

MATEMATICKÉ METÓDY PLÁNOVANIA A VYHODNOCOVANIA EXPERIMENTU V PRÍKLADOCH

Oleg PALUMBÍNY

2008



SLOVENSKÁ TECHNICKÁ UNIVERZITA V BRATISLAVE
Materiálovotechnologická fakulta

© Doc. RNDr. Oleg Palumbíny, CSc.

Recenzenti: Doc. RNDr. Anna Tirpáková, CSc.
 Doc. RNDr. Mária Mišútová, PhD.

Jazyková korektúra: Mgr. Valéria Krahulcová

Schválila Vedecká rada Materiálovotechnologickej fakulty STU ako vysokoškolské skriptum dňa 12. decembra 2007 pre študijný program Informatika Materiálovotechnologickej fakulty STU v Trnave

ISBN 978-80-8096-054-4
EAN 9788080960544

1 ÚVOD

Táto publikácia je zbierkou riešených úloh z predmetov venovaných plánovaniu a vyhodnocovaniu experimentu. Je predovšetkým určená pre dennú formu štúdia predmetu *Matematické metódy plánovania a vyhodnocovania experimentu*. Môže však byť vhodnou pomôckou aj pre poslucháčov diaľkového štúdia ako, aj pre prezenčnú aj dištančnú formu štúdia predmetu *Matematické metódy vyhodnocovania experimentu*, ktorý má polovičnú hodinovú dotáciu ako spomínané Matematické metódy plánovania a vyhodnocovania experimentu.

Z obsahovej stránky možno prácu rozdeliť na tri časti – základy teórie pravdepodobnosti, základy matematickej štatistiky a samotnej teórie plánovania a vyhodnocovania experimentu. Spomínané prvé dva celky sú nevyhnutné na pochopenie tretieho. Navyše sa u poslucháčov predpokladá znalosť učiva z prvých troch semestrov matematiky bakalárskeho štúdia.

Obsah tejto zbierky nadväzuje na prednášky z predmetu Matematické metódy plánovania a vyhodnocovania experimentu. Tieto prednášky sú vedené štýlom obvyklým v matematike a jej aplikáciách. Skladajú sa z **definícií**, kde je presne vymedzený zavádzaný pojem (*vytlačený kurzívou*), **viet** – v nich sú sformulované základné vlastnosti vymedzených pojmov, **príkladov** – v nich sú uvedené aplikácie definovaných pojmov založené na uvedených vetách a **poznámok** – kde sú uvedené doplňujúce a vysvetľujúce informácie o zavedených pojmoch. K úspešnému zvládnutiu predmetu je nevyhnutné vedieť všetky definície a vety. Tiež je potrebné pochopiť a zapamätať si obsah poznámok. Na základe takto zvládnutej teórie je potom možné pristúpiť k jej použitiu v praxi. Vhodnou pomôckou na takúto jej aplikáciu môže byť aj táto publikácia.

Samotná zbierka sa skladá z dvanástich kapitol, ktoré približne kopírujú obsah prednášok z predmetu Matematické metódy plánovania a vyhodnocovania experimentu.

Na záver by som vám chcel zaželať, aby sa tento učebný text stal pre vás užitočnou pomôckou pri osvojovaní si základov metód plánovania a vyhodnocovania experimentu.

Autor

2 HUSTOTA, DISTRIBUČNÁ FUNKCIA, KVANTILOVÁ FUNKCIA A MOMENTY

Príklad 1. Určte c tak, aby (jednopredpisová) funkcia $f_X(x) = c/(x^2 + 1)$, $x \in (-\infty; \infty)$ bola hustotou pravdepodobnosti spojitej náhodnej premennej X . Tiež stanovte distribučnú a kvantilovú funkciu tejto náhodnej premennej, ako aj jej medián a modus.

Riešenie. Hodnotu konštanty c určíme z vlastností hustoty takto:

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} f_X(x) dx = \int_{-\infty}^0 f_X(x) dx + \int_0^{\infty} f_X(x) dx = 2 \int_0^{\infty} f_X(x) dx = 2c \int_0^{\infty} \frac{1}{x^2 + 1} dx = \\ &= 2c \lim_{x \rightarrow \infty} [\arctg t]_{t=0}^x = 2c \lim_{x \rightarrow \infty} [\arctg x - \arctg 0] = 2c \lim_{x \rightarrow \infty} \arctg x = 2c \frac{\pi}{2} = c\pi. \end{aligned}$$

Odtiaľ $c = \frac{1}{\pi}$ a následne $f_X(x) = \frac{1}{\pi} \frac{1}{x^2 + 1}$, $x \in (-\infty; \infty)$. Poznamenávame, že pri výpočte sme využili párnosť funkcie $\frac{1}{x^2 + 1}$.

Pre distribučnú funkciu $F_X(x)$ platí:

$$\begin{aligned} F_X(x) &= \int_{-\infty}^x f_X(t) dt = \int_{-\infty}^x \frac{1}{\pi} \frac{1}{t^2 + 1} dt = \lim_{s \rightarrow -\infty} \left[\frac{1}{\pi} \arctg t \right]_{t=s}^x = \\ &= \frac{1}{\pi} \lim_{s \rightarrow -\infty} [\arctg x - \arctg s] = \frac{1}{\pi} \left[\arctg x - \frac{-\pi}{2} \right] = \frac{1}{2} + \frac{\arctg x}{\pi}, \quad x \in (-\infty; \infty). \end{aligned}$$

Kvantilová funkcia $F_X^{-1}(\alpha)$ je inverzná funkcia k distribučnej funkcii $F_X(x)$. Určíme ju riešením rovnice $\alpha = F_X(x_\alpha) = \frac{1}{2} + \frac{\arctg x_\alpha}{\pi}$ s neznámou x_α , pričom x_α je α -kvantil rozdelenia pravdepodobnosti náhodnej premennej X a $\alpha \in (0; 1)$ je známy parameter. Vychádza

$$x_\alpha = \operatorname{tg} \left[\pi \left(\alpha - \frac{1}{2} \right) \right].$$

Preto kvantilová funkcia (ktorá je funkciou premennej α) $F_X^{-1}(\alpha) = \operatorname{tg} \left[\pi \left(\alpha - \frac{1}{2} \right) \right]$, $\alpha \in (0; 1)$.

Pre medián Me platí, že $Me = F_X^{-1}\left(\frac{1}{2}\right) = \operatorname{tg} \left[\pi \left(\frac{1}{2} - \frac{1}{2} \right) \right] = \operatorname{tg} 0 = 0$. Modus Mo určíme ako to reálne číslo, v ktorom hustota $f_X(x)$ nadobudne svoje lokálne maximum. Z nutnej podmienky extrém dostaneme rovnicu (s neznámou x)

$$0 = \frac{df_X(x)}{dx} = \frac{d}{dx} \left(\frac{1}{\pi} \frac{1}{x^2 + 1} \right) = \frac{1}{\pi} \frac{-2x}{(x^2 + 1)^2}.$$

Jej riešením vychádza $x = 0$. Ďalej

$$\left[\frac{d^2 f_X(x)}{dx^2} \right]_{x=0} = \left[\frac{1}{\pi} \frac{-2(x^2 + 1)^2 + 2x \cdot 2(x^2 + 1) \cdot 2x}{(x^2 + 1)^4} \right]_{x=0} = -\frac{2}{\pi} < 0,$$

odkiaľ vyplýva, že hustota $f_X(x)$ nadobúda v bode 0 svoje (jediné) lokálne maximum. Preto rozdelenie pravdepodobnosti náhodnej premennej X je monomodálne a platí, že $Mo = 0$.

Príklad 2. Vypočítajte strednú hodnotu, disperziu, šikmosť a špicatosť spojitej náhodnej premennej X s (dvojpredpisovou) hustotou $f_X(x) = 3x^2$, $x \in \langle -1; 1 \rangle$, $f_X(x) = 0$, $x \in \mathbb{R} - \langle -1; 1 \rangle$.

Riešenie. Pre strednú hodnotu $E(X)$ platí:

$$\begin{aligned} E(x) &= \int_{-\infty}^{\infty} x f_X(x) dx = \int_{-\infty}^{-1} x f_X(x) dx + \int_{-1}^1 x f_X(x) dx + \int_1^{\infty} x f_X(x) dx = \\ &= \int_{-\infty}^{-1} x \cdot 0 dx + \int_{-1}^1 x \cdot 3x^2 dx + \int_1^{\infty} x \cdot 0 dx = 0 + \int_{-1}^1 3x^3 dx + 0 = \left[\frac{3}{4} x^4 \right]_{x=-1}^1 = 0. \end{aligned}$$

Pre disperziu $D(X)$ (ktorá sa tiež značí symbolom σ^2) vychádza:

$$\begin{aligned} D(x) &= \int_{-\infty}^{\infty} (x - E(X))^2 f_X(x) dx = \int_{-\infty}^{-1} (x - E(X))^2 f_X(x) dx + \int_{-1}^1 (x - E(X))^2 f_X(x) dx + \\ &+ \int_1^{\infty} (x - E(X))^2 f_X(x) dx = \int_{-\infty}^{-1} 0 dx + \int_{-1}^1 (x - E(X))^2 3x^2 dx + \int_1^{\infty} 0 dx = \\ &= 0 + \int_{-1}^1 (x - 0)^2 3x^2 dx + 0 = \int_{-1}^1 3x^4 dx = \left[\frac{3}{5} x^5 \right]_{x=-1}^1 = \frac{6}{5}. \end{aligned}$$

$$\text{Smerodajná odchýlka } \sigma = \sqrt{D(X)} = \sqrt{\frac{6}{5}}.$$

Šikmosť $\alpha(X)$ sa počíta podľa vzorca $\alpha(X) = \frac{\mu_3}{\sigma^3}$, kde

$$\begin{aligned} \mu_3 &= \int_{-\infty}^{\infty} (x - E(X))^3 f_X(x) dx = \int_{-\infty}^{-1} (x - E(X))^3 f_X(x) dx + \int_{-1}^1 (x - E(X))^3 f_X(x) dx + \\ &+ \int_1^{\infty} (x - E(X))^3 f_X(x) dx = \int_{-\infty}^{-1} 0 dx + \int_{-1}^1 (x - E(X))^3 3x^2 dx + \int_1^{\infty} 0 dx = \\ &= 0 + \int_{-1}^1 (x - 0)^3 3x^2 dx + 0 = \int_{-1}^1 3x^5 dx = \left[\frac{3}{6} x^6 \right]_{x=-1}^1 = 0. \end{aligned}$$

Preto $\alpha(X) = 0$. Pre špicatosť $\beta(X) = \frac{\mu_4}{\sigma^4} - 3$ platí:

$$\begin{aligned} \mu_4 &= \int_{-\infty}^{\infty} (x - E(X))^4 f_X(x) dx = \int_{-\infty}^{-1} (x - E(X))^4 f_X(x) dx + \int_{-1}^1 (x - E(X))^4 f_X(x) dx + \\ &+ \int_1^{\infty} (x - E(X))^4 f_X(x) dx = \int_{-\infty}^{-1} 0 dx + \int_{-1}^1 (x - E(X))^4 3x^2 dx + \int_1^{\infty} 0 dx = \\ &= 0 + \int_{-1}^1 (x - 0)^4 3x^2 dx + 0 = \int_{-1}^1 3x^6 dx = \left[\frac{3}{7} x^7 \right]_{x=-1}^1 = \frac{6}{7}. \end{aligned}$$

$$\text{Z toho } \beta(X) = \frac{\mu_4}{\sigma^4} - 3 = \frac{6/7}{(6/5)^2} - 3 = \frac{101}{42}.$$

Príklad 3. Vypočítajte všetky počiatkové a všetky centrálné momenty spojitej náhodnej premennej X s (dvojpredpisovou) hustotou $f_X(x) = 1/3$, $x \in \langle 0; 3 \rangle$, $f_X(x) = 0$, $x \in R - \langle 0; 3 \rangle$.

Riešenie. Pre k -ty počiatkový moment ν_k , kde $k \in N$, platí:

$$\begin{aligned}\nu_k &= \int_{-\infty}^{\infty} x^k f_X(x) dx = \int_{-\infty}^0 x^k f_X(x) dx + \int_0^3 x^k f_X(x) dx + \int_3^{\infty} x^k f_X(x) dx = \\ &= \int_{-\infty}^0 0 dx + \int_0^3 x^k \frac{1}{3} dx + \int_3^{\infty} 0 dx = 0 + \left[\frac{1}{3(k+1)} x^{k+1} \right]_{x=0}^3 + 0 = \frac{3^{k+1}}{3(k+1)} = \frac{3^k}{k+1}.\end{aligned}$$

Špeciálne $E(X) = \nu_1 = \frac{3^1}{1+1} = \frac{3}{2}$. Pre k -ty centrálny moment μ_k , kde $k \in N$, vychádza:

$$\begin{aligned}\mu_k &= \int_{-\infty}^{\infty} (x - \nu_1)^k f_X(x) dx = \int_{-\infty}^0 (x - \nu_1)^k f_X(x) dx + \int_0^3 (x - \nu_1)^k f_X(x) dx + \\ &\quad + \int_3^{\infty} (x - \nu_1)^k f_X(x) dx = \int_{-\infty}^0 0 dx + \int_0^3 \left(x - \frac{3}{2}\right)^k \frac{1}{3} dx + \int_3^{\infty} 0 dx = \\ &= 0 + \left[\frac{1}{3(k+1)} \left(x - \frac{3}{2}\right)^{k+1} \right]_{x=0}^3 + 0 = \frac{1}{3(k+1)} \left[\left(\frac{3}{2}\right)^{k+1} - \left(-\frac{3}{2}\right)^{k+1} \right] = \frac{1}{3(k+1)} \left(\frac{3}{2}\right)^{k+1} [1 + (-1)^{k+1}]\end{aligned}$$

Posledný výraz sa pre nepárne k rovná číslu $\frac{1}{k+1} \left(\frac{3}{2}\right)^k$, pre párne k nule. Preto platí:

$$\mu_k = \frac{1}{k+1} \left(\frac{3}{2}\right)^k \text{ pre } k \text{ nepárne,}$$

$$\mu_k = 0 \text{ pre } k \text{ párne.}$$

3 ŠTATISTICKÝ SÚBOR, ZNAK, VARIÁČNÉ TRIEDENIE, VÝBEROVÉ KVANTILY A VÝBEROVÉ MOMENTY

Príklad 4. Vykonajte 1-stupňové triedenie súboru V podľa kvalitatívneho znaku X , ktorý udáva, či je daná osoba muž alebo žena, pričom

$$V = \{Jano, Peter, Jozef, Mário, Patrik, Jana, Monika, Petra, Karolína\}.$$

Riešenie. Daný súbor bude rozdelený do nasledujúcich dvoch tried:

$$\begin{aligned} T_1 &= \{x_i \in V, X(x_i) = \text{muž}\} = \{Jano, Peter, Jozef, Mário, Patrik\} \\ T_2 &= \{x_i \in V, X(x_i) = \text{žena}\} = \{Jana, Monika, Petra, Karolína\}. \end{aligned}$$

Príklad 5. Vykonajte 2-stupňové triedenie súboru V podľa znakov X_1, X_2 pričom prvý znak udáva pohlavie osoby a druhý znak jej vek. Daný súbor tvoria nasledujúce osoby:

$$V = \{Jano, Peter, Mário, Klaudia, Alena, Zuzana, Nina\},$$

pričom Jano má 25 rokov, Peter 31, Mário 29, Alena 43, Klaudia 23, Zuzana 27 a Nina 33. Znak $X_1: V \rightarrow \{\text{muž}, \text{žena}\}$, $X_2: V \rightarrow \{\langle 0; 29 \rangle, \langle 30; 31 \rangle, \langle 32; 100 \rangle\}$.

Riešenie. Daný súbor bude rozdelený podľa prvého znaku do dvoch 1-indexových tried:

$$\begin{aligned} T_1 &= \{x_i \in V, X(x_i) = \text{muž}\} = \{Jano, Peter, Mário\} \\ T_2 &= \{x_i \in V, X(x_i) = \text{žena}\} = \{Klaudia, Alena, Zuzana, Nina\}. \end{aligned}$$

Každá z týchto dvoch tried bude podľa druhého znaku rozdelená do troch 2-indexových tried:

$$\begin{aligned} T_{11} &= \{x_i \in T_1, X_2(x_i) \in \langle 0; 29 \rangle\} = \{Jano, Mário\} \\ T_{12} &= \{x_i \in T_1, X_2(x_i) \in \langle 30; 31 \rangle\} = \{Peter\} \\ T_{13} &= \{x_i \in T_1, X_2(x_i) \in \langle 32; 100 \rangle\} = \{\} \\ T_{21} &= \{x_i \in T_2, X_2(x_i) \in \langle 0; 29 \rangle\} = \{Klaudia, Zuzana\} \\ T_{22} &= \{x_i \in T_2, X_2(x_i) \in \langle 30; 31 \rangle\} = \{\} \\ T_{23} &= \{x_i \in T_2, X_2(x_i) \in \langle 32; 100 \rangle\} = \{Nina, Alena\}, \end{aligned}$$

pričom symbol $\{\}$ znamená prázdnu množinu, t.j. taký súbor, ktorý neobsahuje žiadny prvok.

Príklad 6. Vykonajte 3-stupňové triedenie súboru V podľa znakov X_1, X_2, X_3 , pričom prvý znak udáva pohlavie osoby, druhý typ vzdelania a tretí jej vek. Daný súbor tvoria osoby:

$$V = \{Jano, Rudo, Michal, Karol, Soňa, Petra, Diana\},$$

pričom Jano má 25 rokov a ZŠ, Rudo 31 a SŠ, Michal 29 a SŠ, Karol 43 a VŠ, Soňa 43 a VŠ, Petra 28 a VŠ, Diana 31 a SŠ. Uvedené znaky sú nasledujúce funkcie definované na V : $X_1: V \rightarrow \{\text{muž}, \text{žena}\}$, $X_2: V \rightarrow \{ZŠ, SŠ, VŠ\}$, $X_3: V \rightarrow \{\langle 0; 30 \rangle, \langle 31; 35 \rangle, \langle 36; 100 \rangle\}$.

Riešenie. Daný súbor bude rozdelený podľa prvého znaku do dvoch 1-indexových tried:

$$\begin{aligned} T_1 &= \{x_i \in V, X(x_i) = \text{muž}\} = \{Jano, Rudo, Michal, Karol\} \\ T_2 &= \{x_i \in V, X(x_i) = \text{žena}\} = \{Soňa, Petra, Diana\}. \end{aligned}$$

Každá z týchto dvoch tried bude podľa druhého znaku rozdelená do troch 2-indexových tried:

$$\begin{aligned}
T_{11} &= \{x_i \in T_1, X_2(x_i) \in \{Z\check{S}\}\} = \{Jano\} \\
T_{12} &= \{x_i \in T_1, X_2(x_i) \in \{S\check{S}\}\} = \{Rudo, Michal\} \\
T_{13} &= \{x_i \in T_1, X_2(x_i) \in \{V\check{S}\}\} = \{Karol\} \\
T_{21} &= \{x_i \in T_1, X_2(x_i) \in \{Z\check{S}\}\} = \{\} \\
T_{22} &= \{x_i \in T_1, X_2(x_i) \in \{S\check{S}\}\} = \{Diana\} \\
T_{23} &= \{x_i \in T_1, X_2(x_i) \in \{V\check{S}\}\} = \{So\check{n}o, Petra\}.
\end{aligned}$$

Každá z týchto šiestich tried bude podľa tretieho znaku rozdelená do troch 3-indexových tried:

$$\begin{aligned}
T_{111} &= \{x_i \in T_{11}, X_3(x_i) \in \langle 0; 30 \rangle\} = \{Jano\} \\
T_{112} &= \{x_i \in T_{11}, X_3(x_i) \in \langle 31; 35 \rangle\} = \{\} \\
T_{113} &= \{x_i \in T_{11}, X_3(x_i) \in \langle 36; 100 \rangle\} = \{\} \\
T_{121} &= \{x_i \in T_{11}, X_3(x_i) \in \langle 0; 30 \rangle\} = \{Michal\} \\
T_{122} &= \{x_i \in T_{11}, X_3(x_i) \in \langle 31; 35 \rangle\} = \{Rudo\} \\
T_{123} &= \{x_i \in T_{11}, X_3(x_i) \in \langle 36; 100 \rangle\} = \{\} \\
T_{131} &= \{x_i \in T_{11}, X_3(x_i) \in \langle 0; 30 \rangle\} = \{\} \\
T_{132} &= \{x_i \in T_{11}, X_3(x_i) \in \langle 31; 35 \rangle\} = \{\} \\
T_{133} &= \{x_i \in T_{11}, X_3(x_i) \in \langle 36; 100 \rangle\} = \{Karol\} \\
T_{211} &= \{x_i \in T_{11}, X_3(x_i) \in \langle 0; 30 \rangle\} = \{\} \\
T_{212} &= \{x_i \in T_{11}, X_3(x_i) \in \langle 31; 35 \rangle\} = \{\} \\
T_{213} &= \{x_i \in T_{11}, X_3(x_i) \in \langle 36; 100 \rangle\} = \{\} \\
T_{221} &= \{x_i \in T_{11}, X_3(x_i) \in \langle 0; 30 \rangle\} = \{\} \\
T_{222} &= \{x_i \in T_{11}, X_3(x_i) \in \langle 31; 35 \rangle\} = \{Diana\} \\
T_{223} &= \{x_i \in T_{11}, X_3(x_i) \in \langle 36; 100 \rangle\} = \{\} \\
T_{231} &= \{x_i \in T_{11}, X_3(x_i) \in \langle 0; 30 \rangle\} = \{Petra\} \\
T_{232} &= \{x_i \in T_{11}, X_3(x_i) \in \langle 31; 35 \rangle\} = \{\} \\
T_{233} &= \{x_i \in T_{11}, X_3(x_i) \in \langle 36; 100 \rangle\} = \{So\check{n}o\}.
\end{aligned}$$

Príklad 7. Vypočítajte výberové kvartily kvantitatívneho znaku X , ak sme namerali nasledujúce hodnoty tohto znaku: 7, 5, -2, 10, 8, 12, 15.

