

10 Odhady, Testovanie hypotéz

Príklad 10.1. Nech $X_1, \dots, X_n$ je náhodný výber z $N(0, \sigma^2)$, pričom σ^2 nepoznáme. Sformulujte test pre

$$H_0 : \sigma^2 = \sigma_0^2 \text{ vs. } H_1 : \sigma^2 = \sigma_1^2, \quad \text{kde } \sigma_0^2 < \sigma_1^2,$$

na hladine α , ktorý má medzi všetkými testami pre tieto hypotézy na hladine α najmenšiu pravdepodobnosť chyby 2. druhu.

Príklad 10.2. Nech $X_1, \dots, X_n$ je náhodný výber z rozdelenia $\text{Bin}(m, p)$, kde m je známe a $p \in (0, 1)$ je neznámy parameter. a) Ukážte, že $\hat{p}_n = \frac{1}{nm} \sum_{i=1}^n X_i$ je odhad parametra p pomocou metódy maximálnej vierohodnosti.

b) Slovenské národné divadlo predáva na jedno predstavenie 100 lístkov. Počas poslednej divadelnej sezóny sa mu podarilo vypredať všetky predstavenia. Divák, ktorý si kúpil lístok, príde na predstavenie s pravdepodobnosťou $p \in (0, 1)$ bez ohľadu na predstavenie a diváci sa rozhodujú nezávisle. Šatniari spočítali na ôsmich predstaveniach, koľko divákov naozaj prišlo. Výsledky sú dané tabuľkou

predstavenie	1	2	3	4	5	6	7	8
# divákov	93	94	89	91	90	91	95	90

Odhadnite pravdepodobnosť p , že divák s lístkom príde na predstavenie SND, pomocou metódy maximálnej vierohodnosti.

Príklad 10.3. Nech $X_1, \dots, X_n$ je náhodný výber z rozdelenia $\text{Exp}(\lambda)$, kde $\lambda > 0$ je neznámy parameter, t.j. X_i majú hustotu $\lambda e^{-\lambda x}$ pre $x > 0$. Napíšte predpis pre odhad $\hat{\lambda}_n$ parametra λ pomocou metódy maximálnej vierohodnosti ako funkciu premenných $X_1, \dots, X_n$.

Príklad 10.4. Nech $X_1, \dots, X_n$ je náhodný výber z rozdelenia $N(\mu, \sigma^2)$, kde μ a $\sigma^2 > 0$ sú neznáme parametre. a) Nájdite odhad parametrov μ a σ^2 pomocou metódy maximálnej vierohodnosti ako funkciu premenných $X_1, \dots, X_n$. b) Merali sme priemer $n = 10$ súčastok. Výsledky sú dané tabuľkou:

1	2	3	4	5	6	7	8	9	10
1010	1035	997	978	1001	988	963	982	999	1031

Predpokladajme, že priemer súčastok má rozdelenie $N(\mu, \sigma^2)$. Odhadnite parametre μ a σ^2 metódou maximálnej vierohodnosti.

Príklad 10.5. Merali sme výšku 12 rôznych detí vo vekoch od $x_1 = 18, x_2 = 19, \dots, x_{12} = 29$ mesiacov. Predpokladáme, že priemerná výška (v danom vekovom rozmedzí) má lineárny trend v závislosti od veku, pričom individuálne odchýlky od priemernej výšky majú normálne rozdelenie s rozptylom nezávislým od veku. Formálne: pre dieťa vo veku x_i je nameraná výška y_i realizáciou náhodnej premennej $Y_i = ax_i + b + \varepsilon_i$, kde $\varepsilon_i \sim N(0, \sigma^2)$. Nájdite regresnú priamku $\hat{a}x + \hat{b}$, kde $\hat{a}$ a $\hat{b}$ sú odhady parametrov a, b metódou najmenších štvorcov. Platí: $\sum_{i=1}^{12} x_i = 282$; $\sum_{i=1}^{12} x_i^2 = 6770$; $\sum_{i=1}^{12} y_i = 958,5$; $\sum_{i=1}^{12} y_i x_i = 22711,3$.

