což je 1.

1. Metoda momentů:

$$\begin{aligned}
\bar{x} &= \frac{17}{7}\\
\mathrm{E}X &= \int_{\infty}^{\infty} x\cdot f_X(x)\,\mathrm{d}x = \int_a^{\infty} x\cdot \mathrm{e}^{a-x}\,\mathrm{d}x = \left[-x\cdot\mathrm{e}^{a-x} - \mathrm{e}^{a-x}\right]_a^{\infty} = a+1\\
a+1 &= \frac{17}{7}\\
a &= \frac{10}{7}
\end{aligned}$$

Metoda momentů v tomto případě nedává smysluplný výsledek.

2. Metoda maximální věrohodnosti:

$$\begin{aligned}
\Lambda(a) &= \mathrm{e}^{a-1}\cdot\mathrm{e}^{a-2}\cdot\mathrm{e}^{a-2}\cdot\mathrm{e}^{a-2}\cdot\mathrm{e}^{a-3}\cdot\mathrm{e}^{a-3}\cdot\mathrm{e}^{a-4} \quad \text{pro } a\in(-\infty,1\rangle\,,\\
a &= 1
\end{aligned}$$

Řešení 8:

1. Metoda momentů:

$$\begin{aligned}
\bar{x} &= \frac{2.00}{5} = 0.40\\
\mathrm{E}X &= \int_{\infty}^{\infty} x\cdot f_X(x)\,\mathrm{d}x = (p+1)(p+2)\int_0^1 x^2\,(1-x)^p\,\mathrm{d}x\\
&= (p+1)(p+2)\left[-\frac{x^2\,(1-x)^{p+1}}{p+1} - \frac{2x\,(1-x)^{p+2}}{(p+1)(p+2)} - \frac{2\,(1-x)^{p+3}}{(p+1)(p+2)(p+3)}\right]_0^1\\
&= \frac{2}{p+3}\\
\frac{2}{p+3} &= 0.40\\
p &= 2
\end{aligned}$$

2. Metoda maximální věrohodnosti:

$$\begin{aligned}
\Lambda(p) &= \prod_{i=1} 5(p+1)(p+2)\cdot x_i\,(1-x_i)^p \\
&= (p+1)^5(p+2)^5\cdot 0.20\cdot 0.80^p\cdot 0.20\cdot 0.80^p\cdot 0.25\cdot 0.75^p\cdot 0.50\cdot 0.50^p\cdot 0.85\cdot 0.15^p \\
&= (p+1)^5(p+2)^5\cdot 0.00425\cdot 0.036^p \\
\lambda(p) &= 5\,\ln(p+1)+5\,\ln(p+2)+\ln(0.00425)+p\,\ln(0.036) \\
\lambda'(p) &= \frac{5}{p+1}+\frac{5}{p+2}+\ln(0.036) \;=\; \frac{5}{p+1}+\frac{5}{p+2}-3.3242
\end{aligned}$$

$$\begin{aligned}
\frac{5}{p+1}+\frac{5}{p+2}-3.3242 &= 0 \\
5p+10+5p+5-3.3242\,(p^2+3p+2) &= 0 \\
-3.3242\,p^2+0.027291\,p+8.3515 &= 0
\end{aligned}$$

Řešení kvadratické rovnice: $p\in\{-1.5809, 1.5891\}$; z toho vyhovuje pouze $p \doteq 1.5891$.