Riešenie. Dané hodnoty nie sú usporiadané podľa veľkosti. Ak ich zoradíme vzostupne podľa veľkosti, dostaneme usporiadaný výber podľa znaku X :

$$(x_1^*, x_2^*, x_3^*, x_4^*, x_5^*, x_6^*, x_7^*) = (-2, 5, 7, 8, 10, 12, 15).$$

Kvartily rozdelia daný usporiadaný výber na 4 skupiny po rovnakom počte prvkov (v našom prípade skupiny budú obsahovať po jednom prvku). Preto dolný kvartil $\tilde{x}_{25} = 5$, stredný kvartil (t.j. medián) $\tilde{x}_{50} = 8$ a horný kvartil $\tilde{x}_{75} = 12$.

Príklad 8. Vypočítajte výberový medián kvantitatívneho znaku X , ak sme namerali nasledujúce hodnoty tohto znaku: -3, -7, -10, 10, 8, 15, 2.

Riešenie. Dané hodnoty nie sú usporiadané podľa veľkosti. Ak ich zoradíme vzostupne podľa veľkosti, dostaneme usporiadaný výber podľa znaku X :

$$(x_1^*, x_2^*, x_3^*, x_4^*, x_5^*, x_6^*, x_7^*) = (-10, -7, -3, 2, 8, 10, 15).$$

Medián rozdelí daný usporiadaný výber na 2 skupiny o rovnakom počte prvkov (v našom prípade skupiny budú obsahovať po troch prvkoch, ktorých hranicou je hodnota $x_4^* = 2$). Preto medián $\tilde{x}_{50} = 2$.

Príklad 9. Vypočítajte výberový medián kvantitatívneho znaku X , ak sme namerali nasledujúce hodnoty tohto znaku: 3, 7, 10, 8, 15, 2.

Riešenie. Dané hodnoty nie sú usporiadané podľa veľkosti. Ak ich zoradíme vzostupne podľa veľkosti, dostaneme usporiadaný výber podľa znaku X :

$$(x_1^*, x_2^*, x_3^*, x_4^*, x_5^*, x_6^*) = (2, 3, 7, 8, 10, 15).$$

Výberový medián rozdelí daný usporiadaný výber na 2 skupiny o rovnakom počte prvkov (v našom prípade skupiny budú obsahovať po troch prvkoch). Keďže rozsah výberu je 6 (t.j. párne číslo), výberový medián dostaneme ako aritmetický priemer tretieho a štvrtého prvkov v usporiadanom výbere, teda

$$\tilde{x}_{50} = \frac{x_3^* + x_4^*}{2} = \frac{7 + 8}{2} = 7,5.$$

Príklad 10. Vypočítajte tretí a siedmy výberový decil kvantitatívneho znaku X , ak sme namerali nasledujúce hodnoty tohto znaku: -10, -8, -9, -5, -7, -6, -4, -2, -3, -1.

Riešenie. Dané hodnoty nie sú usporiadané podľa veľkosti. Ak ich zoradíme vzostupne podľa veľkosti, dostaneme usporiadaný výber podľa znaku X :

$$(x_1^*, x_2^*, x_3^*, x_4^*, x_5^*, x_6^*, x_7^*, x_8^*, x_9^*, x_{10}^*) = (-10, -9, -8, -7, -6, -5, -4, -3, -2, -1).$$

Výberové decily rozdelia potom daný usporiadaný výber na desať skupín o rovnakom počte prvkov (v našom prípade skupiny budú obsahovať po jednom prvku). Preto

$$\begin{aligned}\tilde{x}_{30} &= \frac{x_3^* + x_4^*}{2} = \frac{-8 + (-7)}{2} = -7,5, \\ \tilde{x}_{70} &= \frac{x_7^* + x_8^*}{2} = \frac{-4 + (-3)}{2} = -3,5.\end{aligned}$$

Príklad 11. Vypočítajte tridsiaty, päťdesiaty a sedemdesiaty druhý výberový percentil kvantitatívneho znaku X , ak namerané hodnoty tohto znaku sú kladné celé čísla od 1 do 10000.

Riešenie. Ak týchto 10000 hodnôt usporiadame podľa veľkosti, dostaneme usporiadaný výber podľa znaku X :

$$(x_1^*, x_2^*, x_3^*, \dots, x_{9999}^*, x_{10000}^*) = (1, 2, 3, \dots, 9999, 10000).$$

Výberové percentily rozdelia potom daný usporiadaný výber na sto skupín o rovnakom počte prvkov (v našom prípade skupiny budú obsahovať po sto prvkov). Preto

$$\begin{aligned}\tilde{x}_{30} &= \frac{x_{30}^* + x_{31}^*}{2} = \frac{30 + 31}{2} = 30,5, \\ \tilde{x}_{50} &= \frac{x_{50}^* + x_{51}^*}{2} = \frac{50 + 51}{2} = 50,5, \\ \tilde{x}_{72} &= \frac{x_{72}^* + x_{73}^*}{2} = \frac{72 + 73}{2} = 72,5.\end{aligned}$$

Príklad 12. Vypočítajte tretí výberový decil kvantitatívneho znaku X , ak namerané hodnoty tohto znaku sú čísla 1, 1, 2, 2, 3, 2, 3, 3, 2, 3.

Riešenie. Dané hodnoty nie sú usporiadané podľa veľkosti. Ak ich zoradíme vzostupne podľa veľkosti, dostaneme usporiadaný výber podľa znaku X :

$$(x_1^*, x_2^*, x_3^*, x_4^*, x_5^*, x_6^*, x_7^*, x_8^*, x_9^*, x_{10}^*) = (1, 1, 2, 2, 2, 2, 3, 3, 3, 3).$$

Čísla sa v danom výbere opakujú, ale to neprekáža. Pre tretí výberový decil aj v tomto prípade platí

$$\tilde{x}_{30} = \frac{x_3^* + x_4^*}{2} = \frac{2 + 2}{2} = 2.$$

Príklad 13. Vypočítajte štvrtý výberový decil kvantitatívneho znaku X , ak namerané hodnoty tohto znaku sú čísla 1, 1, 2, 2, 3.

Riešenie. Hľadaný štvrtý výberový decil neexistuje, lebo rozsah výberu je len 5, a preto výberové decily nemôžu rozdeliť týchto 5 prvkov do 10 skupín po rovnakom počte prvkov.

Príklad 14. Vypočítajte výberové momenty $m_1(0)$, $m_2(-2)$ a $m_5(\bar{x})$ kvantitatívneho znaku X , ak namerané hodnoty tohto znaku sú čísla 1, 2, 3, 4, 5.

Riešenie. Hľadané výberové momenty sú:

$$\begin{aligned} m_1(0) &= \bar{x} = \frac{1}{5} \sum_{i=1}^5 x_i = \frac{1+2+3+4+5}{5} = 3, \\ m_2(-2) &= \frac{1}{5} \sum_{i=1}^5 [(x_i - (-2))^2] = \frac{(1+2)^2 + (2+2)^2 + (3+2)^2 + (4+2)^2 + (5+2)^2}{5} = 27, \\ m_5(\bar{x}) &= \frac{1}{5} \sum_{i=1}^5 [(x_i - \bar{x})^5] = \frac{(1-3)^5 + (2-3)^5 + (3-3)^5 + (4-3)^5 + (5-3)^5}{5} = 0. \end{aligned}$$

Príklad 15. Vypočítajte výberové momenty $m_1(0)$, $m_2(4)$ a $m_4(\bar{x})$ kvantitatívneho znaku X , ak namerané hodnoty tohto znaku sú čísla -2, -1, 1, 2.

Riešenie. Hľadané výberové momenty sú:

$$\begin{aligned} m_1(0) &= \bar{x} = \frac{1}{4} \sum_{i=1}^4 x_i = \frac{-2 + (-1) + 1 + 2}{4} = 0, \\ m_2(4) &= \frac{1}{4} \sum_{i=1}^4 [(x_i - 4)^2] = \frac{(-1-4)^2 + (-2-4)^2 + (1-4)^2 + (2-4)^2}{4} = 18,5, \\ m_4(\bar{x}) &= \frac{1}{4} \sum_{i=1}^4 [(x_i - \bar{x})^4] = \frac{(-1-0)^4 + (-2-0)^4 + (1-0)^4 + (2-0)^4}{4} = 8,5. \end{aligned}$$

Príklad 16. Vypočítajte výberový priemer kvantitatívneho znaku X , ak namerané hodnoty tohto znaku sú všetky celé čísla od 1 do 100.

Riešenie. Hľadaný výberový priemer je:

$$\bar{x} = m_1(0) = \frac{1}{100} \sum_{i=1}^{100} x_i = \frac{1+2+3+\dots+99+100}{100} = \frac{1}{100} \frac{1}{2} 100 (1+100) = \frac{10100}{200} = 50,5.$$

Príklad 17. Vypočítajte výberový priemer kvantitatívneho znaku X , ak namerané hodnoty tohto znaku sú čísla 1, 1, 1, 2, 2, 3, 3, 3, 3, 3.

Riešenie. Hľadaný výberový priemer je:

$$\bar{x} = m_1(0) = \frac{1}{10} \sum_{i=1}^{10} x_i = \frac{1 \cdot 3 + 2 \cdot 2 + 3 \cdot 5}{10} = \frac{3+4+15}{10} = \frac{22}{10} = 2,2.$$

4 HISTOGRAM, POLYGÓN A OGIVNÁ KRIVKA

Príklad 18. Daný je nasledovný výber rozsahu 24 podľa kvantitatívneho znaku X :

$-10, -3, 3, 5, -10, -3, 4, 6, -8, -3, 5, 6, -8, 2, 5, 7, -8, 3, 5, 9, -8, 3, 5, 9$.

Zostrojte podľa neho usporiadaný výber, jeho tabuľku početností a modifikovanú tabuľku početností vzhľadom na množiny $M_1 = \{-10, -8, -3\}$, $M_2 = \{2, 3, 4, 5\}$, $M_3 = \{6, 7, 9\}$.

Riešenie. Usporiadaný výber má tvar usporiadanej 24-tice, v ktorej sú číselné hodnoty usporiadané vzostupne podľa veľkosti:

$(-10, -10, -8, -8, -8, -8, -3, -3, -3, 2, 3, 3, 3, 4, 5, 5, 5, 5, 5, 6, 6, 7, 9, 9)$.

Tabuľka početností má tvar:

x_i	-10	-8	-3	2	3	4	5	6	7	9
f_i	2	4	3	1	3	1	5	2	1	2

V hornom riadku tabuľky sú hodnoty x_i znaku X a v dolnom riadku sú absolútne početnosti f_i týchto hodnôt. Modifikovaná tabuľka početností má tvar:

M_i	M_1	M_2	M_3
p_i	9	10	5

kde v hornom riadku sú uvedené množiny M_i a v dolnom absolútne početnosti p_i týchto množín, t.j. počty prvkov x_i nachádzajúcich sa v týchto množinách.

Príklad 19. Nech je daný nasledovný výber rozsahu 30 podľa kvantitatívneho znaku X : $-20, -9, 0, -20, -9, 1, -20, -9, 1, -18, -9, 2, -18, -9, 3, -18, -9, 3, -15, -9, 3, -15, 0, 9, -15, 0, 9$. Zostrojte podľa neho usporiadaný výber, jeho tabuľku početností a modifikovanú tabuľku početností vzhľadom na množiny $M_1 = \{-20, -18\}$, $M_2 = \{-15, -9\}$, $M_3 = \{0, 1, 2\}$, $M_4 = \{3, 9\}$.

Riešenie. Usporiadaný výber má tvar usporiadanej 30-tice, v ktorej sú číselné hodnoty usporiadané vzostupne podľa veľkosti:

$(-20, -20, -20, -18, -18, -18, -15, -15, -15, -15, -9, -9, -9, -9, -9, -9, -9, 0, 0, 0, 0, 0, 1, 1, 2, 3, 3, 3, 9, 9, 9)$.

Tabuľka početností má tvar:

x_i	-20	-18	-15	-9	0	1	2	3	9
f_i	3	3	4	7	4	2	1	3	3

V hornom riadku tabuľky sú hodnoty x_i znaku X a v dolnom riadku sú absolútne početnosti f_i týchto hodnôt. Modifikovaná tabuľka početností má tvar:

M_i	M_1	M_2	M_3	M_4
p_i	6	11	7	6

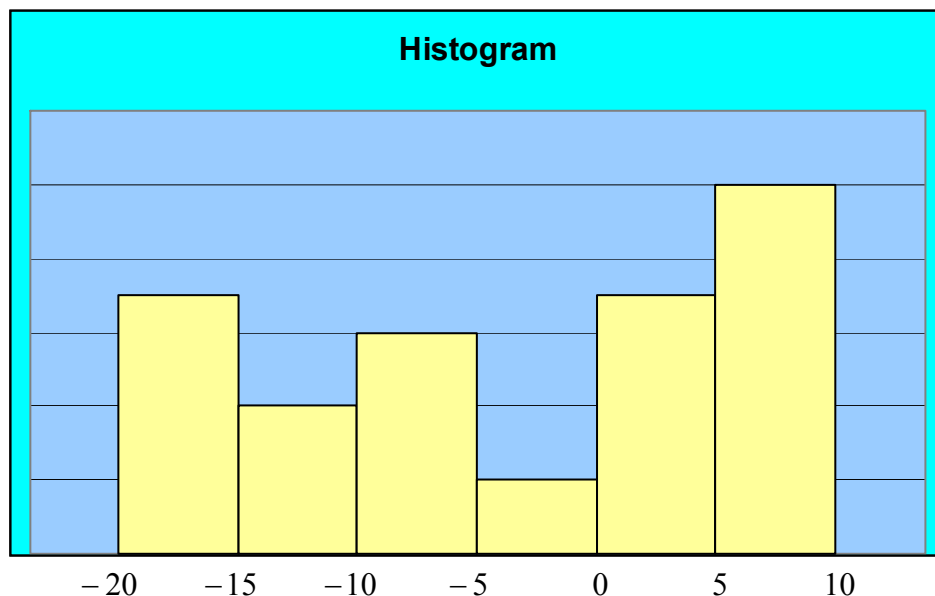
V hornom riadku sú uvedené množiny M_i a v dolnom absolútne početnosti p_i týchto množín, t.j. počty prvkov x_i nachádzajúcich sa v týchto množinách.

Príklad 20. Nech je daný usporiadaný výber rozsahu 36 podľa kvantitatívneho znaku X : $(-20, -20, -18, -17, -17, -16, -16, -15, -14, -13, -13, -10, -10, -8, -8, -7, -6, -5, -5, 1, 1, 2, 2, 3, 3, 4, 5, 5, 6, 6, 6, 7, 7, 8, 8, 10)$. Zostrojte podľa neho veľkú tabuľku početností, histogram, polygón a ogivnú krivku.

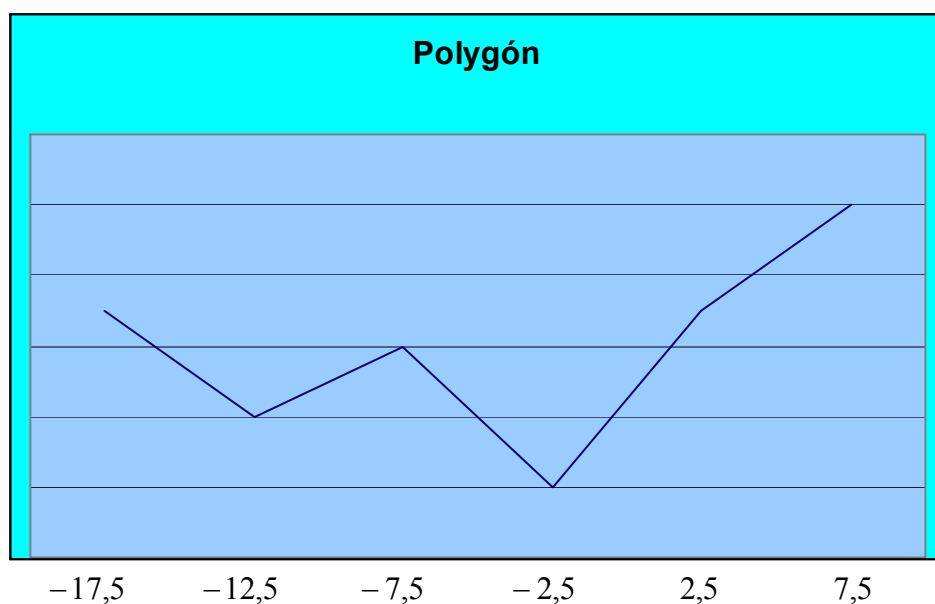
Riešenie. Počet ohraničených polouzavretých intervalov na ktoré daný súbor rozdelíme, bude $m = \sqrt{n} = \sqrt{36} = 6$. Variačné rozpätie $r = x_{\max} - x_{\min} = 10 - (-20) = 30$. Budeme voliť ohraničené intervaly $M_i, i = 1, \dots, 6$ rovnakej dĺžky $d = r/m = 30/6 = 5$. Ak označíme stred intervalu M_i symbolom s_i , absolútnu početnosť intervalu M_i symbolom n_i , relatívnu početnosť intervalu M_i symbolom f_i , kumulatívnu absolútnu početnosť intervalu M_i symbolom N_i a relatívnu kumulatívnu početnosť intervalu M_i symbolom F_i , dostaneme nasledujúcu veľkú tabuľku početností daného výberu:

i	M_i	s_i	d	n_i	f_i	N_i	F_i
1	$\langle -20; -15 \rangle$	-17,5	5	7	7/36	7	7/36
2	$\langle -15; -10 \rangle$	-12,5	5	4	4/36	11	11/36
3	$\langle -10; -5 \rangle$	-7,5	5	6	6/36	17	17/36
4	$\langle -5; 0 \rangle$	-2,5	5	2	2/36	19	19/36
5	$\langle 0; 5 \rangle$	2,5	5	7	7/36	26	26/36
6	$\langle 5; 10 \rangle$	7,5	5	10	10/36	36	1

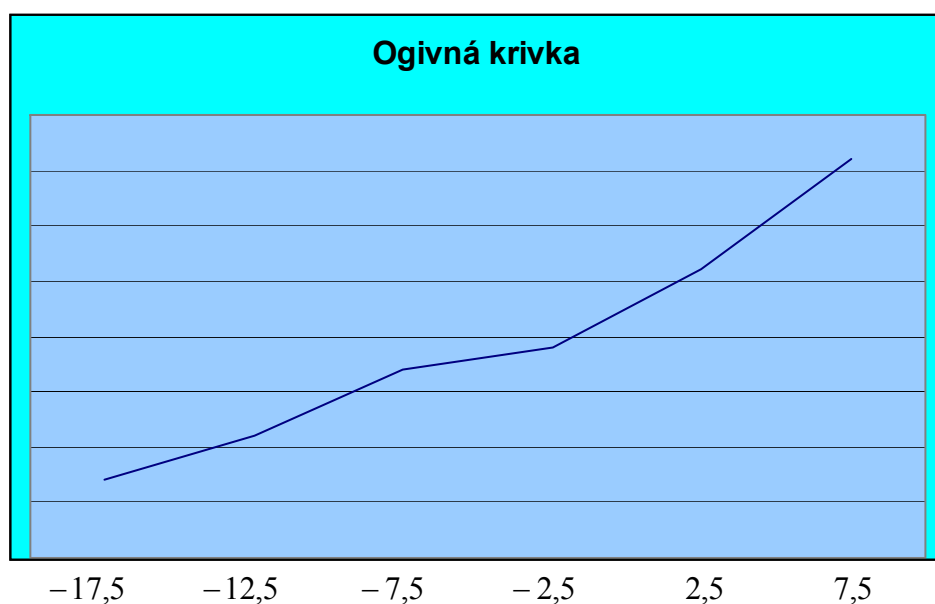
Histogram je na nasledujúcom obrázku. Je to systém 6 obdĺžnikov v rovine, pričom ich základne sú umiestnené na horizontálnej súradnicovej osi. Dĺžka týchto základní je rovnaká a má hodnotu $d = 5$, pričom základne obdĺžnikov sú intervaly $M_i, i = 1, \dots, 6$. Výška i -teho obdĺžnika je volená tak, aby jeho plošný obsah bol rovný relatívnej početnosti f_i .