Príklady na precvičenie

Príklad 10.6. Nech $X_1, \dots, X_n$ je náhodný výber z rozdelenia $Geo(p)$ (začínajúceho od 0), kde $p \in (0, 1)$ je neznámy parameter. Ukážte, že odhad p pomocou metódy maximálnej vierohodnosti je $\hat{p} = \frac{n}{n + \sum_{i=1}^n X_i}$.

Príklad 10.7. Nech $X_1, \dots, X_n$ je náhodný výber z tzv. Rayleighovho rozdelenia s neznámym parametrom σ^2 , t.j. X_i majú hustotu $f(x) = \frac{x}{\sigma^2} e^{-x^2/(2\sigma^2)}$ pre $x > 0$ a $f(x)$ je 0 inde. Nájdite odhad $\hat{\sigma}^2$ parametra σ^2 pomocou metódy maximálnej vierohodnosti ako funkciu premenných $X_1, \dots, X_n$. Určte strednú hodnotu odhadu $\hat{\sigma}^2$, keď vieme, že $E(X_i) = \sigma\sqrt{\pi/2}$ a $D(X_i) = \frac{4-\pi}{2}\sigma^2$.

Príklad 10.8. Meriame rýchlosť pohybujúceho sa objektu v časoch $t_i = i$ sekúnd, pre $i = 1, 2, 3, 4, 5$. Z fyzikálnej podstaty problému a vlastností meracích zariadení vieme, že nameraná rýchlosť objektu v čase t je náhodná premenná s rozdelením $N(at + b, \sigma^2)$, kde a je zrýchlenie objektu, b je jeho počiatočná rýchlosť a σ^2 je konštanta charakterizujúca chybovosť meraní. Naviac vieme, že merania sú navzájom nezávislé. Konkrétne namerané hodnoty rýchlosti boli $v_1 = 99, 27$; $v_2 = 108, 80$; $v_3 = 119, 39$; $v_4 = 128, 30$ a $v_5 = 139, 88$ m/s. Platí: $\sum_{i=1}^5 v_i = 595, 64$ a $\sum_{i=1}^5 t_i v_i = 1887, 64$. Nájdite odhad $\hat{a}$ zrýchlenia a $\hat{b}$ počiatočnej rýchlosti metódou najmenších štvorcov.

Príklad 10.9. Uvažujte lineárny model $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, $i = 1, \dots, n$, kde $\varepsilon_1, \dots, \varepsilon_n$ majú rozdelenie $N(0, \sigma^2)$ a sú nezávislé. **a)** Nájdite odhad parametrov β_0 a β_1 metódou maximálnej vierohodnosti. **b)** Nájdite odhad parametrov β_0 a β_1 metódou najmenších štvorcov (priamym postupom). **c)** Nájdite odhad parametrov β_0 a β_1 metódou najmenších štvorcov pomocou vzorca $\hat{\beta} = (F^T F)^{-1} F^T y$.

Príklad* 10.10 (Test dobrej zhody). Počas dvoch týždňov sme pozorovali, ako dlho trvá bankovému úradníkovi obsluha jedného zákazníka. Výsledné dĺžky obsluhy $y_1, \dots, y_{200}$ sú v tabuľke združené po skupinách (napr. menej ako 5 minút obsluhoval 58 zákazníkov).

dĺžka obsluhy (v minútach)	[0,5)	[5,10)	[10,20)	[20,30)	[30,∞)
počet zákazníkov	58	42	58	26	16

Pomocou testu dobrej zhody na intervaloch daných tabuľkou testujte na 5%-nej hladine hypotézu H_0 , že dĺžka obsluhy zákazníkov bankovým úradníkom má exponenciálne rozdelenie s parametrom $\lambda = 1/10$. Vieme, že 5%-ná kritická hodnota rozdelenia χ^2_4 je $\chi^2_4(0,05) = 9,49$.