Na ďalšom obrázku je polygón. Je to lomená čiara v rovine skladajúca sa z 5 úsečiek, pričom x -ové súradnice ich koncových bodov sú stredy s_i intervalov $M_i, i = 1, \dots, 6$; y -ové súradnice ich koncových bodov sú relatívne početnosti f_i .



Na nasledujúcom obrázku je ogivná krivka. Je to lomená čiara v rovine skladajúca sa z 5 úsečiek, pričom x -ové súradnice ich koncových bodov sú stredy s_i intervalov $M_i, i = 1, \dots, 6$; y -ové súradnice ich koncových bodov sú relatívne kumulatívne početnosti F_i .



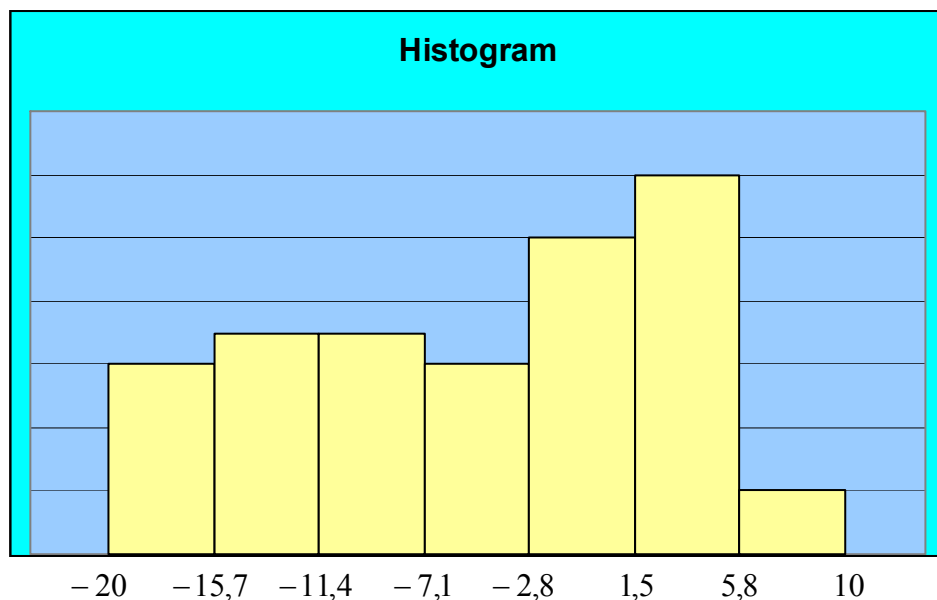
Príklad 21. Nech je daný usporiadaný výber rozsahu 50 podľa kvantitatívneho znaku X : $(-20, -20, -19, -19, -17, -16, -15, -15, -14, -14, -13, -13, -13, -10, -9, -9, -9, -9, -8, -8, -7, -7, -6, -5, -4, -3, -2, -1, -1, -1, 0, 0, 0, 1, 1, 1, 2, 2, 2, 2, 3, 3, 3, 4, 4, 5, 5, 5, 9, 10)$. Zostrojte podľa neho veľkú tabuľku početností, histogram, polygón a ogivnú krivku.

Riešenie. Počet ohraničených polouzavretých intervalov na ktoré daný súbor rozdelíme, bude po zaokrúhlení $m = \sqrt{n} = \sqrt{50} \approx 7$. Variačné rozpätie $r = x_{\max} - x_{\min} = 10 - (-20) = 30$. Budeme voliť ohraničené intervaly $M_i, i = 1, \dots, 7$ rovnakej dĺžky $d = r/m = 30/7 \approx 4,3$. Ak označíme stred intervalu M_i symbolom s_i , absolútnu početnosť intervalu M_i symbolom n_i , relatív-

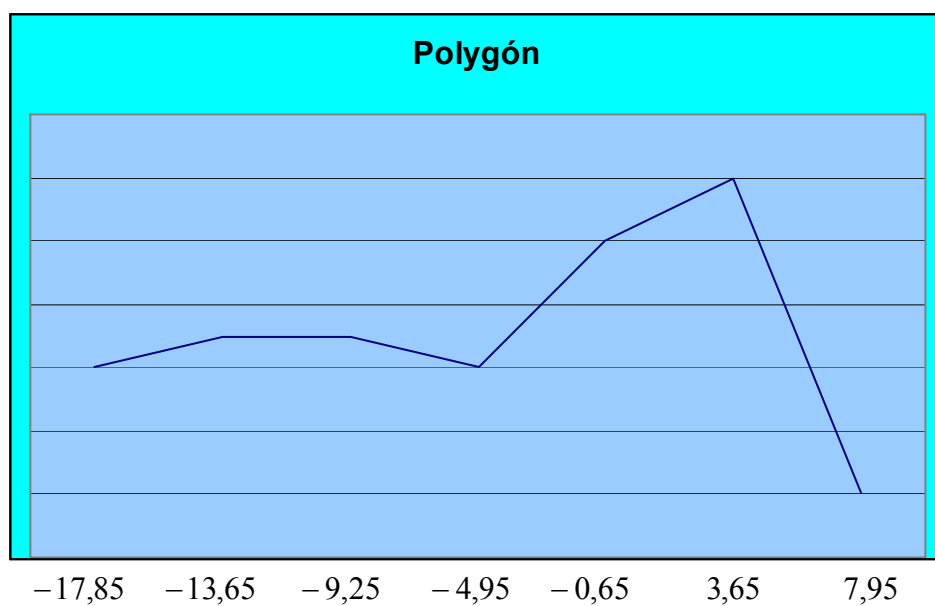
nu početnosť intervalu M_i symbolom f_i , kumulatívnu absolútnu početnosť intervalu M_i symbolom N_i a relatívnu kumulatívnu početnosť intervalu M_i symbolom F_i , dostaneme nasledujúcu veľkú tabuľku početností daného výberu:

i	M_i	s_i	d	n_i	f_i	N_i	F_i
1	$\langle -20; -15,7 \rangle$	-17,85	4,3	6	6/50	6	6/50
2	$\langle -15,7; -11,4 \rangle$	-13,55	4,3	7	7/50	13	13/50
3	$\langle -11,4; -7,1 \rangle$	-9,25	4,3	7	7/50	20	20/50
4	$\langle -7,1; -2,8 \rangle$	-4,95	4,3	6	6/50	26	26/50
5	$\langle -2,8; 1,5 \rangle$	-0,65	4,3	10	10/50	36	36/50
6	$\langle 1,5; 5,8 \rangle$	3,65	4,3	12	12/50	48	48/50
7	$\langle 5,8; 10 \rangle$	7,95	4,3	2	2/50	50	1

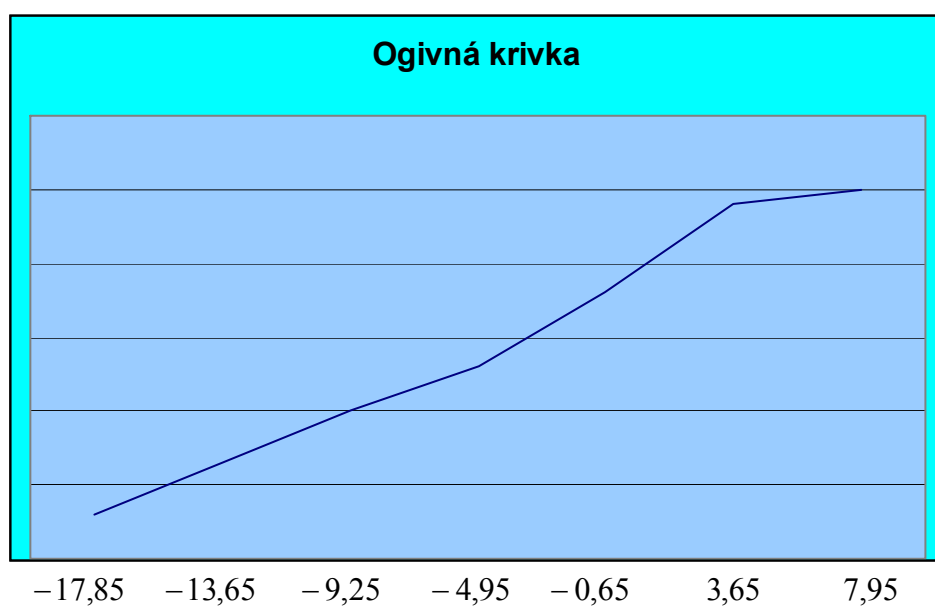
Histogram je na nasledujúcom obrázku. Je to systém 7 obdĺžnikov v rovine, pričom ich základne sú umiestnené na horizontálnej súradnicovej osi. Dĺžka týchto základní je rovnaká a má hodnotu $d = 4,3$, pričom základne obdĺžnikov sú intervaly $M_i, i = 1, \dots, 7$. Výška i -teho obdĺžnika je volená tak, aby jeho plošný obsah bol rovný relatívnej početnosti f_i .



Na ďalšom obrázku je polygón. Je to lomená čiara v rovine skladajúca sa zo 7 úsečiek, pričom x -ové súradnice ich koncových bodov sú stredy s_i intervalov $M_i, i = 1, \dots, 7$; y -ové súradnice ich koncových bodov sú relatívne početnosti f_i .



Na nasledujúcom obrázku je ogivná krivka. Je to lomená čiara v rovine skladajúca sa zo 6 úsečiek, pričom x -ové súradnice ich koncových bodov sú stredy s_i intervalov $M_i, i = 1, \dots, 7$; y -ové súradnice ich koncových bodov sú relatívne kumulatívne početnosti F_i .



5 CHARAKTERISTIKY POLOHY A VARIABILITY

Príklad 22. Nech spojitý kvantitatívny znak X udáva dĺžku oceľových trámov v centimetroch. Nech je daný usporiadaný výber rozsahu 10 podľa tohoto znaku:

(207, 207, 207, 208, 208, 208, 208, 208, 208, 209).

Odhadnite priemernú dĺžku týchto trámov.

Riešenie. Priemernú dĺžku trámov môžeme odhadnúť ľubovoľnou z charakteristík polohy. My zvolíme dve z nich – najčastejšie používaný výberový priemer $\bar{x}$, a tiež často používaný výberový medián $\tilde{x}$.

$$\bar{x} = \frac{1}{10} \sum_{i=1}^{10} x_i = \frac{1}{10} (3 \cdot 207 + 6 \cdot 208 + 1 \cdot 209) = \frac{621 + 1248 + 209}{10} = 207,8$$

$$\tilde{x} = \frac{x_5^* + x_6^*}{2} = \frac{208 + 208}{2} = 208.$$

Vidíme, že obe hodnoty sa málo od seba odlišujú; na odhad priemernej dĺžky trámov sú obe rovnako vhodné. Pretože kvantitatívny znak X bol meraný v centimetroch, z toho vyplýva, že aj obe vypočítané charakteristiky polohy sú uvedené v tejto jednotke dĺžky.

Príklad 23. Nech spojitý kvantitatívny znak X udáva dĺžku kovových hriadeľov v milimetroch. Nech je daný výber rozsahu 15 podľa tohoto znaku:

25, 24, 24, 25, 26, 25, 25, 24, 24, 25, 24, 25, 26, 25, 25

Odhadnite priemernú dĺžku týchto trámov.

Riešenie. Priemernú dĺžku trámov môžeme odhadnúť ľubovoľnou z charakteristík polohy. My opäť použijeme výberový priemer $\bar{x}$ a výberový medián $\tilde{x}$. Pre výberový priemer máme:

$$\bar{x} = \frac{1}{15} \sum_{i=1}^{15} x_i = \frac{1}{15} (5 \cdot 24 + 8 \cdot 25 + 2 \cdot 26) = \frac{120 + 200 + 52}{15} = \frac{372}{15} = \frac{124}{5} = 24,8.$$

Pri výpočte výberového mediánu budeme potrebovať usporiadaný výber. Je ním nasledujúca usporiadaná n -ca, kde $n = 15$: (24, 24, 24, 24, 25, 25, 25, 25, 25, 25, 25, 25, 26, 26). Pretože rozsah výberu $n = 15$ je nepárne číslo, výberový medián bude ôsma najmenšia hodnota uvažovaného znaku X , t.j. $\tilde{x} = x_8^* = 25$. Pretože kvantitatívny znak X bol meraný v milimetroch, z toho vyplýva, že aj obe vypočítané charakteristiky polohy sú uvedené v tejto jednotke dĺžky.

Príklad 24. Je daná nasledujúca veľká tabuľka početností:

i	M_i	s_i	d	n_i	f_i	N_i	F_i
1	$\langle 6; 8 \rangle$	7	2	3	3/87	3	3/87
2	$\langle 8; 10 \rangle$	9	2	8	8/87	11	8/87
3	$\langle 10; 12 \rangle$	11	2	15	15/87	26	26/87
4	$\langle 12; 14 \rangle$	13	2	27	27/87	53	53/87
5	$\langle 14; 16 \rangle$	15	2	19	19/87	72	72/87
6	$\langle 16; 18 \rangle$	17	2	10	10/87	82	82/87
7	$\langle 18; 20 \rangle$	19	2	5	5/87	87	1

Na jej základe vypočítajte výberový modus $\hat{x}$.

Riešenie. Z predchádzajúcej tabuľky vidíme, že najväčšiu absolútnu početnosť má zrejme interval M_4 . To znamená, že hľadaný výberový modus bude niekde v tomto intervale. Vieme,

že $\hat{x} = -\frac{b}{2a}$, pričom čísla a, b spĺňajú nasledujúci systém troch lineárnych rovníc o troch neznámych a, b, c :

$$as_3^2 + bs_3 + c = n_3$$

$$as_4^2 + bs_4 + c = n_4$$

$$as_5^2 + bs_5 + c = n_5$$

Hodnoty $n_i, s_i, i = 3, 4, 5$ nájdeme vo veľkej tabuľke početností, čím dostaneme systém

$$11^2a + 11b + c = 15$$

$$13^2a + 13b + c = 27$$

$$15^2a + 15b + c = 19$$

Pretože číslo c nepotrebujeme, systém vyriešime tak, že c vypočítame (napríklad) z tretej rovnice a dosadíme do predchádzajúcich dvoch rovníc. Dostaneme tak nasledujúci systém dvoch rovníc o dvoch neznámych a, b :

$$-104a - 4b = -4$$

$$-56a - 2b = 8$$

Jeho riešením dostaneme hodnoty $a = -2,5, b = 66$; odtiaľ potom pre výberový modus dostávame hodnotu $\hat{x} = 13,2$.

Príklad 25. Nech je daný usporiadaný výber rozsahu 50 podľa kvantitatívneho znaku X udávajúci náboj v coulomboch: $(-20, -20, -19, -19, -17, -16, -15, -15, -14, -14, -13, -13, -13, -10, -9, -9, -9, -9, -8, -8, -7, -7, -6, -5, -4, -3, -2, -1, -1, -1, 0, 0, 0, 1, 1, 1, 2, 2, 2, 2, 3, 3, 3, 4, 4, 5, 5, 5, 9, 10)$. Na jeho základe vypočítajte výberový modus $\hat{x}$.

Riešenie. Počet ohraničených polouzavretých intervalov na ktoré daný súbor rozdelíme, bude po zaokrúhlení $m = \sqrt{n} = \sqrt{50} \approx 7$. Variačné rozpätie $r = x_{\max} - x_{\min} = 10 - (-20) = 30$. Budeme voliť ohraničené intervaly $M_i, i = 1, \dots, 7$ rovnakej dĺžky $d = r/m = 30/7 \approx 4,3$. Ak označíme stred intervalu M_i symbolom s_i , absolútnu početnosť intervalu M_i symbolom n_i , relatívnu početnosť intervalu M_i symbolom f_i , kumulatívnu absolútnu početnosť intervalu M_i symbolom N_i a relatívnu kumulatívnu početnosť intervalu M_i symbolom F_i , dostaneme nasledujúcu veľkú tabuľku početností daného výberu:

i	M_i	s_i	d	n_i	f_i	N_i	F_i
1	$\langle -20; -15,7 \rangle$	-17,85	4,3	6	6/50	6	6/50
2	$\langle -15,7; -11,4 \rangle$	-13,55	4,3	7	7/50	13	13/50
3	$\langle -11,4; -7,1 \rangle$	-9,25	4,3	7	7/50	20	20/50
4	$\langle -7,1; -2,8 \rangle$	-4,95	4,3	6	6/50	26	26/50
5	$\langle -2,8; 1,5 \rangle$	-0,65	4,3	10	10/50	36	36/50
6	$\langle 1,5; 5,8 \rangle$	3,65	4,3	12	12/50	48	48/50
7	$\langle 5,8; 10 \rangle$	7,95	4,3	2	2/50	50	1

Z predchádzajúcej tabuľky vidíme, že najväčšiu absolútnu početnosť má interval M_6 . To znamená, že hľadaný výberový modus bude niekde v tomto intervale. Vieme, že $\hat{x} = -\frac{b}{2a}$, pričom čísla a, b spĺňajú nasledujúci systém troch lineárnych rovníc o troch neznámych a, b, c :

$$\begin{aligned}as_5^2 + bs_5 + c &= n_5 \\as_6^2 + bs_6 + c &= n_6 \\as_7^2 + bs_7 + c &= n_7\end{aligned}$$

Hodnoty $n_i, s_i, i = 5, 6, 7$ nájdeme vo veľkej tabuľke početností, čím dostaneme systém

$$\begin{aligned}(-0,65)^2 a - 0,65b + c &= 10 \\3,65^2 a + 3,65b + c &= 12 \\7,95^2 a + 7,95b + c &= 2\end{aligned}$$

Pretože číslo c nepotrebujeme, systém vyriešime tak, že c vypočítame (napríklad) z tretej rovnice a dosadíme do predchádzajúcich dvoch rovníc. Dostaneme tak nasledujúci systém dvoch rovníc o dvoch neznámych a, b :

$$\begin{aligned}-62,78a - 8,5b &= 8 \\-49,88a - 4,3b &= 10\end{aligned}$$

Jeho riešením dostaneme (po zaokrúhlení) hodnoty $a = -0,3285, b = 1,1567$; odtiaľ potom pre výberový modus dostávame (s presnosťou na dve desatinné miesta) hodnotu $\hat{x} = 1,76$. Pretože kvantitatívny znak X bol meraný v coulomboch, z toho vyplýva, že aj vypočítaný výberový modus je v tejto jednotke elektrického náboja.

Príklad 26. Nech 14, 21, 12, 15, 17 a 23 je náhodný výber rozsahu $n = 6$ z kvantitatívneho znaku X . Vypočítajte nasledujúce charakteristiky variability tohto kvantitatívneho znaku: variačné rozpätie, priemernú odchýlku, výberový rozptyl a výberovú smerodajnú odchýlku.

Riešenie. Variačné rozpätie $r = x_{\max} - x_{\min} = 23 - 12 = 11$. Na výpočet ďalších charakteristík variability budeme potrebovať výberový priemer $\bar{x}$:

$$\bar{x} = \frac{1}{6} \sum_{i=1}^6 x_i = \frac{1}{6} (14 + 21 + 12 + 15 + 17 + 23) = \frac{102}{6} = 17.$$

Pre priemernú odchýlku d potom máme

$$d = \frac{1}{6} \sum_{i=1}^6 |x_i - \bar{x}| = \frac{1}{6} (|14-17| + |21-17| + |12-17| + |15-17| + |17-17| + |23-17|) = \\ = \frac{1}{6} (3 + 4 + 5 + 2 + 0 + 6) = \frac{20}{6} = \frac{10}{3} \approx 3,3333.$$

Pre výberový rozptyl s^2 máme vzťah:

$$s^2 = \frac{1}{6} \sum_{i=1}^6 [(x_i - \bar{x})^2] = \frac{(14-17)^2 + (21-17)^2 + (12-17)^2 + (15-17)^2 + (17-17)^2 + (23-17)^2}{6} = \\ = \frac{3^2 + 4^2 + 5^2 + 2^2 + 0^2 + 6^2}{6} = \frac{90}{6} = \frac{30}{2} = 15.$$

Pre výberovú smerodajnú odchýlku s platí, že $s = \sqrt{s^2} = \sqrt{15} \approx 3,873$.

Príklad 27. Nech spojitý kvantitatívny znak X udáva dĺžku kovových hriadeľov v milimetroch. Nech je daný výber rozsahu 15 podľa tohto znaku:

25, 24, 24, 25, 26, 25, 25, 24, 24, 25, 24, 25, 26, 25, 25.

Vypočítajte nasledujúce charakteristiky variability tohto kvantitatívneho znaku: variačné rozpätie, priemernú odchýlku, výberový rozptyl a výberovú smerodajnú odchýlku.

Riešenie. Variačné rozpätie $r = x_{\max} - x_{\min} = 26 - 24 = 2$. Na výpočet ďalších charakteristík variability budeme potrebovať výberový priemer $\bar{x}$:

$$\bar{x} = \frac{1}{15} \sum_{i=1}^{15} x_i = \frac{1}{15} (5 \cdot 24 + 8 \cdot 25 + 2 \cdot 26) = 24,8.$$

Pre priemernú odchýlku d potom máme

$$d = \frac{1}{15} \sum_{i=1}^{15} |x_i - \bar{x}| = \frac{1}{15} (5 \cdot |24 - 24,8| + 8 \cdot |25 - 24,8| + 2 \cdot |26 - 24,8|) = \\ = \frac{1}{15} (5 \cdot 0,8 + 8 \cdot 0,2 + 2 \cdot 1,2) = \frac{4 + 1,6 + 2,4}{15} = \frac{7,8}{15} = \frac{2,6}{5} = 0,52.$$

Pre výberový rozptyl s^2 máme vzťah:

$$s^2 = \frac{1}{15} \sum_{i=1}^{15} [(x_i - \bar{x})^2] = \frac{1}{15} (5 \cdot (24 - 24,8)^2 + 8 \cdot (25 - 24,8)^2 + 2 \cdot (26 - 24,8)^2) = \\ = \frac{1}{15} (5 \cdot 0,8^2 + 8 \cdot 0,2^2 + 2 \cdot 1,2^2) = \frac{3,2 + 0,32 + 2,88}{15} = \frac{6,4}{15} = \frac{32}{75} \approx 0,4267.$$

Pre výberovú smerodajnú odchýlku s platí, že $s = \sqrt{s^2} = \sqrt{\frac{32}{75}} \approx 0,6532$. Pretože kvantitatívny

znak X bol meraný v milimetroch, z toho vyplýva, že aj všetky uvedené charakteristiky variability (okrem výberovej disperzie) sú uvedené v tejto jednotke dĺžky. Výberová disperzia je uvedená v mm^2 .

6 CHARAKTERISTIKY ŠIKMOSTI A ŠPICATOSTI

Príklad 28. Nech spojitý kvantitatívny znak X nadobudol hodnoty

-19, -18, -15, -8, -6, -4, 1, 2, 4, 5, 9, 15.

Vypočítajte jeho charakteristiky symetrie a koncentrácie.

Riešenie. Charakteristiky symetrie (šikmosti) sú kvantilová miera šikmosti s_K , Pearsonova miera šikmosti s_P a momentová miera šikmosti s_M . Charakteristikou koncentrácie (špicatosti) je momentová miera špicatosti k_M .

Výberové kvantily sú $\tilde{x}_{25} = -11,5$, $\tilde{x}_{50} = -1,5$ a $\tilde{x}_{75} = 4,5$. Preto

$$s_K = \frac{(\tilde{x}_{75} - \tilde{x}_{50}) - (\tilde{x}_{50} - \tilde{x}_{25})}{(\tilde{x}_{75} - \tilde{x}_{50}) + (\tilde{x}_{50} - \tilde{x}_{25})} = \frac{(4,5 - (-1,5)) - (-1,5 + 11,5)}{(4,5 - (-1,5)) + (-1,5 + 11,5)} = -\frac{1}{4} = -0,25.$$

Keďže $s_K < 0$, z toho vyplýva, že daný výber je zošikmený doprava podľa kvantilovej miery šikmosti.

Pearsonova miera šikmosti $s_P = \frac{\bar{x} - \hat{x}}{s}$ (kde $\bar{x}$ je výberový priemer, $\hat{x}$ je výberový modus a s je

výberová smerodajná odchýlka) neexistuje, lebo neexistuje výberový modus $\hat{x}$.

Na výpočet momentovej miery šikmosti budeme potrebovať výberový priemer $\bar{x}$ a výberovú smerodajnú odchýlku s .

$$\bar{x} = \frac{1}{12} \sum_{i=1}^{12} x_i = \frac{-19 - 18 - 15 - 8 - 6 - 4 + 1 + 2 + 4 + 5 + 9 + 15}{12} = \frac{-34}{12} = \frac{-17}{6} \approx 2,833.$$

$$s = \sqrt{\frac{1}{12} \sum_{i=1}^{12} (x_i - \bar{x})^2} \approx \sqrt{107} \approx 10,3.$$

Preto $s_M = \frac{1}{12s^3} \sum_{i=1}^{12} ((x_i - \bar{x})^3) \approx -0,106$. Keďže $s_M < 0$, z toho vyplýva, že daný výber je zošikmený doprava podľa momentovej miery šikmosti.

Momentovú mieru špicatosti vypočítame podľa vzorca $k_M = \frac{1}{12s^4} \sum_{i=1}^{12} ((x_i - \bar{x})^4) - 3$. Približne

vychádza hodnota $k_M \approx 1,092$. Keďže $k_M > 0$, z toho vyplýva, že daný výber je viac koncentrovaný ako referenčný výber s normálnym rozdelením.

Príklad 29. Nech je daný usporiadaný výber rozsahu 50 podľa kvantitatívneho znaku X : (-20, -20, -19, -19, -17, -16, -15, -15, -14, -14, -13, -13, -13, -10, -9, -9, -9, -9, -8, -8, -7, -7, -6, -5, -4, -3, -2, -1, -1, -1, 0, 0, 0, 1, 1, 1, 2, 2, 2, 2, 3, 3, 3, 4, 4, 5, 5, 5, 9, 10). Na jeho základe vypočítajte Pearsonovu mieru šikmosti znaku X .

Riešenie. Počet ohraničených polouzavretých intervalov, na ktoré daný súbor rozdelíme, bude po zaokrúhlení $m = \sqrt{n} = \sqrt{50} \approx 7$. Variačné rozpätie $r = x_{\max} - x_{\min} = 10 - (-20) = 30$. Budeme voliť ohraničené intervaly M_i , $i = 1, \dots, 7$ rovnakej dĺžky $d = r/m = 30/7 \approx 4,3$. Ak označíme stred intervalu M_i symbolom s_i , absolútnu početnosť intervalu M_i symbolom n_i , relatívnu početnosť intervalu M_i symbolom f_i , kumulatívnu absolútnu početnosť intervalu M_i sym-

bolom N_i a relatívnu kumulatívnu početnosť intervalu M_i symbolom F_i , dostaneme nasledujúcu veľkú tabuľku početností daného výberu:

i	M_i	s_i	d	n_i	f_i	N_i	F_i
1	$\langle -20; -15,7 \rangle$	-17,85	4,3	6	6/50	6	6/50
2	$\langle -15,7; -11,4 \rangle$	-13,55	4,3	7	7/50	13	13/50
3	$\langle -11,4; -7,1 \rangle$	-9,25	4,3	7	7/50	20	20/50
4	$\langle -7,1; -2,8 \rangle$	-4,95	4,3	6	6/50	26	26/50
5	$\langle -2,8; 1,5 \rangle$	-0,65	4,3	10	10/50	36	36/50
6	$\langle 1,5; 5,8 \rangle$	3,65	4,3	12	12/50	48	48/50
7	$\langle 5,8; 10 \rangle$	7,95	4,3	2	2/50	50	1

Z predchádzajúcej tabuľky vidíme, že najväčšiu absolútnu početnosť má interval M_6 . To znamená, že hľadaný výberový modus bude niekde v tomto intervale. Vieme, že $\hat{x} = -\frac{b}{2a}$, pričom čísla a, b spĺňajú nasledujúci systém troch lineárnych rovníc o troch neznámych a, b, c :

$$\begin{aligned}as_5^2 + bs_5 + c &= n_5 \\as_6^2 + bs_6 + c &= n_6 \\as_7^2 + bs_7 + c &= n_7\end{aligned}$$

Hodnoty $n_i, s_i, i = 5, 6, 7$ nájdeme vo veľkej tabuľke početností, čím dostaneme systém

$$\begin{aligned}(-0,65)^2 a - 0,65b + c &= 10 \\3,65^2 a + 3,65b + c &= 12 \\7,95^2 a + 7,95b + c &= 2\end{aligned}$$

Jeho riešením dostaneme približné hodnoty $a = -0,3285, b = 1,1567$; odtiaľ potom pre výberový modus dostávame aproximatívnu hodnotu $\hat{x} = 1,76$.

$$\bar{x} = \frac{1}{50} \sum_{i=1}^{50} x_i = \frac{2 \cdot (-20) + 2 \cdot (-19) - 17 - 16 + 2 \cdot (-15) + 2 \cdot (-14) + 3 \cdot (-13) + \dots + 9 + 10}{50} = -4,9.$$

$$s = \sqrt{\frac{1}{50} \sum_{i=1}^{50} ((x_i - \bar{x})^2)} = \sqrt{65,85} \approx 8,115.$$

Pre hľadajú Pearsonovu mieru šikmosti potom máme $s_p = \frac{\bar{x} - \hat{x}}{s} \approx \frac{-4,9 - 1,76}{8,115} \approx -0,82 < 0$.

Z toho vyplýva, že daný výber je podľa tejto miery zošikmený doprava.

Príklad 30. Nech medzi premennými x, y je lineárny vzťah $y = ax + b$. Nech pre 10 hodnôt nezávisle premennej x sme namerali 10 hodnôt závisle premennej y podľa nasledujúcej tabuľky:

x_i	10	20	30	40	50	60	70	80	90	100
y_i	19	38	62	75	96	119	143	158	181	203

Na jej základe vypočítajte výberový koeficient korelácie r_{XY} .

Riešenie. Výberový koeficient korelácie vypočítame podľa vzťahu

$$r_{XY} = \frac{\frac{1}{10} \sum_{i=1}^{10} ((x_i - \bar{x})(y_i - \bar{y}))}{\sqrt{\frac{1}{10} \sum_{i=1}^{10} ((x_i - \bar{x})^2)} \sqrt{\frac{1}{10} \sum_{i=1}^{10} ((y_i - \bar{y})^2)}},$$

kde

$$\bar{x} = \frac{1}{10} \sum_{i=1}^{10} x_i, \quad \bar{y} = \frac{1}{10} \sum_{i=1}^{10} y_i$$

sú výberové priemery znakov X, Y . Po vyčíslení vychádza $r_{XY} = 0,999155579$. To znamená, že medzi nameranými hodnotami veličín X, Y je veľká miera lineárnej závislosti, lebo r_{XY} je blízko jednotky. To, že výberový koeficient korelácie je v absolútnej hodnote rôzny od jednej signalizuje, že hodnoty y_i boli namerané s určitými chybami.

Príklad 31. Nech medzi premennými x, y je lineárny vzťah $y = ax + b$. Nech pre 10 hodnôt nezávisle premennej x sme namerali 10 hodnôt závisle premennej y podľa nasledujúcej tabuľky:

x_i	10	20	30	40	50	60	70	80	90	100
y_i	20	40	60	80	100	120	140	160	180	200

Na jej základe vypočítajte výberový koeficient korelácie r_{XY} .

Riešenie. Výberový koeficient korelácie vypočítame podľa vzťahu

$$r_{XY} = \frac{\frac{1}{10} \sum_{i=1}^{10} ((x_i - \bar{x})(y_i - \bar{y}))}{\sqrt{\frac{1}{10} \sum_{i=1}^{10} ((x_i - \bar{x})^2)} \sqrt{\frac{1}{10} \sum_{i=1}^{10} ((y_i - \bar{y})^2)}},$$

kde

$$\bar{x} = \frac{1}{10} \sum_{i=1}^{10} x_i, \quad \bar{y} = \frac{1}{10} \sum_{i=1}^{10} y_i$$

sú výberové priemery znakov X, Y . Po vyčíslení vychádza $r_{XY} = 1$. To signalizuje, že medzi nameranými hodnotami veličín X, Y je presná lineárna závislosť, t.j. namerané hodnoty y_i veličiny Y nie sú zaťažené chybami.

Príklad 32. Nech pre 5 hodnôt nezávisle premennej x sme namerali 5 hodnôt závisle premennej y podľa nasledujúcej tabuľky:

x_i	-1	-2	0	1	2
y_i	2	1	-6	2	1

Na jej základe vypočítajte výberový koeficient korelácie r_{XY} .

Riešenie. Pretože $\bar{x} = \bar{y} = 0$, pre výberový koeficient korelácie dostávame

$$r_{XY} = \frac{\frac{1}{5} \sum_{i=1}^5 ((x_i - \bar{x})(y_i - \bar{y}))}{\sqrt{\frac{1}{5} \sum_{i=1}^5 ((x_i - \bar{x})^2)} \sqrt{\frac{1}{5} \sum_{i=1}^5 ((y_i - \bar{y})^2)}} = \frac{\sum_{i=1}^5 (x_i y_i)}{\sqrt{\sum_{i=1}^5 (x_i^2)} \sqrt{\sum_{i=1}^5 (y_i^2)}} = \frac{0}{\sqrt{10} \sqrt{46}} = 0.$$

To podľa definície znamená, že premenné X, Y sú nekorelované.

7 LINEÁRNY MODEL A METÓDA NAJMENŠÍCH ŠTVORCOV

Príklad 33. Závislosť výšky y dieťaťa od veku x je v období piateho až desiateho roku lineárnou funkciou. Odhadnite túto funkciu $y(x) = \beta_1 + \beta_2 x$ metódou najmenších štvorcov, ak boli namerané údaje

x_i	5,0	5,5	6,0	6,5	7,0	7,5	8,0	8,5	9,0	9,5	10,0
y_i	105	112	119	123	129	134	139	142	148	152	157

pričom hodnoty x_i sú v uvedených rokoch a y_i v centimetroch. Tiež odhadnite výšku dieťaťa vo veku 7,2 roku.

Riešenie. Neznáme parametre β_1, β_2 určíme ako riešenia nasledujúceho normálneho systému rovníc

$$\begin{bmatrix} n, & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i, & \sum_{i=1}^n x_i^2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i y_i \end{bmatrix},$$

kde $n = 11$ je počet meraní, $\sum_{i=1}^{11} x_i = 82,5$, $\sum_{i=1}^{11} x_i^2 = 646,25$, $\sum_{i=1}^{11} y_i = 1460$, $\sum_{i=1}^{11} x_i y_i = 11227,5$. Po dosadení dostaneme systém

$$\begin{bmatrix} 11, & 82,5 \\ 82,5, & 646,25 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 1460 \\ 11227,5 \end{bmatrix},$$

ktorý má riešenie $\hat{\beta}_1 = 56,97$, $\hat{\beta}_2 = 10,08$. Strieška nad symbolmi β_1, β_2 značí, že sú to len odhady neznámych parametrov β_1, β_2 , pričom tieto odhady sme získali metódou najmenších štvorcov. Preto hľadanú lineárnu funkciu odhadneme formulou

$$y(x) = \beta_1 + \beta_2 x \approx \hat{\beta}_1 + \hat{\beta}_2 x = 56,97 + 10,08x.$$

Ak do posledného vzťahu dosadíme hodnotu $x = 7,2$, po vyčíslení dostaneme $y(7,2) = 129,7$. Odtiaľ odhadneme, že výška dieťaťa vo veku 7,2 roku bude asi 129,7 centimetra.

Príklad 34. Závislosť dĺžky y koľajnice od teploty x je v rozsahu teplôt -20°C až 50°C lineárnou funkciou. Odhadnite túto funkciu $y(x) = \beta_1 + \beta_2 x$ metódou najmenších štvorcov, ak boli namerané údaje

x_i	5	10	15	20	25	30	35	40
y_i	800,0	800,1	800,2	800,3	800,5	800,6	800,7	800,9

Hodnoty x_i sú v uvedené v °C a y_i v centimetroch. Odhadnite tiež dĺžku koľajnice pri teplote -10°C .

Riešenie. Neznáme parametre β_1, β_2 určíme ako riešenia nasledujúceho normálneho systému rovníc

$$\begin{bmatrix} n, & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i, & \sum_{i=1}^n x_i^2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i y_i \end{bmatrix},$$

kde $n = 8$ je počet meraní, $\sum_{i=1}^8 x_i = 180$, $\sum_{i=1}^8 x_i^2 = 5100$, $\sum_{i=1}^8 y_i = 6403,3$, $\sum_{i=1}^8 x_i y_i = 144101$. Po dosadení dostaneme systém

$$\begin{bmatrix} 8, & 180 \\ 180, & 5100 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 6403,3 \\ 144101 \end{bmatrix},$$

ktorý má riešenie $\hat{\beta}_1 = 799,841$, $\hat{\beta}_2 = 0,0254$. Strieška nad symbolmi β_1, β_2 značí, že sú to len odhady neznámych parametrov β_1, β_2 , pričom tieto odhady sme získali metódou naj-menších štvorcov. Preto hľadanú lineárnu funkciu odhadneme formulou

$$y(x) = \beta_1 + \beta_2 x \approx \hat{\beta}_1 + \hat{\beta}_2 x = 799,841 + 0,0254x.$$

Ak do nej dosadíme hodnotu $x = -10$, po vyčíslení dostaneme $y(-10) = 799,587$. Odtiaľ odhadneme, že dĺžka koľajnice pri teplote -10°C bude asi 799,6 centimetra.

Príklad 35. Závislosť ročného obchodného obratu y v štáte je lineárnou funkciou priemernej mzdy x za daný rok. Odhadnite túto funkciu $y(x) = \beta_1 + \beta_2 x$ metódou najmenších štvorcov, ak boli namerané údaje

x_i	10,5	11,0	11,5	12,0	12,5	13,0	13,5	14,0
y_i	300	320	318	342	356	382	403	418

pričom hodnoty x_i sú v uvedené v tisícoch a y_i v miliardách. Tiež odhadnite veľkosť ročného obchodného obratu pri priemernej mzde 16 tisíc.

Riešenie. Neznáme parametre β_1, β_2 určíme ako riešenia nasledujúceho normálneho systému rovníc

$$\begin{bmatrix} n, & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i, & \sum_{i=1}^n x_i^2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i y_i \end{bmatrix},$$

kde $n=8$ je počet meraní, $\sum_{i=1}^8 x_i = 98$, $\sum_{i=1}^8 x_i^2 = 28253,75$, $\sum_{i=1}^8 y_i = 2839$, $\sum_{i=1}^8 x_i y_i = 37978,5$. Po dosadení dostaneme systém

$$\begin{bmatrix} 8, & 98 \\ 98, & 28253,75 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 2839 \\ 37978,5 \end{bmatrix},$$

ktorý má riešenie $\hat{\beta}_1 = -67,13$, $\hat{\beta}_2 = 34,4$. Strieška nad symbolmi β_1 , β_2 značí, že sú to len odhady neznámych parametrov β_1 , β_2 , pričom tieto odhady sme získali metódou naj-menších štvorcov. Preto hľadanú lineárnu funkciu odhadneme formulou

$$y(x) = \beta_1 + \beta_2 x \approx \hat{\beta}_1 + \hat{\beta}_2 x = -67,13 + 34,4x.$$

Ak do nej dosadíme hodnotu $x = 16$, po vyčíslení dostaneme $y(16) = 484$. Odtiaľ odhadneme, že veľkosť ročného obchodného obratu pri priemernej mzde 16 tisíc bude asi 484 miliárd.

Príklad 36. Spotreba benzínu y je kvadraticky závislá od okamžitej rýchlosti x automobilu. Odhadnite túto závislosť $y(x) = \beta_1 + \beta_2 x + \beta_3 x^2$ metódou najmenších štvorcov, ak boli namerané údaje

x_i	50	55	65	75	80	90	100
y_i	4,8	5,2	5,9	6,4	7,9	10,0	12,4

Hodnoty x_i sú v uvedené v km/h a y_i v litroch na 100 kilometrov. Tiež odhadnite spotrebu benzínu pri rýchlosti 70 km/h.

Riešenie. Neznáme parametre β_1 , β_2 , β_3 určíme ako riešenia nasledujúceho normálneho systému rovníc

$$\begin{bmatrix} n, & \sum_{i=1}^n x_i, & \sum_{i=1}^n x_i^2, \\ \sum_{i=1}^n x_i, & \sum_{i=1}^n x_i^2, & \sum_{i=1}^n x_i^3, \\ \sum_{i=1}^n x_i^2, & \sum_{i=1}^n x_i^3, & \sum_{i=1}^n x_i^4, \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i y_i \\ \sum_{i=1}^n x_i^2 y_i \end{bmatrix},$$

kde $n = 7$ je počet meraní priemernej spotreby, $\sum_{i=1}^8 x_i = 515$, $\sum_{i=1}^8 x_i^2 = 39875$, $\sum_{i=1}^8 x_i^3 = 3228875$,

$\sum_{i=1}^8 x_i^4 = 271461875$, $\sum_{i=1}^8 y_i = 52,6$, $\sum_{i=1}^8 x_i y_i = 4161,5$, $\sum_{i=1}^8 x_i^2 y_i = 344217,5$. Po dosadení dostaneme systém

$$\begin{bmatrix} 7, & 515, & 39875 \\ 515, & 39875, & 3228875 \\ 39875, & 3228875, & 271146875 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 52,6 \\ 4161,5 \\ 344217,5 \end{bmatrix},$$

ktorý má riešenie $\hat{\beta}_1 = 11,43$, $\hat{\beta}_2 = -0,27$, $\hat{\beta}_3 = 0,0028$. Strieška nad symbolmi $\beta_1, \beta_2, \beta_3$ značí, že sú to odhady neznámych parametrov $\beta_1, \beta_2, \beta_3$, pričom tieto odhady sme získali metódou najmenších štvorcov. Odtiaľ potom hľadanú kvadratickú funkciu odhadneme formulou

$$y(x) = \beta_1 + \beta_2 x + \beta_3 x^2 \approx \hat{\beta}_1 + \hat{\beta}_2 x + \hat{\beta}_3 x^2 = 11,43 - 0,27x + 0,0028x^2.$$

Ak do nej dosadíme hodnotu $x = 70$, po vyčíslení dostaneme $y(70) = 6,3$. Odtiaľ odhadneme, že priemerná spotreba automobilu pri rýchlosti 70 km/h je asi 6,3 litra na 100 kilometrov.

8 METÓDA MAXIMÁLNEJ VIEROHODNOSTI A MOMENTOVÁ METÓDA

Príklad 37. Na základe náhodného výberu $x_1, \dots, x_n$ z rozdelenia pravdepodobnosti s hustotou $f(x, a) = 2axe^{-ax^2}$, $x \geq 0$, $f(x, a) = 0$, $x < 0$, odhadnite neznámy kladný parameter a metódou maximálnej vierohodnosti.

Riešenie. Z nezápornosti funkcie $f(x, a)$ na celom E_1 a z toho, že

$$\int_{-\infty}^{\infty} f(x, a) dx = \int_{-\infty}^0 0 dx + \int_0^{\infty} 2axe^{-ax^2} dx = \lim_{t \rightarrow \infty} \left[-e^{-ax^2} \right]_{x=0}^t = 1,$$

vyplýva, že funkcia $f(x, a)$ je hustotou nejakej náhodnej premennej X pre každú prípustnú hodnotu parametra a . Pre funkciu vierohodnosti, ktorá je funkciou premennej a , platí

$$L(a) = \prod_{i=1}^n f(x_i, a) = \prod_{i=1}^n (2ax_i e^{-ax_i^2}) = (2a)^n \left(\prod_{i=1}^n x_i \right) e^{-a \sum_{i=1}^n x_i^2},$$

lebo o číslach $x_1, \dots, x_n$ môžeme, vzhľadom na tvar hustoty, bez újmy na všeobecnosti predpokladať, že sú kladné. Zlogaritmovaním funkcie $L(a)$ dostaneme funkciu $\ell(a)$, kde

$$\ell(a) = \ln L(a) = n \ln(2a) + \sum_{i=1}^n \ln x_i - a \sum_{i=1}^n x_i^2.$$

Nutná podmienka existencie extrému funkcie jednej premennej dáva nasledujúcu rovnicu s neznámou a

$$0 = \frac{d\ell(a)}{da} = \frac{d}{da} \left(n \ln(2a) + \sum_{i=1}^n \ln x_i - a \sum_{i=1}^n x_i^2 \right) = \frac{n}{a} - \sum_{i=1}^n x_i^2,$$

lebo $x_1, \dots, x_n$ sú vzhľadom na premennú a konštanty. Riešením poslednej rovnice dostávame

$$a = \frac{1}{\frac{1}{n} \sum_{i=1}^n x_i^2} = \frac{n}{\sum_{i=1}^n x_i^2}.$$

Pretože

$$\left(\frac{d^2 \ell(a)}{da^2} \right)_{a = \frac{n}{\sum_{i=1}^n x_i^2}} = \left(-\frac{n}{a^2} \right)_{a = \frac{n}{\sum_{i=1}^n x_i^2}} = -\frac{n}{\left(\frac{n}{\sum_{i=1}^n x_i^2} \right)^2} < 0,$$

z toho vyplýva, že funkcia $\ell(a)$ nadobudne v bode $a = \frac{n}{\sum_{i=1}^n x_i^2}$ svoje maximum. Preto číslo

$\hat{a} = \frac{n}{\sum_{i=1}^n x_i^2}$ je hľadaným odhadom neznámeho parametra a získaným metódou maximálnej vierohodnosti.

Príklad 38. Na základe náhodného výberu $x_1, \dots, x_n$ z rozdelenia pravdepodobnosti s hustotou $f(x, a) = ax^{-a-1}$, $x \geq 1$, $f(x, a) = 0$, $x < 1$, odhadnite neznámy kladný parameter a metódou maximálnej vierohodnosti.

Riešenie. Z nezápornosti funkcie $f(x, a)$ na celom E_1 a z toho, že

$$\int_{-\infty}^{\infty} f(x, a) dx = \int_{-\infty}^1 0 dx + \int_1^{\infty} ax^{-a-1} dx = \lim_{t \rightarrow \infty} \left[-x^{-a} \right]_{x=1}^t = 1,$$

vyplýva, že funkcia $f(x, a)$ je hustotou nejakej náhodnej premennej X pre každú prípustnú hodnotu parametra a . Pre funkciu vierohodnosti, ktorá je funkciou premennej a , platí

$$L(a) = \prod_{i=1}^n f(x_i, a) = \prod_{i=1}^n (ax_i^{-a-1}) = a^n \left(\prod_{i=1}^n x_i \right)^{-a-1},$$

lebo o číslach $x_1, \dots, x_n$ môžeme, vzhľadom na tvar hustoty, bez újmy na všeobecnosti predpokladať, že sú väčšie ako jedna. Zlogaritmovaním funkcie $L(a)$ dostaneme funkciu $\ell(a)$, kde

$$\ell(a) = \ln L(a) = n \ln a - (a+1) \ln \prod_{i=1}^n x_i.$$

Nutná podmienka existencie extrému funkcie jednej premennej dáva nasledujúcu rovnicu s neznámou a

$$0 = \frac{d\ell(a)}{da} = \frac{d}{da} \left(n \ln a - (a+1) \ln \prod_{i=1}^n x_i \right) = \frac{n}{a} - \ln \prod_{i=1}^n x_i,$$

lebo $x_1, \dots, x_n$ sú vzhľadom na premennú a konštanty. Riešením poslednej rovnice dostávame

$$a = \frac{n}{\ln \prod_{i=1}^n x_i}.$$

Pretože

$$\left(\frac{d^2 \ell(a)}{da^2} \right)_{a=\frac{n}{\sum_{i=1}^n x_i^2}} = \left(-\frac{n}{a^2} \right)_{a=\frac{n}{\ln \prod_{i=1}^n x_i}} = -\frac{n}{\left(\frac{n}{\ln \prod_{i=1}^n x_i} \right)^2} < 0,$$

z toho vyplýva, že funkcia $\ell(a)$ nadobudne v bode $a = \frac{n}{\sum_{i=1}^n x_i^2}$ svoje maximum. Preto číslo

$\hat{a} = \frac{n}{\ln \prod_{i=1}^n x_i}$ je hľadaným odhadom neznámeho parametra a získaným metódou maximálnej vierohodnosti.

Príklad 39. Na základe náhodného výberu $x_1, \dots, x_n$ z rozdelenia pravdepodobnosti s hustotou $f(x, a) = ax^{a-1}$, $x \in \langle 0; 1 \rangle$, $f(x, a) = 0$, $x \in E_1 - \langle 0; 1 \rangle$, odhadnite neznámy kladný parameter a momentovou metódou.

Riešenie. Z nezápornosti funkcie $f(x, a)$ na celom E_1 a z toho, že

$$\int_{-\infty}^{\infty} f(x, a) dx = \int_{-\infty}^0 0 dx + \int_0^1 ax^{a-1} dx + \int_1^{\infty} 0 dx = \left[x^a \right]_{x=0}^1 = 1$$

vyplýva, že funkcia $f(x, a)$ je hustotou nejakej náhodnej premennej X pre každú prípustnú hodnotu parametra a . Pre hodnotu prvého počiatočného momentu ν_1 máme

$$\begin{aligned} \nu_1 = E(X) &= \int_{-\infty}^{\infty} xf(x, a) dx = \int_{-\infty}^0 0 dx + \int_0^1 xax^{a-1} dx + \int_1^{\infty} 0 dx = \int_0^1 ax^a dx = \\ &= \left[\frac{a}{a+1} x^{a+1} \right]_{x=0}^1 = \frac{a}{a+1} = \frac{a+1-1}{a+1} = 1 - \frac{1}{a+1}. \end{aligned}$$

Keď v poslednom vzťahu nahradíme prvý počiatočný moment ν_1 výberovým počiatočným momentom $m_1(0) = \frac{1}{n} \sum_{i=1}^n x_i$ a parameter a symbolom $\hat{a}$, dostaneme nasledujúcu rovnicu s neznámou $\hat{a}$:

$$m_1(0) = 1 - \frac{1}{\hat{a} + 1}.$$

Jej riešením dostaneme $\hat{a} = \frac{m_1(0)}{1 - m_1(0)} = \frac{\frac{1}{n} \sum_{i=1}^n x_i}{1 - \frac{1}{n} \sum_{i=1}^n x_i}$. Číslo $\hat{a}$ je hľadaným odhadom neznámeho

parametra a získaným momentovou metódou.

Príklad 40. Na základe náhodného výberu $x_1, \dots, x_n$ z rozdelenia pravdepodobnosti s hustotou $f(x, a) = 2axe^{-ax^2}$, $x \geq 0$, $f(x, a) = 0$, $x < 0$, odhadnite neznámy kladný parameter a momentovou metódou.

Riešenie. Z nezápornosti funkcie $f(x, a)$ na celom E_1 a z toho, že

$$\int_{-\infty}^{\infty} f(x, a) dx = \int_{-\infty}^0 0 dx + \int_0^{\infty} 2axe^{-ax^2} dx = \lim_{t \rightarrow \infty} \left[-e^{-ax^2} \right]_{x=0}^t = 1$$

vyplýva, že funkcia $f(x, a)$ je hustotou nejakej náhodnej premennej X pre každú prípustnú hodnotu parametra a . Pre hodnotu prvého počiatočného momentu ν_1 máme

$$\nu_1 = E(X) = \int_{-\infty}^{\infty} xf(x, a) dx = \int_{-\infty}^0 0 dx + \int_0^{\infty} x2axe^{-ax^2} dx.$$

Na výpočet posledného integrálu použijeme metódu per partes. Ak $u = x$, $v' = 2axe^{-ax^2}$, potom $u' = 1$, $v = -e^{-ax^2}$. Odtiaľ máme

$$\int_0^{\infty} x2axe^{-ax^2} dx = \left[x(-e^{-ax^2}) \right]_{x=0}^{\infty} - \int_0^{\infty} -e^{-ax^2} dx = \int_0^{\infty} e^{-ax^2} dx,$$

lebo (keďže $a > 0$) L'Hospitalovo pravidlo dáva $\lim_{x \rightarrow \infty} [x(-e^{-ax^2})] = \lim_{x \rightarrow \infty} \frac{-x}{e^{ax^2}} = \lim_{x \rightarrow \infty} \frac{-1}{2axe^{ax^2}} = 0$.

Posledný integrál riešime substitúciou $s = \sqrt{2a} x$, ktorá vzhľadom na známy Laplaceov (nevlastný) integrál $\int_0^{\infty} e^{-\frac{x^2}{2}} dx = \sqrt{\frac{\pi}{2}}$ dáva

$$\nu_1 = \int_0^{\infty} e^{-ax^2} dx = \frac{1}{\sqrt{2a}} \int_0^{\infty} e^{-\frac{s^2}{2}} ds = \frac{1}{2} \sqrt{\frac{\pi}{a}}.$$

Keď v poslednom vzťahu nahradíme prvý počiatočný moment ν_1 výberovým počiatočným momentom $m_1(0) = \frac{1}{n} \sum_{i=1}^n x_i$ a parameter a symbolom $\hat{a}$, dostaneme nasledujúcu rovnicu s neznámou $\hat{a}$:

$$m_1(0) = \frac{1}{2} \sqrt{\frac{\pi}{\hat{a}}}.$$

Jej riešením dostaneme $\hat{a} = \frac{\pi}{4(m_1(0))^2} = \frac{\pi}{4\left(\frac{1}{n} \sum_{i=1}^n x_i\right)^2}$. Číslo $\hat{a}$ je hľadaným odhadom neznámeho parametra a získaným momentovou metódou.

9 INTERVALY SPOL' AHLIVOSTI

Príklad 41. Merali sme priemer včkového hriadeľa na $n = 100$ súčiastkach. Predpokladáme, že priemer včkového hriadeľa je spojitá náhodná premenná s normálnym rozdelením s neznámymi parametrami μ a σ^2 . Určte všetky intervaly spoľahlivosti pre parametre μ, σ, σ^2 so spoľahlivosťou $1 - \alpha = 0,95$, ak $\bar{x}_{100} = 8,45 \text{ mm}, s_{100}^2 = 1,24 \text{ mm}^2$.

Riešenie. Máme určiť celkovo 9 intervalov spoľahlivosti; po tri pre každý z parametrov μ, σ, σ^2 . Podľa príslušných vzorcov dostávame:

a1. Obojstranný interval spoľahlivosti pre parameter μ .

Na jeho určenie popri číslach $n, \bar{x}_n, s_n$ potrebujeme ešte kvantil t -rozdelenia $t_{n-1}^{kv}\left(1 - \frac{\alpha}{2}\right)$.

Nájdeme ho v štatistických tabuľkách, odkiaľ dostaneme hodnotu $t_{99}^{kv}(0,975) = 1,98$. Preto hľadaný interval má tvar

$$\begin{aligned} \left\langle \bar{x}_{100} - \frac{s_{100} \cdot t_{99}^{kv}(0,975)}{\sqrt{100}}; \bar{x}_{100} + \frac{s_{100} \cdot t_{99}^{kv}(0,975)}{\sqrt{100}} \right\rangle = \\ = \left\langle 8,45 - \frac{1,11 \cdot 1,98}{10}; 8,45 + \frac{1,11 \cdot 1,98}{10} \right\rangle = \langle 8,22; 8,67 \rangle. \end{aligned}$$

a2. Dolný (tiež pravostranný) interval spoľahlivosti pre parameter μ .

Na jeho určenie popri číslach $n, \bar{x}_n, s_n$ potrebujeme ešte kvantil t -rozdelenia $t_{n-1}^{kv}(1 - \alpha)$. Nájdeme ho v štatistických tabuľkách, odkiaľ dostaneme hodnotu $t_{99}^{kv}(0,95) = 1,66$. Preto hľadaný interval má tvar

$$\left\langle \bar{x}_{100} - \frac{s_{100} \cdot t_{99}^{kv}(0,95)}{\sqrt{100}}; \infty \right\rangle = \left\langle 8,45 - \frac{1,11 \cdot 1,66}{\sqrt{100}}; \infty \right\rangle = \langle 8,27; \infty \rangle.$$

a3. Horný (tiež ľavostranný) interval spoľahlivosti pre parameter μ .

Na jeho určenie popri číslach $n, \bar{x}_n, s_n$ potrebujeme ešte kvantil t -rozdelenia $t_{n-1}^{kv}(1 - \alpha)$. Nájdeme ho v štatistických tabuľkách, odkiaľ dostaneme hodnotu $t_{99}^{kv}(0,95) = 1,66$. Preto hľadaný interval má tvar

$$\left\langle -\infty; \bar{x}_{100} + \frac{s_{100} \cdot t_{99}^{kv}(0,95)}{\sqrt{100}} \right\rangle = \left\langle -\infty; 8,45 + \frac{1,11 \cdot 1,66}{\sqrt{100}} \right\rangle = \langle -\infty; 8,63 \rangle.$$

b1. Obojstranný interval spoľahlivosti pre parameter σ^2 .

Na jeho určenie popri číslach $n, \bar{x}_n, s_n^2$ potrebujeme ešte dva kvantily χ^2 -rozdelenia $\chi_{n-1}^{2,kv}\left(\frac{\alpha}{2}\right), \chi_{n-1}^{2,kv}\left(1 - \frac{\alpha}{2}\right)$. Nájdeme ich v štatistických tabuľkách, odkiaľ dostaneme hodnoty

$\chi_{99}^{2,kv}(0,025) = 73,4$, $\chi_{99}^{2,kv}(0,975) = 128,4$. Preto hľadaný interval má tvar

$$\left\langle \frac{(n-1) \cdot s_n^2}{\chi_{n-1}^{2,kv}(0,975)}; \frac{(n-1) \cdot s_n^2}{\chi_{n-1}^{2,kv}(0,025)} \right\rangle = \left\langle \frac{99 \cdot 1,24}{128,4}; \frac{99 \cdot 1,24}{73,4} \right\rangle = \langle 0,95; 1,67 \rangle.$$

b2. Dolný (tiež pravostranný) interval spoľahlivosti pre parameter σ^2 .

Na jeho určenie popri číslach n , $\bar{x}_n$, s_n potrebujeme ešte kvantil χ^2 -rozdelenia $\chi_{n-1}^{2,kv}(1-\alpha)$. Nájďme ho v štatistických tabuľkách, odkiaľ dostaneme nasledujúcu hodnotu $\chi_{99}^{2,kv}(0,95) = 123,2$. Preto hľadaný interval má tvar

$$\left\langle \frac{(n-1) \cdot s_n^2}{\chi_{n-1}^{2,kv}(0,95)}; \infty \right\rangle = \left\langle \frac{99 \cdot 1,24}{123,2}; \infty \right\rangle = \langle 0,996; \infty \rangle.$$

b3. Horný (tiež ľavostranný) interval spoľahlivosti pre parameter σ^2 .

Na jeho určenie popri číslach n , $\bar{x}_n$, s_n potrebujeme ešte kvantil χ^2 -rozdelenia $\chi_{n-1}^{2,kv}(\alpha)$. Nájďme ho v štatistických tabuľkách, odkiaľ dostaneme, že $\chi_{99}^{2,kv}(0,05) = 77,0$. Preto hľadaný interval má tvar

$$\left\langle 0; \frac{(n-1) \cdot s_n^2}{\chi_{n-1}^{2,kv}(0,05)} \right\rangle = \left\langle 0; \frac{99 \cdot 1,24}{77,0} \right\rangle = \langle 0; 1,59 \rangle.$$

Ostávajúce tri intervaly spoľahlivosti pre parameter σ získame odmocnením číselných hodnôt v príslušných intervaloch spoľahlivosti pre parameter σ^2 . Vychádza tak

c1. Obojstranný interval spoľahlivosti pre parameter σ .

$$\left\langle \sqrt{\frac{(n-1) \cdot s_n^2}{\chi_{n-1}^{2,kv}(0,975)}}; \sqrt{\frac{(n-1) \cdot s_n^2}{\chi_{n-1}^{2,kv}(0,025)}} \right\rangle = \left\langle \sqrt{\frac{99 \cdot 1,24}{128,4}}; \sqrt{\frac{99 \cdot 1,24}{73,4}} \right\rangle = \langle 0,97; 1,29 \rangle.$$

c2. Dolný (tiež pravostranný) interval spoľahlivosti pre parameter σ .

$$\left\langle \sqrt{\frac{(n-1) \cdot s_n^2}{\chi_{n-1}^{2,kv}(0,95)}}; \infty \right\rangle = \left\langle \sqrt{\frac{99 \cdot 1,24}{123,2}}; \infty \right\rangle = \langle 0,998; \infty \rangle.$$

c3. Horný (tiež ľavostranný) interval spoľahlivosti pre parameter σ .

$$\left\langle 0; \sqrt{\frac{(n-1) \cdot s_n^2}{\chi_{n-1}^{2,kv}(0,05)}} \right\rangle = \left\langle 0; \sqrt{\frac{99 \cdot 1,24}{77,0}} \right\rangle = \langle 0; 1,26 \rangle.$$

Príklad 42. Spojitá náhodná premenná má normálne rozdelenie s neznámymi parametrami μ a σ^2 . Namerali sme hodnoty tejto náhodnej premennej a zapísali do nasledujúcej tabuľky

x_i	9	11	12	14	15	16	17	18	20	21
p_i	1	2	3	4	7	5	4	3	2	1

kde x_i sú namerané hodnoty a p_i ich absolútne početnosti.

Riešenie. Najprv určíme neznáme hodnoty n , $\bar{x}_n$, s_n a s_n^2 . Po kratšom výpočte vychádza $n = 32$, $\bar{x}_{32} = 15,34$, $s_{32} = 2,719$ a $s_{32}^2 = 7,394$. Podobne ako v predchádzajúcom príklade potom určíme hľadaných 9 intervalov spoľahlivosti pre μ, σ, σ^2 .

a1. Obojstranný interval spoľahlivosti pre parameter μ .

Na jeho určenie popri číslach n , $\bar{x}_n$, s_n potrebujeme ešte kvantil t -rozdelenia $t_{n-1}^{kv}\left(1 - \frac{\alpha}{2}\right)$.

Nájdeme ho v štatistických tabuľkách, odkiaľ dostaneme hodnotu $t_{31}^{kv}(0,975) = 2,04$. Preto

$$\left\langle \bar{x}_{32} - \frac{s_{32} \cdot t_{31}^{kv}(0,975)}{\sqrt{32}}; \bar{x}_{32} + \frac{s_{32} \cdot t_{31}^{kv}(0,975)}{\sqrt{32}} \right\rangle = \left\langle 15,34 - \frac{2,719 \cdot 2,04}{5,657}; 15,34 + \frac{2,719 \cdot 2,04}{5,657} \right\rangle = \langle 14,36; 16,32 \rangle.$$

a2. Dolný (tiež pravostranný) interval spoľahlivosti pre parameter μ .

Na jeho určenie popri číslach n , $\bar{x}_n$, s_n potrebujeme ešte kvantil t -rozdelenia $t_{n-1}^{kv}(1 - \alpha)$. Nájdeme ho v štatistických tabuľkách, odkiaľ dostaneme hodnotu $t_{31}^{kv}(0,95) = 1,70$. Preto hľadaný interval má tvar

$$\left\langle \bar{x}_{32} - \frac{s_{32} \cdot t_{31}^{kv}(0,95)}{\sqrt{32}}; \infty \right\rangle = \left\langle 15,34 - \frac{2,719 \cdot 1,70}{5,657}; \infty \right\rangle = \langle 8,27; \infty \rangle.$$

a3. Horný (tiež ľavostranný) interval spoľahlivosti pre parameter μ .

Na jeho určenie popri číslach n , $\bar{x}_n$, s_n potrebujeme ešte kvantil t -rozdelenia $t_{n-1}^{kv}(1 - \alpha)$. Nájdeme ho v štatistických tabuľkách, odkiaľ dostaneme hodnotu $t_{31}^{kv}(0,95) = 1,70$. Preto hľadaný interval má tvar

$$\left\langle -\infty; \bar{x}_{32} + \frac{s_{32} \cdot t_{31}^{kv}(0,95)}{\sqrt{32}} \right\rangle = \left\langle -\infty; 15,34 + \frac{2,719 \cdot 1,70}{5,657} \right\rangle = \langle -\infty; 16,16 \rangle.$$

b1. Obojstranný interval spoľahlivosti pre parameter σ^2 .

Na jeho určenie popri číslach n , $\bar{x}_n$, s_n^2 potrebujeme ešte dva kvantily χ^2 -rozdelenia $\chi_{n-1}^{2,kv}\left(\frac{\alpha}{2}\right)$, $\chi_{n-1}^{2,kv}\left(1 - \frac{\alpha}{2}\right)$. Nájdeme ich v štatistických tabuľkách, odkiaľ dostaneme hodnoty

$\chi_{31}^{2,kv}(0,025) = 17,5$, $\chi_{31}^{2,kv}(0,975) = 48,2$. Preto hľadaný interval má tvar

$$\left\langle \frac{(n-1) \cdot s_n^2}{\chi_{n-1}^{2,kv}(0,975)}; \frac{(n-1) \cdot s_n^2}{\chi_{n-1}^{2,kv}(0,025)} \right\rangle = \left\langle \frac{31 \cdot 7,394}{48,2}; \frac{31 \cdot 7,394}{17,5} \right\rangle = \langle 4,75; 13,10 \rangle.$$

b2. Dolný (tiež pravostranný) interval spoľahlivosti pre parameter σ^2 .

Na jeho určenie popri číslach n , $\bar{x}_n$, s_n potrebujeme ešte kvantil χ^2 -rozdelenia $\chi_{n-1}^{2,kv}(1-\alpha)$. Nájdeme ho v štatistických tabuľkách, odkiaľ dostaneme nasledujúcu hodnotu kvantilu $\chi_{31}^{2,kv}(0,95) = 45,0$. Preto hľadaný interval má tvar

$$\left\langle \frac{(n-1) \cdot s_n^2}{\chi_{n-1}^{2,kv}(0,95)}; \infty \right\rangle = \left\langle \frac{31 \cdot 7,394}{45,0}; \infty \right\rangle = \langle 5,09; \infty \rangle.$$

b3. Horný (tiež ľavostranný) interval spoľahlivosti pre parameter σ^2 .

Na jeho určenie popri číslach n , $\bar{x}_n$, s_n potrebujeme ešte kvantil χ^2 -rozdelenia $\chi_{n-1}^{2,kv}(\alpha)$. Nájdeme ho v štatistických tabuľkách, odkiaľ dostaneme, že $\chi_{31}^{2,kv}(0,05) = 19,3$. Preto hľadaný interval má tvar

$$\left\langle 0; \frac{(n-1) \cdot s_n^2}{\chi_{n-1}^{2,kv}(0,05)} \right\rangle = \left\langle 0; \frac{31 \cdot 7,394}{19,3} \right\rangle = \langle 0; 11,88 \rangle.$$

Ostávajúce tri intervaly spoľahlivosti pre parameter σ získame odmocnením číselných hodnôt v príslušných intervaloch spoľahlivosti pre parameter σ^2 . Vychádza tak

c1. Obojstranný interval spoľahlivosti pre parameter σ .

$$\left\langle \sqrt{\frac{(n-1) \cdot s_n^2}{\chi_{n-1}^{2,kv}(0,975)}}; \sqrt{\frac{(n-1) \cdot s_n^2}{\chi_{n-1}^{2,kv}(0,025)}} \right\rangle = \left\langle \sqrt{\frac{31 \cdot 7,394}{48,2}}; \sqrt{\frac{31 \cdot 7,394}{17,5}} \right\rangle = \langle 2,18; 3,62 \rangle.$$

c2. Dolný (tiež pravostranný) interval spoľahlivosti pre parameter σ .

$$\left\langle \sqrt{\frac{(n-1) \cdot s_n^2}{\chi_{n-1}^{2,kv}(0,95)}}; \infty \right\rangle = \left\langle \sqrt{\frac{31 \cdot 7,394}{45,0}}; \infty \right\rangle = \langle 2,26; \infty \rangle.$$

c3. Horný (tiež ľavostranný) interval spoľahlivosti pre parameter σ .

$$\left\langle 0; \sqrt{\frac{(n-1) \cdot s_n^2}{\chi_{n-1}^{2,kv}(0,05)}} \right\rangle = \left\langle 0; \sqrt{\frac{31 \cdot 7,394}{19,3}} \right\rangle = \langle 0; 3,45 \rangle.$$

10 TESTOVANIE HYPOTÉZ

Príklad 43. Výšku dievčat vo veku 9,5 až 10 rokov možno považovať za normálne rozdelenú náhodnú premennú X . Nech celoštátny priemer výšky dievčat je $\mu_0 = 131$ cm. V okrese Liptovský Mikuláš sme náhodne vybrali 20 dievčat v danom veku a odmerali ich výšku. Dostali sme tak nasledujúce hodnoty x_i v centimetroch, pričom p_i sú absolútne početnosti týchto hodnôt:

x_i	124	126	127	128	129	130	131	133	136
p_i	1	3	6	4	1	1	1	1	1

Na hladine $\alpha = 0,05$ otestujte, či je priemerná výška dievčat v danom veku v okrese Liptovský Mikuláš taká istá ako celoštátny priemer. Ak nie je, zistite, či je menšia alebo väčšia.

Riešenie. a) Najprv otestujeme, či je priemerná výška dievčat v danom veku v okrese Liptovský Mikuláš taká istá ako celoštátny priemer. Matematicky to znamená, že máme otestovať dvojicu hypotéz $H_0 : \mu = 131 \not\asymp H_1 : \mu \neq 131$. Ak je výška dievčat v okrese Liptovský Mikuláš taká istá ako celoštátny priemer, znamená to platnosť nulovej hypotézy $H_0 : \mu = 131$. Ak to tak nie je, znamená to platnosť alternatívnej hypotézy $H_1 : \mu \neq 131$. Teraz zostrojíme kritickú oblasť W pre test tejto dvojice hypotéz. Platí, že

$$W = \left\{ (x_1, \dots, x_{20}) \in R^{20} : \frac{|\bar{x}_{20} - 131|}{s_{20}} \sqrt{20} > t_{19}^{kv}(0,975) \right\}.$$

V našom prípade $\bar{x}_{20} = 128,3$ cm, $s_{20} = 2,754$ cm, $t_{19}^{kv}(0,975) = 2,09$. Z toho vyplýva, že platí

$$\frac{|\bar{x}_{20} - 131|}{s_{20}} \sqrt{20} = \frac{|128,3 - 131|}{2,754} \sqrt{20} = 4,38 > 2,09 = t_{19}^{kv}(0,975).$$

Preto $(x_1, \dots, x_{20}) = (124, 126, 126, 126, 127, \dots, 133, 136) \in W$. Z toho vyplýva, že zamietame hypotézu $H_0 : \mu = 131$ a prijímame hypotézu $H_1 : \mu \neq 131$, čo slovne formulujeme nasledovne: *Bolo štatisticky preukázané, že priemerná výška dievčat vo veku 9,5 až 10 rokov v okrese Liptovský Mikuláš je iná ako celoštátny priemer.*

b) Pretože $128,3 < 131$, budeme testovať, či priemerná výška dievčat vo veku 9,5 až 10 rokov v okrese Liptovský Mikuláš je menšia ako celoštátny priemer. Matematicky to znamená, že máme otestovať dvojicu hypotéz $H_0 : \mu = 131 \not\asymp H_1 : \mu < 131$. Ak je výška dievčat v okrese Liptovský Mikuláš taká istá ako celoštátny priemer, znamená to platnosť nulovej hypotézy $H_0 : \mu = 131$. Ak to tak nie je, znamená to platnosť alternatívnej hypotézy $H_1 : \mu < 131$. Teraz zostrojíme kritickú oblasť W pre test tejto dvojice hypotéz. Platí, že

$$W = \left\{ (x_1, \dots, x_{20}) \in R^{20} : \frac{\bar{x}_{20} - 131}{s_{20}} \sqrt{20} < -t_{19}^{kv}(0,95) \right\}.$$

V našom prípade $t_{19}^{kv}(0,95) = 1,73$. Z toho vyplýva, že platí

$$\frac{\bar{x}_{20} - 131}{s_{20}} \sqrt{20} = \frac{128,3 - 131}{2,754} \sqrt{20} = -4,38 < -1,73 = -t_{19}^{kv}(0,95).$$

Preto $(x_1, \dots, x_{20}) = (124, 126, 126, 126, 127, \dots, 133, 136) \in W$. Z toho vyplýva, že zamietame hypotézu $H_0 : \mu = 131$ a prijímame hypotézu $H_1 : \mu < 131$, čo slovne formulujeme nasledovne: *Bolo štatisticky preukázané, že priemerná výška dievčat vo veku 9,5 až 10 rokov v okrese Liptovský Mikuláš je menšia ako celoštátny priemer.*

Príklad 44. V petrochemickej firme sa plní motorový olej do litrových nádob. Plniaca linka je nastavená správne, ak plní olej do nádob na priemernú hodnotu $\mu_0 = 1$ liter s odchýlkou $\sigma_0 \leq 0,01$ litra. Na hladine $\alpha = 0,05$ otestujte, či je plniaca linka nastavená správne vzhľadom na hodnoty μ_0, σ_0 , ak sme namerali nasledujúce údaje

x_i	0,96	0,97	0,98	0,99	1,00	1,02	1,05
p_i	1	12	65	8	11	2	1

pričom hodnoty x_i sú namerané v litroch (s presnosťou na dve desatinné miesta) a p_i sú absolútne početnosti týchto hodnôt.

Riešenie. a) Najprv otestujeme, či je plniaca linka nastavená správne vzhľadom na priemerné množstvo oleja v nádobách. Matematicky to znamená, že máme otestovať dvojicu hypotéz $H_0 : \mu = 1 \not\asymp H_1 : \mu \neq 1$. Ak je plniaca linka nastavená správne vzhľadom na priemerné množstvo oleja v nádobách, znamená to platnosť nulovej hypotézy $H_0 : \mu = 1$. Ak to tak nie je, znamená to platnosť alternatívnej hypotézy $H_1 : \mu \neq 1$. Teraz zostrojíme kritickú oblasť W pre test tejto dvojice hypotéz. Platí, že

$$W = \left\{ (x_1, \dots, x_{100}) \in R^{100} : \frac{|\bar{x}_{100} - 1|}{s_{100}} \sqrt{100} > t_{99}^{kv}(0,975) \right\}.$$

V našom prípade $\bar{x}_{100} = 0,9831$ litra, $s_{20} = 0,012$ litra, $t_{99}^{kv}(0,975) = 1,98$. Z toho vyplýva, že platí

$$\frac{|\bar{x}_{100} - 1|}{s_{100}} \sqrt{100} = \frac{|0,9831 - 1|}{0,012} \sqrt{100} = 14,08 > 1,98 = t_{99}^{kv}(0,975).$$

Preto $(x_1, \dots, x_{100}) = (0,96; 0,97; 0,97; 0,97; \dots; 1,02; 1,05) \in W$. Z toho vyplýva, že zamietame hypotézu $H_0 : \mu = 1$ a prijímame hypotézu $H_1 : \mu \neq 1$, čo slovne formulujeme nasledovne: *Bolo štatisticky preukázané, že plniaca linka neplní olej do nádob na deklarovanú priemernú hodnotu 1 liter.*

b) Ukázali sme, že plniaca linka neplní olej do nádob na deklarovanú priemernú hodnotu 1 liter. Z hľadiska správcu plniacej linky to znamená, že musí zaistiť jej opravu. Z hľadiska spotrebiteľa je však dôležité, či ho firma neokráda, t.j. či nedáva do nádob priemerne menej oleja

ako deklaruje. Pretože $0,9831 < 1$, kontrolór zastupujúci záujmy verejnosti bude testovať, či plniaca linka plní olej do nádob na menšiu hodnotu ako deklarovaný 1 liter alebo nie. Matematicky to znamená, že bude testovať dvojicu hypotéz $H_0: \mu = 1 \not\propto H_1: \mu < 1$. Ak je priemerné množstvo oleja v nádobách rovné 1 liter, znamená to platnosť nulovej hypotézy $H_0: \mu = 1$. Ak to tak nie je, znamená to platnosť alternatívnej hypotézy $H_1: \mu < 1$. Teraz zostrojíme kritickú oblasť W pre test tejto dvojice hypotéz. Platí, že

$$W = \left\{ (x_1, \dots, x_{100}) \in R^{100} : \frac{\bar{x}_{100} - 1}{s_{100}} \sqrt{100} < -t_{99}^{kv}(0,95) \right\}.$$

V našom prípade $t_{99}^{kv}(0,95) = 1,67$. Z toho vyplýva, že platí

$$\frac{\bar{x}_{100} - 1}{s_{100}} \sqrt{100} = \frac{0,9831 - 1}{0,012} \sqrt{100} = -14,08 < -1,67 = -t_{99}^{kv}(0,95).$$

Preto $(x_1, \dots, x_{100}) = (124, 126, 126, 126, 127, \dots, 133, 136) \in W$. Z toho vyplýva, že zamietame hypotézu $H_0: \mu = 1$ a prijímame hypotézu $H_1: \mu < 1$, čo slovne formulujeme nasledovne: *Bolo štatisticky preukázané, že petrochemická firma plní olej do nádob na priemernú hodnotu menšiu ako deklarovaný 1 liter.* To znamená, že firma okráda spotrebiteľa a preto dostane primeranú pokutu.

c) Nakoniec otestujeme, či plniaca linka neprekračuje povolenú hodnotu rozptylu priemerného množstva oleja v nádobách. Matematicky to znamená, že máme otestovať dvojicu hypotéz $H_0: \sigma = 0,01 \not\propto H_1: \sigma > 0,01$. Ak je plniaca linka nastavená správne vzhľadom na rozptyl, znamená to platnosť nulovej hypotézy $H_0: \sigma = 0,01$. Ak to tak nie je, znamená to platnosť alternatívnej hypotézy $H_1: \sigma > 0,01$. Teraz zostrojíme kritickú oblasť W pre test tejto dvojice hypotéz. Platí, že

$$W = \left\{ (x_1, \dots, x_{100}) \in R^{100} : \frac{99 \cdot s_{100}^2}{0,01^2} = \frac{99 \cdot 0,000144}{0,0001} = 142,56 > 123,2 = \chi_{99}^{2,kv}(0,975) \right\}.$$

V našom prípade $s_{100} = 0,012$ litra, $\chi_{99}^{2,kv}(0,95) = 123,2$. Z toho vyplýva, že platí

$$\frac{99 \cdot s_{100}^2}{0,01^2} = \frac{99 \cdot 0,000144}{0,0001} = 142,56 > 123,2 = \chi_{99}^{2,kv}(0,975).$$

Preto $(x_1, \dots, x_{100}) = (0,96; 0,97; 0,97; 0,97; \dots; 1,02; 1,05) \in W$. Z toho vyplýva, že zamietame hypotézu $H_0: \sigma = 0,01$ a prijímame hypotézu $H_1: \sigma > 0,01$, čo slovne formulujeme nasledovne: *Bolo štatisticky preukázané, že plniaca linka plní olej do nádob s väčším rozptylom, ako je povolená hodnota 0,01 litra.* To znamená, že firma dostane pokutu aj za prekročenie povolenej hodnoty rozptylu.

Príklad 45. V skúšobni sa porovnávali pneumatiky firmy X a firmy Y . Skúmala sa výška dezénu v milimetroch po najazdení 50 000 kilometrov. Každá firma dodala po 10 pneumatík. Výška dezénu je v nasledujúcej tabuľke

X	2,4	3,8	1,5	2,8	2,4	3,5	2,4	2,8	2,7	2,2
Y	3,1	2,9	2,9	3,4	2,8	2,5	3,4	3,0	2,9	3,2

V prvom riadku sú výšky dezénov x_i typu pneumatík od firmy X a v druhom riadku sú výšky dezénov y_i typu pneumatík od firmy Y . Na hladine $\alpha = 0,05$ otestujte, či sú obidva druhy pneumatík rovnako kvalitné.

Riešenie. Obidva druhy pneumatík sú rovnako kvalitné, ak priemerná výška dezénu po najazdení 50 000 kilometrov je u oboch typov rovnaká. Alebo, čo je to isté, priemerná hodnota rozdielov $z_i = x_i - y_i$ je rovná nule. Ak tieto rozdiely zapíšeme do tabuľky, dostaneme s ohľadom na ich znamienka

Z	-0,7	0,9	-1,4	-0,6	-0,4	1,0	-1,0	-0,2	-0,2	-1,0
-----	------	-----	------	------	------	-----	------	------	------	------

To, že obidva druhy pneumatík sú rovnako kvalitné, matematicky znamená, že máme otestovať dvojicu hypotéz $H_0 : \mu = 0 \propto H_1 : \mu \neq 0$. Ak sú obidva typy pneumatík rovnako kvalitné, znamená to platnosť nulovej hypotézy $H_0 : \mu = 0$. Ak to tak nie je, znamená to platnosť alternatívnej hypotézy $H_1 : \mu \neq 0$. Teraz zostrojíme kritickú oblasť W pre test tejto dvojice hypotéz. Platí, že

$$W = \left\{ (x_1, \dots, x_{10}) \in R^{10} : \frac{|\bar{z}_{10} - 0|}{s_{10}} \sqrt{10} > t_9^{kv}(0,975) \right\}.$$

V našom prípade $\bar{z}_{10} = -0,36$ mm, $s_{20} = 0,786$ mm, $t_9^{kv}(0,975) = 2,26$. Z toho vyplýva, že platí

$$\frac{|\bar{z}_{10} - 0|}{s_{10}} \sqrt{10} = \frac{|-0,36 - 0|}{0,786} \sqrt{10} = 1,45 < 2,26 = t_9^{kv}(0,975).$$

Preto $(z_1, \dots, z_{10}) = (-0,7; 0,9; -1,4; -0,6; \dots; -0,2; -1,0) \notin W$. Z toho vyplýva, že zamietame alternatívnu hypotézu $H_1 : \mu \neq 0$ a prijímame nulovú hypotézu $H_0 : \mu = 0$, čo slovne formulujeme nasledovne: *Bolo štatisticky preukázané, že obidva typy pneumatík sú rovnako kvalitné; vzniknuté rozdiely možno pripísať náhodným vplyvom.* Záverom vypočítajme ešte priemerné výšky dezénov $\bar{x}_{10}$, $\bar{y}_{10}$ oboch druhov pneumatík. Po príslušnom výpočte vychádzajú hodnoty $\bar{x}_{10} = 2,65$ mm, $\bar{y}_{10} = 3,01$ mm. Z toho, že $\bar{x}_{10} < \bar{y}_{10}$ by sa dalo usudzovať, že typ X je menej kvalitný ako typ Y . Test ukázal, že to tak nie je.

11 PLÁNOVANIE EXPERIMENTU

Príklad 46. Uvažujme funkciu dvoch premenných $f(x, y) = x^3 + y^2 + xy$. Stanovte výraz $\sum_{k=0}^3 \frac{d^k f(a, b, x, y)}{k!}$, kde $d^k f(a, b, x, y)$ je k -ty diferenciál funkcie dvoch premenných, pričom a) $(a, b) = (0, 0)$, b) $(a, b) = (2, 3)$, c) $(a, b) = (-1, 1)$.

Riešenie. Vzorce pre k -ty diferenciál sú nasledovné:

$$\begin{aligned} d^0 f(a, b, x, y) &= f(a, b), \\ d^1 f(a, b, x, y) &= df(a, b, x, y) = f'_x(a, b)(x - a) + f'_y(a, b)(y - b), \\ d^2 f(a, b, x, y) &= f''_{xx}(a, b) \cdot (x - a)^2 + 2f''_{xy}(a, b)(x - a)(y - b) + f''_{yy}(a, b) \cdot (y - b)^2, \\ d^3 f(a, b, x, y) &= f'''_{xxx}(a, b) \cdot (x - a)^3 + 3f'''_{xxy}(a, b)(x - a)^2(y - b) + \\ &\quad + 3f'''_{xyy}(a, b)(x - a)(y - b)^2 + f'''_{yyy}(a, b)(y - b)^3, \end{aligned}$$

kde napr. $f'''_{xxy}(a, b)$ značí parciálnu deriváciu tretieho rádu podľa premenných x, y v počítaní v bode (a, b) ; podobný význam majú ostatné symboly. Po príslušnom výpočte dostávame

$$\begin{aligned} f'_x(a, b) &= 3a^2 + b, f'_y(a, b) = 2b + a, f''_{xx}(a, b) = 6a, f''_{xy}(a, b) = 1, f''_{yy}(a, b) = 0, \\ f'''_{xxx}(a, b) &= 6, f'''_{xxy}(a, b) = f'''_{xyy}(a, b) = f'''_{yyy}(a, b) = 0. \end{aligned}$$

a) Ak v predchádzajúcich vzorcoch dosadíme $(a, b) = (0, 0)$, dostaneme

$$\begin{aligned} d^0 f(0, 0, x, y) &= 0, \\ d^1 f(0, 0, x, y) &= 0, \\ d^2 f(0, 0, x, y) &= 2xy + 2y^2, \\ d^3 f(0, 0, x, y) &= 6x^3. \end{aligned}$$

Odtiaľ $\sum_{k=0}^3 \frac{d^k f(0, 0, x, y)}{k!} = x^3 + y^2 + xy.$

b) Ak v predchádzajúcich vzorcoch dosadíme $(a, b) = (2, 3)$, dostaneme

$$\begin{aligned} d^0 f(2, 3, x, y) &= 23, \\ d^1 f(2, 3, x, y) &= 15(x - 2) + 8(y - 3), \\ d^2 f(2, 3, x, y) &= 12(x - 2)^2 + 2(x - 2)(y - 3) + 2(y - 3)^2, \\ d^3 f(2, 3, x, y) &= 6(x - 2)^3. \end{aligned}$$

Odtiaľ

$$\sum_{k=0}^3 \frac{d^k f(2, 3, x, y)}{k!} = 23 + 15(x - 2) + 8(y - 3) + 6(x - 2)^2 + (x - 2)(y - 3) + (x - 2)^2 + (x - 2)^3.$$

c) Ak v predchádzajúcich vzorcoch dosadíme $(a, b) = (-1, 1)$, dostaneme

$$\begin{aligned}d^0 f(-1, 1, x, y) &= -1, \\d^1 f(-1, 1, x, y) &= 4(x+1) + 1(y-1), \\d^2 f(-1, 1, x, y) &= -6(x+1)^2 + 2(x+1)(y-1) + 2(y-1)^2, \\d^3 f(-1, 1, x, y) &= 6(x+1)^3.\end{aligned}$$

Odtiaľ

$$\sum_{k=0}^3 \frac{d^k f(-1, 1, x, y)}{k!} = -1 + 4(x+1) + (y-1) - 3(x+1)^2 + (x+1)(y-1) + (y-1)^2 + (x+1)^3.$$

Príklad 47. Uvažujme funkciu dvoch premenných $f(x, y) = e^x \sin y$. Nájdite nasledujúce vý-

razy: a) $\sum_{k=0}^3 \frac{d^k f(0, 0, x, y)}{k!}$, b) $\sum_{k=0}^3 \frac{d^k f\left(0, 0, \frac{1}{10}, \frac{1}{10}\right)}{k!}$. c) Na kalkulačke vypočítajte hodnotu

$f\left(\frac{1}{10}, \frac{1}{10}\right)$ a porovnajte ju s hodnotou výrazu $\sum_{k=0}^3 \frac{d^k f\left(0, 0, \frac{1}{10}, \frac{1}{10}\right)}{k!}$.

Riešenie. a) Po príslušnom výpočte dostávame

$$\begin{aligned}f'_x(0, 0) &= 0, f'_y(0, 0) = 1, f''_{xx}(0, 0) = 0, f''_{xy}(0, 0) = 1, f''_{yy}(0, 0) = 0, \\f'''_{xxx}(a, b) &= 0, f'''_{xxy}(a, b) = 1, f'''_{xyy}(a, b) = 0, f'''_{yyy}(a, b) = -1.\end{aligned}$$

Odtiaľ
$$\sum_{k=0}^3 \frac{d^k f(0, 0, x, y)}{k!} = y + xy + \frac{x^2 y}{2} - \frac{y^3}{6}.$$

b) Ak v predchádzajúcom vzorci dosadíme $(x, y) = \left(\frac{1}{10}, \frac{1}{10}\right)$, dostaneme

$$\sum_{k=0}^3 \frac{d^k f\left(0, 0, \frac{1}{10}, \frac{1}{10}\right)}{k!} = \frac{1}{10} + \frac{1}{100} + \frac{1}{2000} - \frac{1}{6000} \approx 0,110333333.$$

c) Pomocou kalkulačky dostávame $f\left(\frac{1}{10}, \frac{1}{10}\right) = e^{\frac{1}{10}} \sin \frac{1}{10} \approx 0,110332988$. Porovnaním posledných dvoch desatinných čísel vidíme, že sa zhodujú až na 5 desatinných miest. To znamená, že Taylorov polynóm funkcie dvoch premenných možno použiť na približný výpočet funkčných hodnôt funkcie dvoch premenných.

Príklad 48. Nech funkcia $f(x_1, x_2)$ závisí od dvoch faktorov x_1, x_2 , pričom o týchto faktoroch vieme, že $x_1 \in \langle 0, 20 \rangle$, $x_2 \in \langle 0, 30 \rangle$. Na základe týchto skutočností naplánujte experiment typu a) 2^2 , b) 3^2 , c) 4^2 .

Riešenie. Naplánovanie experimentu spočíva v určení množiny M bodov (x_1, x_2) , v ktorých potom vykonávame meranie funkčných hodnôt funkcie $f(x_1, x_2)$.

a) Experiment typu 2^2 znamená, že meranie uskutočňujeme v krajných bodoch dvojrozmerného intervalu

$$\langle 0, 20 \rangle \times \langle 0, 30 \rangle = \{(x_1, x_2) \in R^2 : x_1 \in \langle 0, 20 \rangle, x_2 \in \langle 0, 30 \rangle\}$$

To znamená, že množina M má tvar

$$M = \{(x_1, x_2) \in R^2 : x_1 \in M_1, x_2 \in M_2\}, \text{ kde } M_1 = \{0, 20\}, M_2 = \{0, 30\}.$$

Po dosadení vychádza $M = \{(0, 0), (20, 0), (0, 30), (20, 30)\}$.

b) Experiment typu 3^2 znamená, že meranie uskutočňujeme v krajných a stredových bodoch dvojrozmerného intervalu

$$\langle 0, 20 \rangle \times \langle 0, 30 \rangle = \{(x_1, x_2) \in R^2 : x_1 \in \langle 0, 20 \rangle, x_2 \in \langle 0, 30 \rangle\}$$

To znamená, že množina M má tvar

$$M = \{(x_1, x_2) \in R^2 : x_1 \in M_1, x_2 \in M_2\}, \text{ kde } M_1 = \{0, 10, 20\}, M_2 = \{0, 15, 30\}.$$

Vychádza $M = \{(0, 0), (10, 0), (20, 0), (0, 15), (10, 15), (20, 15), (0, 30), (10, 30), (20, 30)\}$.

c) Pre experiment typu 4^2 to znamená, že meranie uskutočňujeme v $4^2 = 16$ bodoch dvojrozmerného intervalu

$$\langle 0, 20 \rangle \times \langle 0, 30 \rangle = \{(x_1, x_2) \in R^2 : x_1 \in \langle 0, 20 \rangle, x_2 \in \langle 0, 30 \rangle\}$$

To znamená, že množina M má tvar

$$M = \{(x_1, x_2) \in R^2 : x_1 \in M_1, x_2 \in M_2\}, \text{ kde } M_1 = \left\{0, \frac{20}{3}, \frac{40}{3}, 20\right\}, M_2 = \{0, 10, 20, 30\}.$$

Po dosadení vychádza

$$M = \left\{ (0,0), \left(\frac{20}{3}, 0\right), \left(\frac{40}{3}, 0\right), (20,0), (0,10), \left(\frac{20}{3}, 10\right), \left(\frac{40}{3}, 10\right), (20,10), \right. \\ \left. (0,20), \left(\frac{20}{3}, 20\right), \left(\frac{40}{3}, 20\right), (20,20), (0,30), \left(\frac{20}{3}, 30\right), \left(\frac{40}{3}, 30\right), (20,30) \right\}.$$

Príklad 49. Nech funkcia $f(x_1, x_2, x_3)$ závisí od troch faktorov x_1, x_2, x_3 pričom o týchto faktoroch vieme, že $x_1 \in \langle -10, 20 \rangle$, $x_2 \in \langle -20, 30 \rangle$, $x_3 \in \langle 10, 40 \rangle$. Na základe tohto naplánujte experiment typu 2^3 .

Riešenie. Naplánovanie experimentu spočíva v určení množiny M bodov (x_1, x_2, x_3) , v ktorých potom vykonávame meranie funkčných hodnôt funkcie $f(x_1, x_2, x_3)$. Experiment typu 2^3 znamená, že meranie uskutočňujeme v krajných bodoch trojrozmerného intervalu

$$\langle -10, 20 \rangle \times \langle -20, 30 \rangle \times \langle 10, 40 \rangle = \{(x_1, x_2, x_3) \in R^3 : x_1 \in \langle -10, 20 \rangle, x_2 \in \langle -20, 30 \rangle, x_3 \in \langle 10, 40 \rangle\}.$$

To znamená, že množina M má tvar $M = \{(x_1, x_2, x_3) \in R^3 : x_1 \in M_1, x_2 \in M_2, x_3 \in M_3\}$, kde

$$M_1 = \{-10, 20\}, M_2 = \{-20, 30\}, M_3 = \{10, 40\}.$$

Po dosadení vychádza

$$M = \{(-10, -20, 10), (-10, -20, 40), (-10, 30, 10), (-10, 30, 40), \\ (20, -20, 10), (20, -20, 40), (20, 30, 10), (20, 30, 40)\}.$$

Príklad 50. Nech $x \in \langle 10, 50 \rangle$. Zavedme novú premennú t pomocou lineárnej transformácie $t = \frac{x-s}{\lambda}$, kde $s = \frac{10+50}{2} = 30$ je stred intervalu $\langle 10, 50 \rangle$ a $\lambda = \frac{50-10}{2} = 20$ je jeho polovičná dĺžka. Presvedčte sa, že transformácia t zobrazí interval $\langle 10, 50 \rangle$ na interval $\langle -1, 1 \rangle$.

Riešenie. Lineárna transformácia má tvar

$$t = \frac{x-s}{\lambda} = \frac{x-30}{20}.$$

Pretože na intervale $\langle 10, 50 \rangle$ platí $\frac{dt}{dx} = \frac{1}{20} > 0$, z toho vyplýva, že funkcia $t = \frac{x-30}{20}$ je na tomto intervale rastúca. Preto sa interval $\langle 10, 50 \rangle$ zobrazí na interval $\langle t(10), t(50) \rangle$. Ale

$$t(10) = \frac{10-30}{20} = -1, \quad t(50) = \frac{50-30}{20} = 1,$$

odkiaľ tvrdenie príkladu bezprostredne vyplýva. Poznamenávame, že toto tvrdenie tiež vyplýva z príslušnej vety z prednášky. Ktoorej ?

12 REGRESNÁ FUNKCIA A TRANSFORMOVANÁ REGRESNÁ FUNKCIA

Príklad 51. Uvažujme experiment typu 2^2 , pričom $x_1 \in \langle 10; 80 \rangle$, $x_2 \in \langle 30; 60 \rangle$. Regresná funkcia je lineárneho typu, pričom v každom zo 4 bodov sme vykonali po 2 merania, čím sme získali nasledujúce hodnoty:

$$\begin{aligned} y_1 = f(10, 30) = 5, & \quad y_3 = f(10, 60) = 9, & \quad y_5 = f(80, 30) = 2, & \quad y_7 = f(80, 60) = 5, \\ y_2 = f(10, 30) = 7, & \quad y_4 = f(10, 60) = 8, & \quad y_6 = f(80, 30) = 3, & \quad y_8 = f(80, 60) = 4. \end{aligned}$$

Odhadnite $\hat{b}$ -koeficienty transformovanej regresnej funkcie lineárneho typu $\hat{f}_T(t_1, t_2)$ metódou najmenších štvorcov. Tiež stanovte $\hat{\beta}$ -koeficienty pôvodnej regresnej funkcie lineárneho typu $\hat{f}(x_1, x_2)$.

Riešenie. Transformovaná regresná funkcia lineárneho typu má tvar:

$$\hat{f}_T(t_1, t_2) = \hat{b}_0 + \hat{b}_1 t_1 + \hat{b}_2 t_2 + \hat{b}_{12} t_1 t_2.$$

Transformované namerané hodnoty sú:

$$\begin{aligned} y_1 = f_T(-1, -1) = 5, & \quad y_3 = f_T(-1, 1) = 9, & \quad y_5 = f_T(1, -1) = 2, & \quad y_7 = f_T(1, 1) = 5, \\ y_2 = f_T(-1, -1) = 7, & \quad y_4 = f_T(-1, 1) = 8, & \quad y_6 = f_T(1, -1) = 3, & \quad y_8 = f_T(1, 1) = 4. \end{aligned}$$

Na výpočet $\hat{b}$ -koeficientov budeme potrebovať transformovanú maticu plánu $\mathbf{X}_T$ a vektor meraní $\mathbf{Y}$. Vektor parametrov $\hat{\mathbf{b}}$ potom určíme zo vzťahu

$$\hat{\mathbf{b}} = \frac{1}{m k^n} \mathbf{X}_T' \mathbf{Y},$$

kde $\mathbf{X}_T'$ znamená transponovanú maticu plánu $\mathbf{X}_T$. Pretože $k = m = n = 2$, dostávame

$$\begin{aligned}\hat{\mathbf{b}} &= \begin{bmatrix} \hat{b}_0 \\ \hat{b}_1 \\ \hat{b}_2 \\ \hat{b}_{12} \end{bmatrix} = \frac{1}{mk^n} \mathbf{X}'_T \mathbf{Y} = \frac{1}{2 \cdot 2^2} \cdot \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ 7 \\ 9 \\ 8 \\ 2 \\ 3 \\ 5 \\ 4 \end{bmatrix} = \\ &= \frac{1}{8} \cdot \begin{bmatrix} 5+7+9+8+2+3+5+4 \\ -5-7-9-8+2+3+5+4 \\ -5-7+9+8-2-3+5+4 \\ 5+7-9-8-2-3+5+4 \end{bmatrix} = \frac{1}{8} \cdot \begin{bmatrix} 43 \\ -15 \\ 9 \\ -1 \end{bmatrix} = \begin{bmatrix} 43/8 \\ -15/8 \\ 9/8 \\ -1/8 \end{bmatrix} = \begin{bmatrix} 5,375 \\ -1,875 \\ 1,125 \\ -0,125 \end{bmatrix}.\end{aligned}$$

Stĺpce matice $\mathbf{X}'_T$ majú v prvom riadku číslo 1, v druhom riadku hodnotu premennej t_1 , v treťom riadku hodnotu premennej t_2 a v štvrtom riadku hodnotu súčinu $t_1 t_2$. Pretože sme v každom zo 4 bodov vykonali po 2 merania, sú prvé dva stĺpce rovnaké, druhé dva stĺpce rovnaké, tretie dva stĺpce rovnaké a tiež štvrté dva stĺpce rovnaké. Do prvých dvoch stĺpcov sme dosadili bod $(-1, -1)$, čo znamená $t_1 = -1, t_2 = -1$, do druhých dvoch stĺpcov bod $(-1, 1)$, čo znamená $t_1 = -1, t_2 = 1$, do tretích dvoch stĺpcov bod $(1, -1)$, čo znamená $t_1 = 1, t_2 = -1$ a do štvrtých dvoch stĺpcov bod $(1, 1)$, čo znamená $t_1 = 1, t_2 = 1$. Z vypočítanej hodnoty vektora $\hat{\mathbf{b}}$ dostávame hľadanú transformovanú regresnú funkciu lineárneho typu

$$\hat{f}_T(t_1, t_2) = \hat{b}_0 + \hat{b}_1 t_1 + \hat{b}_2 t_2 + \hat{b}_{12} t_1 t_2 = 5,375 - 1,875 t_1 + 1,125 t_2 - 0,125 t_1 t_2.$$

Na jej základe potom stanovíme pôvodnú regresnú funkciu lineárneho typu $\hat{f}(x_1, x_2)$, kde

$$\hat{f}(x_1, x_2) = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \hat{\beta}_{12} x_1 x_2$$

takto:

Pretože premenné t_1, t_2 sme získali z premenných x_1, x_2 pomocou lineárnych transformácií

$$t_1 = \frac{x_1 - s_1}{\lambda_1}, \quad t_2 = \frac{x_2 - s_2}{\lambda_2},$$

kde $s_1 = 45$ je stred intervalu $\langle 10; 80 \rangle$, číslo $\lambda_1 = 35$ je polovičná dĺžka intervalu $\langle 10; 80 \rangle$; podobne $s_2 = 45$ je stred intervalu $\langle 30; 60 \rangle$, číslo $\lambda_2 = 15$ je polovičná dĺžka intervalu $\langle 30; 60 \rangle$. Preto

$$\begin{aligned}\widehat{f}(x_1, x_2) &= \widehat{f}_T\left(\frac{x_1 - s_1}{\lambda_1}, \frac{x_2 - s_2}{\lambda_2}\right) = \widehat{f}_T\left(\frac{x_1 - 45}{35}, \frac{x_2 - 45}{15}\right) = \\ &= 5,375 - 1,875 \cdot \frac{x_1 - 45}{35} + 1,125 \cdot \frac{x_2 - 45}{15} - 0,125 \cdot \frac{x_1 - 45}{35} \cdot \frac{x_2 - 45}{15}.\end{aligned}$$

Odtiaľ po roznásobení a úprave dostávame

$$\begin{aligned}\widehat{f}(x_1, x_2) &= \left(5,375 + 1,875 \cdot \frac{45}{35} - 1,125 \cdot \frac{45}{15} - 0,125 \cdot \frac{45}{35} \cdot \frac{45}{15}\right) + \left(1,875 \cdot \frac{1}{35} + 0,125 \cdot \frac{1}{35} \cdot \frac{45}{15}\right)x_1 + \\ &+ \left(-1,125 \cdot \frac{45}{15} + 0,125 \cdot \frac{1}{15} \cdot \frac{45}{35}\right)x_2 + \left(-0,125 \cdot \frac{1}{35} \cdot \frac{1}{15}\right)x_1x_2 \approx \\ &\approx 3,9285714 + 0,0642857x_1 - 3,3642857x_2 - 0,0002381x_1x_2.\end{aligned}$$

Príklad 52. Uvažujme experiment typu 2^3 , pričom $x_1 \in \langle 10; 80 \rangle$, $x_2 \in \langle 30; 60 \rangle$, $x_3 \in \langle 30; 60 \rangle$. Regresná funkcia je lineárneho typu, pričom v každom z 8 bodov sme vykonali po 1 meraní, čím sme získali nasledujúce transformované hodnoty:

$$\begin{aligned}y_1 &= f_T(-1, -1, -1) = 2, & y_5 &= f_T(1, -1, -1) = 4, \\ y_2 &= f_T(-1, -1, 1) = 4, & y_6 &= f_T(1, -1, 1) = 5, \\ y_3 &= f_T(-1, 1, -1) = 7, & y_7 &= f_T(1, 1, -1) = 3, \\ y_4 &= f_T(-1, 1, 1) = 4, & y_8 &= f_T(1, 1, 1) = 2.\end{aligned}$$

Odhadnite $\widehat{b}$ -koeficienty transformovanej regresnej funkcie lineárneho typu $\widehat{f}_T(t_1, t_2, t_3)$ metódou najmenších štvorcov.

Riešenie. Transformovaná regresná funkcia lineárneho typu má tvar:

$$\widehat{f}_T(t_1, t_2, t_3) = \widehat{b}_0 + \widehat{b}_1t_1 + \widehat{b}_2t_2 + \widehat{b}_3t_3 + \widehat{b}_{12}t_1t_2 + \widehat{b}_{13}t_1t_3 + \widehat{b}_{23}t_2t_3 + \widehat{b}_{123}t_1t_2t_3.$$

Na výpočet $\widehat{b}$ -koeficientov budeme potrebovať transformovanú maticu plánu $\mathbf{X}_T$ a vektor meraní $\mathbf{Y}$. Vektor parametrov $\widehat{\mathbf{b}}$ potom určíme zo vzťahu

$$\widehat{\mathbf{b}} = \frac{1}{mk^n} \mathbf{X}_T' \mathbf{Y},$$

kde $\mathbf{X}_T'$ znamená transponovanú maticu plánu $\mathbf{X}_T$. Pretože $m = 1$, $k = 2$, $n = 3$, dostávame

$$\hat{\mathbf{b}} = \begin{bmatrix} \hat{b}_0 \\ \hat{b}_1 \\ \hat{b}_2 \\ \hat{b}_3 \\ \hat{b}_{12} \\ \hat{b}_{13} \\ \hat{b}_{23} \\ \hat{b}_{123} \end{bmatrix} = \frac{1}{mk^n} \mathbf{X}'_T \mathbf{Y} = \frac{1}{1 \cdot 2^3} \cdot \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 4 \\ 7 \\ 4 \\ 4 \\ 5 \\ 3 \\ 2 \end{bmatrix} =$$

$$= \frac{1}{8} \cdot \begin{bmatrix} 31 \\ -3 \\ 1 \\ -1 \\ -9 \\ 1 \\ 3 \\ -5 \end{bmatrix} = \begin{bmatrix} 31/8 \\ -3/8 \\ 1/8 \\ -1/8 \\ -9/8 \\ 1/8 \\ 3/8 \\ -5/8 \end{bmatrix} = \begin{bmatrix} 3,875 \\ -0,375 \\ 0,125 \\ -0,125 \\ -1,125 \\ 0,125 \\ 0,375 \\ -0,625 \end{bmatrix}.$$

Stĺpce matice $\mathbf{X}'_T$ majú v prvom riadku číslo 1, v druhom riadku hodnotu premennej t_1 , v treťom riadku hodnotu premennej t_2 , v štvrtom riadku hodnotu premennej t_3 , v piatom riadku hodnotu súčinu $t_1 t_2$, v šiestom súčinu $t_1 t_3$, v siedmom súčinu $t_2 t_3$ a v ôsmom hodnotu súčinu $t_1 t_2 t_3$. Pretože sme v každom z 8 bodov vykonali po 1 meraní, je každý stĺpec iný. Do prvého stĺpca sme dosadili bod $(-1, -1, -1)$, čo znamená $t_1 = -1, t_2 = -1, t_3 = -1$, do druhého bod $(-1, -1, 1)$, čo znamená $t_1 = -1, t_2 = -1, t_3 = 1$, atď. až do ôsmeho sme dosadili posledný bod $(1, 1, 1)$, čo znamená $t_1 = 1, t_2 = 1, t_3 = 1$. Z vypočítanej hodnoty vektora $\hat{\mathbf{b}}$ dostávame nakoniec hľadanú transformovanú regresnú funkciu lineárneho typu

$$\begin{aligned} \hat{f}_T(t_1, t_2, t_3) &= \hat{b}_0 + \hat{b}_1 t_1 + \hat{b}_2 t_2 + \hat{b}_3 t_3 + \hat{b}_{12} t_1 t_2 + \hat{b}_{13} t_1 t_3 + \hat{b}_{23} t_2 t_3 + \hat{b}_{123} t_1 t_2 t_3 = \\ &= 3,875 - 0,375 t_1 + 0,125 t_2 - 0,125 t_3 - 1,125 t_1 t_2 + 0,125 t_1 t_3 + 0,375 t_2 t_3 - 0,625 t_1 t_2 t_3. \end{aligned}$$

13 VYHODNOCOVANIE EXPERIMENTU

Príklad 53. Uvažujme experiment typu 2^2 , kde regresná funkcia je lineárneho typu, pričom v každom zo 4 bodov sme vykonali po 2 merania. Získali sme tak nasledujúce údaje, ktoré po vykonaní príslušnej lineárnej transformácie T , kde

$$T: t_1 = \frac{x_1 - s_1}{\lambda_1}, \quad t_2 = \frac{x_2 - s_2}{\lambda_2}$$

majú tvar:

$$\begin{aligned} y_1 = f_T(-1, -1) = 3, & \quad y_3 = f_T(-1, 1) = 7, & \quad y_5 = f_T(1, -1) = -2, & \quad y_7 = f_T(1, 1) = 2, \\ y_2 = f_T(-1, -1) = 5, & \quad y_4 = f_T(-1, 1) = 5, & \quad y_6 = f_T(1, -1) = -4, & \quad y_8 = f_T(1, 1) = 4. \end{aligned}$$

Vyhodnoťme tento experiment na hladine $\alpha = 0,1$.

Riešenie. Transformovaná regresná funkcia lineárneho typu má tvar:

$$\hat{f}_T(t_1, t_2) = \hat{b}_0 + \hat{b}_1 t_1 + \hat{b}_2 t_2 + \hat{b}_{12} t_1 t_2.$$

Na výpočet $\hat{b}$ -koeficientov budeme potrebovať transformovanú maticu plánu $\mathbf{X}_T$ a vektor meraní $\mathbf{Y}$. Vektor parametrov $\hat{\mathbf{b}}$ potom určíme zo vzťahu

$$\hat{\mathbf{b}} = \frac{1}{m k^n} \mathbf{X}_T' \mathbf{Y},$$

kde $\mathbf{X}_T'$ znamená transponovanú maticu plánu $\mathbf{X}_T$. Pretože $k = m = n = 2$, dostávame

$$\begin{aligned} \hat{\mathbf{b}} &= \begin{bmatrix} \hat{b}_0 \\ \hat{b}_1 \\ \hat{b}_2 \\ \hat{b}_{12} \end{bmatrix} = \frac{1}{m k^n} \mathbf{X}_T' \mathbf{Y} = \frac{1}{2 \cdot 2^2} \cdot \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 5 \\ 7 \\ 5 \\ -2 \\ -4 \\ 2 \\ 4 \end{bmatrix} = \\ &= \frac{1}{8} \cdot \begin{bmatrix} 3+5+7+5-2-4+2+4 \\ -3-5-7-5-2-4+2+4 \\ -3-5+7+5+2+4+2+4 \\ 3+5-7-5+2+4+2+4 \end{bmatrix} = \frac{1}{8} \cdot \begin{bmatrix} 20 \\ -20 \\ 16 \\ 8 \end{bmatrix} = \begin{bmatrix} 20/8 \\ -20/8 \\ 16/8 \\ 8/8 \end{bmatrix} = \begin{bmatrix} 2,5 \\ -2,5 \\ 2 \\ 1 \end{bmatrix}. \end{aligned}$$

Stĺpce matice $\mathbf{X}_T'$ majú v prvom riadku číslo 1, v druhom riadku hodnotu premennej t_1 , v treťom riadku hodnotu premennej t_2 a v štvrtom riadku hodnotu súčinu $t_1 t_2$. Pretože sme v

každom zo 4 bodov vykonali po 2 merania, sú prvé dva stĺpce rovnaké, druhé dva stĺpce rovnaké, tretie dva stĺpce rovnaké, a tiež štvrté dva stĺpce rovnaké. Do prvých dvoch stĺpcov sme dosadili bod $(-1, -1)$, čo znamená $t_1 = -1, t_2 = -1$, do druhých dvoch stĺpcov bod $(-1, 1)$, čo znamená $t_1 = -1, t_2 = 1$, do tretích dvoch stĺpcov bod $(1, -1)$, čo znamená $t_1 = 1, t_2 = -1$ a do štvrtých dvoch stĺpcov bod $(1, 1)$, čo znamená $t_1 = 1, t_2 = 1$. Z vypočítanej hodnoty vektora $\hat{\mathbf{b}}$ dostávame hľadanú transformovanú regresnú funkciu lineárneho typu

$$\hat{f}_T(t_1, t_2) = \hat{b}_0 + \hat{b}_1 t_1 + \hat{b}_2 t_2 + \hat{b}_{12} t_1 t_2 = 2,5 - 2,5 t_1 + 2 t_2 + 1 t_1 t_2$$

alebo priamo koeficienty $\hat{b}_0 = 2,5, \hat{b}_1 = -2,5, \hat{b}_2 = 2, \hat{b}_{12} = 1$.

Vyhodnotenie experimentu vykonáme otestovaním vplyvu jednotlivých faktorov na hodnotu transformovanej regresnej funkcie $\hat{f}_T(t_1, t_2)$. Aby sme tak urobili, vypočítame najprv tzv. vyrovnané hodnoty $\hat{y}_i$ nameraných hodnôt y_i , pričom

$$\begin{aligned}\hat{y}_1 = \hat{y}_2 &= f_T(-1, -1) = 2,5 - 2,5(-1) + 2(-1) + 1(-1)(-1) = 4, \\ \hat{y}_3 = \hat{y}_4 &= f_T(-1, 1) = 2,5 - 2,5(-1) + 2(1) + 1(-1)(1) = 6, \\ \hat{y}_5 = \hat{y}_6 &= f_T(1, -1) = 2,5 - 2,5(1) + 2(-1) + 1(1)(-1) = -3, \\ \hat{y}_7 = \hat{y}_8 &= f_T(1, 1) = 2,5 - 2,5(1) + 2(1) + 1(1)(1) = 3.\end{aligned}$$

Na ich základe vypočítame veličinu R_0^2 , kde

$$\begin{aligned}R_0^2 = \sum_{i=1}^{mk^n} [(y_i - \hat{y}_i)^2] &= (3 - 4)^2 + (5 - 4)^2 + (7 - 6)^2 + (5 - 6)^2 + \\ &+ (-2 + 3)^2 + (-4 + 3)^2 + (2 - 3)^2 + (4 - 3)^2 = 8.\end{aligned}$$

Vplyv spoločného pôsobenia faktorov t_1, t_2 na hodnotu transformovanej regresnej funkcie $\hat{f}_T(t_1, t_2)$ zistíme testovaním dvojice hypotéz $H_0 : b_{12} = 0 \propto H_1 : b_{12} \neq 0$. Kritickú oblasť W tohto testu skonštruujeme pomocou veličín $|t|, t_{(m-1)2^2}^{kv}(1 - \frac{\alpha}{2})$, kde

$$|t| = \frac{|\hat{b}_{12}| k^n \sqrt{m(m-1)}}{\sqrt{R_0^2}} = \frac{1 \cdot 2^2 \cdot \sqrt{2 \cdot 1}}{\sqrt{8}} = 2$$

a $t_{(m-1)2^2}^{kv}(1 - \frac{\alpha}{2}) = t_{4(2-1)}^{kv}(1 - \frac{0,1}{2}) = t_4^{kv}(0,95) = 2,77$ je kvantil t -rozdelenia. Pretože $|t| < t_4^{kv}(0,95)$, namerané hodnoty nepadnú do kritickej oblasti W , a preto prijímame hypotézu $H_0 : b_{12} = 0$, čo slovne formulujeme tak, že *bolo štatisticky preukázané, že spoločné pôsobenie faktorov t_1, t_2 na hodnotu regresnej funkcie má zanedbateľný vplyv.*

Vplyv samostatného pôsobenia faktora t_1 na hodnotu transformovanej regresnej funkcie $\hat{f}_T(t_1, t_2)$ zistíme testovaním dvojice hypotéz $H_0 : b_1 = 0 \propto H_1 : b_1 \neq 0$. Kritickú oblasť W tohto testu skonštruujeme pomocou veličín $|t|, t_{(m-1)2^2}^{kv}(1 - \frac{\alpha}{2})$, kde

$$|t| = \frac{|\hat{b}_1| k^n \sqrt{m(m-1)}}{\sqrt{R_0^2}} = \frac{|-2,5| \cdot 2^2 \cdot \sqrt{2 \cdot 1}}{\sqrt{8}} = 5$$

a $t_{4(m-1)}^{kv}(1 - \frac{\alpha}{2}) = t_{4(2-1)}^{kv}(1 - \frac{0,1}{2}) = t_4^{kv}(0,95) = 2,77$ je kvantil t -rozdelenia. Pretože $|t| > t_4^{kv}(0,95)$, namerané hodnoty padnú do kritickej oblasti W , a preto prijímame hypotézu $H_1 : b_1 \neq 0$, čo slovne formulujeme tak, že *bolo štatisticky preukázané, že samostatné pôsobenie faktora t_1 má na hodnotu regresnej funkcie podstatný vplyv.*

Vplyv samostatného pôsobenia faktora t_2 na hodnotu transformovanej regresnej funkcie $\hat{f}_T(t_1, t_2)$ zistíme testovaním dvojice hypotéz $H_0 : b_2 = 0 \propto H_1 : b_2 \neq 0$. Kritickú oblasť W tohto testu skonštruujeme pomocou veličín $|t|, t_{(m-1)2^2}^{kv}(1 - \frac{\alpha}{2})$, kde

$$|t| = \frac{|\hat{b}_2| k^n \sqrt{m(m-1)}}{\sqrt{R_0^2}} = \frac{|2| \cdot 2^2 \cdot \sqrt{2 \cdot 1}}{\sqrt{8}} = 4$$

a $t_{4(m-1)}^{kv}(1 - \frac{\alpha}{2}) = t_{4(2-1)}^{kv}(1 - \frac{0,1}{2}) = t_4^{kv}(0,95) = 2,77$ je kvantil t -rozdelenia. Pretože $|t| > t_4^{kv}(0,95)$, namerané hodnoty padnú do kritickej oblasti W , a preto prijímame hypotézu $H_1 : b_2 \neq 0$, čo slovne formulujeme tak, že *bolo štatisticky preukázané, že samostatné pôsobenie faktora t_2 má na hodnotu regresnej funkcie podstatný vplyv.*

Príklad 54. Uvažujme experiment typu 2^2 , kde regresná funkcia je lineárneho typu, pričom v každom zo 4 bodov sme vykonali po 3 merania. Získali sme tak nasledujúce údaje, ktoré po vykonaní príslušnej lineárnej transformácie T , kde

$$T : t_1 = \frac{x_1 - s_1}{\lambda_1}, \quad t_2 = \frac{x_2 - s_2}{\lambda_2}$$

majú tvar:

$$\begin{aligned} y_1 = f_T(-1, -1) = 3, & \quad y_4 = f_T(-1, 1) = 4, & \quad y_7 = f_T(1, -1) = 7, & \quad y_{10} = f_T(1, 1) = -2, \\ y_2 = f_T(-1, -1) = 4, & \quad y_5 = f_T(-1, 1) = 8, & \quad y_8 = f_T(1, -1) = 9, & \quad y_{11} = f_T(1, 1) = 0, \\ y_3 = f_T(-1, -1) = 5, & \quad y_6 = f_T(-1, 1) = 12, & \quad y_9 = f_T(1, -1) = 11, & \quad y_{12} = f_T(1, 1) = 2. \end{aligned}$$

Na hladine $\alpha = 0,05$ vyhodnotíme vplyv samostatného pôsobenia faktora t_1 na hodnotu regresnej funkcie.

Riešenie. Na výpočet $\hat{b}$ -koeficientov regresnej funkcie (ktorá má ten istý tvar ako v minulej úlohe) opäť použijeme vzťah $\hat{\mathbf{b}} = \frac{1}{mk^n} \mathbf{X}'_T \mathbf{Y}$, kde v našom prípade $k = n = 2, m = 3$. Preto

$$\hat{\mathbf{b}} = \frac{1}{3 \cdot 2^2} \cdot \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 4 \\ 5 \\ 4 \\ 8 \\ 12 \\ 7 \\ 9 \\ 11 \\ -2 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 5,25 \\ -0,75 \\ -1,25 \\ -3,25 \end{bmatrix}.$$

Odtiaľ $\hat{b}_1 = -0,75$. Pre vyrovnané hodnoty $\hat{y}_i$ nameraných hodnôt y_i potom máme

$$\begin{aligned} \hat{y}_1 = \hat{y}_2 = \hat{y}_3 &= f_T(-1, -1) = 5,25 - 0,75(-1) - 1,25(-1) - 3,25(-1)(-1) = 4, \\ \hat{y}_4 = \hat{y}_5 = \hat{y}_6 &= f_T(-1, 1) = 5,25 - 0,75(-1) - 1,25(1) - 3,25(-1)(1) = 8, \\ \hat{y}_7 = \hat{y}_8 = \hat{y}_9 &= f_T(1, -1) = 5,25 - 0,75(1) - 1,25(-1) - 3,25(1)(-1) = 9, \\ \hat{y}_{10} = \hat{y}_{11} = \hat{y}_{12} &= f_T(1, 1) = 5,25 - 0,75(1) - 1,25(1) - 3,25(1)(1) = 0. \end{aligned}$$

Na ich základe vypočítame veličinu R_0^2 , kde

$$\begin{aligned} R_0^2 &= \sum_{i=1}^{mk^n} [(y_i - \hat{y}_i)^2] = (3 - 4)^2 + (4 - 4)^2 + (5 - 4)^2 + (4 - 8)^2 + (8 - 8)^2 + \\ &+ (12 - 8)^2 + (7 - 9)^2 + (9 - 9)^2 + (11 - 9)^2 + (-2 - 0)^2 + 0^2 + (2 - 0)^2 = 50. \end{aligned}$$

Vplyv samostatného pôsobenia faktora t_1 na hodnotu transformovanej regresnej funkcie $\hat{f}_T(t_1, t_2)$ zistíme testovaním dvojice hypotéz $H_0 : b_1 = 0 \propto H_1 : b_1 \neq 0$. Kritickú oblasť W tohto testu skonštruujeme pomocou veličín $|t|$, $t_{(m-1)2^2}^{kv}(1 - \frac{\alpha}{2})$, kde

$$|t| = \frac{|\hat{b}_1| k^n \sqrt{m(m-1)}}{\sqrt{R_0^2}} = \frac{|-0,75| \cdot 2^2 \cdot \sqrt{3 \cdot 2}}{\sqrt{50}} = 1,039$$

a $t_{4(m-1)}^{kv}(1 - \frac{\alpha}{2}) = t_{4(3-1)}^{kv}(1 - \frac{0,05}{2}) = t_8^{kv}(0,975) = 2,75$ je kvantil t -rozdelenia. Keďže $|t| < t_8^{kv}(0,975)$, namerané hodnoty nepadnú do kritickej oblasti W , a preto prijímame hypotézu $H_0 : b_1 = 0$, čo slovné formulujeme tak, že *bolo štatisticky preukázané, že samostatné pôsobenie faktora t_1 má na hodnotu regresnej funkcie zanedbateľný vplyv.*

14 LITERATÚRA

Pre ďalšie štúdium je vhodným zdrojom poznatkov nasledujúca literatúra:

- [1] GROFÍK Rudolf, BUBAN Ľubomír: *Štatistika*. Nitra: ES VŠP, 1982. (skriptum).
- [2] LAMOŠ František: *Matematická štatistika pre učiteľské štúdium*. Bratislava: PS UK, 1982 (skriptum).
- [3] ŠKRÁŠEK Josef, TICHÝ Zdeněk: *Základy aplikované matematiky III*. Praha: SNTL, 1990.
- [4] LIKEŠ, Jiří, MACHEK Josef: *Matematická statistika*. Praha: SNTL, 1983.
- [5] CYHELSKÝ Ľubomír: *Statistika v příkladech*. Praha: SNTL /SVTL, 1967.
- [6] BÉKÉS Ján, ANDONOV Iva: *Analýza a syntéza strojárskych objektov a procesov*. Bratislava/Sofia: Alfa/Technika, 1986.

Obsah

1 Úvod	3
2 Hustota, distribučná funkcia, kvantilová funkcia a momenty	4
3 Štatistický súbor, znak a variačné triedenie, výberové kvantily a výberové momenty	7
4 Histogram, polygón a ogivná krivka	11
5 Charakteristiky polohy a variability	16
6 Charakteristiky šikmosti a špicatosti	20
7 Lineárny model a metóda najmenších štvorcov	24
8 Metóda maximálnej vierohodnosti a momentová metóda	28
9 Intervaly spoľahlivosti	32
10 Testovanie hypotéz	36
11 Plánovanie experimentu	40
12 Regresná funkcia a transformovaná regresná funkcia	44
13 Vyhodnocovanie experimentu	48
14 Literatúra	52

Autor: Doc. Ing. Oleg Palumbíny, CSc.
Názov: Matematické metódy plánovania a vyhodnocovania
experimentu v príkladoch
Miesto vydania: Trnava
Vydavateľ: AlumniPress
Rok vydania: 2008
Vydanie: prvé
Rozsah: 53 strán
Edičné číslo: 5/AP/2008

ISBN 978-80-8096-054-4
EAN 9788080960544