# 4 Popisná štatistika

**Príklad 1** *V panelárni bola počas 50 týždňov sledovaná výroba panelov a počty chybných panelov sa zaznamenávali:*

$$14, 13, 15, 12, 14, 9, 15, 8, 14, 10, 16, 12, 13, 17, 18, 13, 13, 10, 14, 16, 11, 14, 12, 9, 15,$$
$$11, 14, 15, 11, 15, 10, 16, 10, 12, 13, 11, 13, 11, 9, 13, 8, 12, 16, 12, 17, 12, 13, 11, 13, 12$$

**a)** *Zostrojte tabuľu početnosti pre 9 tried,*

**b)** *bodovo a intervalovo (s hladinou významnosti 5%) odhadnite strednú hodnotu a smerodajnú odchýlku.*

**a)** *Variačné rozpätie je*

$$R = x_{\max} - x_{\min} = 18 - 8 = 10,$$

*a dĺžka intervalu je*

$$h = \frac{R}{k} = \frac{10}{9} = 1.111 \doteq 1.2.$$

*Tabuľka početnosti bude mať tvar*

| | $x_j$ | $n_j$ | $p_j$ | $N_j$ | $M_j$ |
|---|---|---|---|---|---|
| $[8, 9.2)$ | 8.6 | 5 | $\frac{5}{50}$ | 5 | $\frac{5}{50}$ |
| $[9.2, 10.4)$ | 9.8 | 4 | $\frac{4}{50}$ | 9 | $\frac{9}{50}$ |
| $[10.4, 11.6)$ | 11 | 6 | $\frac{6}{50}$ | 15 | $\frac{15}{50}$ |
| $[11.6, 12.8)$ | 12.2 | 8 | $\frac{8}{50}$ | 23 | $\frac{23}{50}$ |
| $[12.8, 14)$ | 13.4 | 9 | $\frac{9}{50}$ | 32 | $\frac{32}{50}$ |
| $[14, 15.2)$ | 14.6 | 11 | $\frac{11}{50}$ | 43 | $\frac{43}{50}$ |
| $[15.2, 16.4)$ | 15.8 | 4 | $\frac{4}{50}$ | 47 | $\frac{47}{50}$ |
| $[16.4, 17.6)$ | 17 | 2 | $\frac{2}{50}$ | 49 | $\frac{49}{50}$ |
| $[17.6, 18.8)$ | 18.2 | 1 | $\frac{1}{50}$ | 50 | $\frac{50}{50} = 1$ |

**b)** *Najlepším bodovým odhadom strednej hodnoty je výberový priemer*

$$\begin{aligned}\overline{x} &= \frac{1}{50}\sum_{j=1}^{9} x_j \cdot n_j = \frac{1}{50}(8.6 \cdot 5 + 9.8 \cdot 4 + 11 \cdot 6 + 12.2 \cdot 8 \\ &\quad +13.4 \cdot 9 + 14.6 \cdot 11 + 15.8 \cdot 4 + 17 \cdot 2 + 18.2 \cdot 1) \\ &\doteq 12.848.\end{aligned}$$

*Najlepším bodovým odhadom rozptylu je výberový rozptyl*

$$\begin{aligned}S_x^2 &= \frac{1}{50-1}\sum_{j=1}^{9} (x_j - \overline{x}) \cdot n_j = \frac{1}{49}((8.6 - 12.848)^2 \cdot 5 + (9.8 - 12.848)^2 \cdot 4 \\ &\quad + (11 - 12.848)^2 \cdot 6 + (12.2 - 12.848)^2 \cdot 8 + (13.4 - 12.848)^2 \cdot 9 \\ &\quad + (14.6 - 12.848)^2 \cdot 11 + (15.8 - 12.848)^2 \cdot 4 + (17 - 12.848)^2 \cdot 2 \\ &\quad + (18.2 - 12.848)^2 \cdot 1) \\ &\doteq 5.831\end{aligned}$$

*a smerodajnej odchýlky je výberová smerodajná odchýlka*

$$S_x = \sqrt{S_x^2} = \sqrt{5.831} = 2.415.$$

*Intervalový odhad pre strednú hodnotu* $\mu = E(\xi)$ *je*

$$\begin{aligned}
\overline{x} - \frac{S_x}{\sqrt{n}} \cdot u_{1-\frac{\alpha}{2}} &\leq \mu \leq \overline{x} + \frac{S_x}{\sqrt{n}} \cdot u_{1-\frac{\alpha}{2}}, \\
12.848 - \frac{2.415}{\sqrt{50}} \cdot u_{0.975} &\leq \mu \leq 12.848 + \frac{2.415}{\sqrt{50}} \cdot u_{0.975}, \\
12.848 - \frac{2.415}{\sqrt{50}} \cdot 1.96 &\leq \mu \leq 12.848 + \frac{2.415}{\sqrt{50}} \cdot 1.96, \\
12.848 - 2.415 &\leq \mu \leq 12.848 + 2.415, \\
10.433 &\leq \mu \leq 15.263
\end{aligned}$$

*a podobne je intervalový odhad pre smerodajnú odchýlku*

**Príklad 2** *Merala sa minimálna kapacita* $n = 60$ *dolaďovacích kondenzátorov s týmito výsledkami:*

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 2.04 | 2.37 | 2.27 | 2.06 | 2.34 | 2.07 | 2.09 | 2.31 | 2.14 | 2.27 |
| 2.02 | 2.51 | 2.16 | 2.31 | 2.23 | 2.51 | 2.27 | 2.39 | 2.58 | 2.40 |
| 2.00 | 2.29 | 2.17 | 2.34 | 2.17 | 2.41 | 2.22 | 2.10 | 2.22 | 2.19 |
| 2.47 | 2.37 | 2.52 | 2.41 | 2.37 | 2.14 | 2.33 | 2.53 | 2.34 | 2.43 |
| 2.17 | 2.54 | 2.12 | 2.58 | 2.09 | 2.47 | 2.17 | 2.27 | 2.24 | 2.63 |
| 2.28 | 2.11 | 2.39 | 2.38 | 2.51 | 2.41 | 2.43 | 2.61 | 2.48 | 2.63 |

**a)** Roztrieďte štatistický súbor do 7 tried, určte tabuľku početností,

**b)** bodovo odhadnite strednú hodnotu $E(\xi)$ a smerodajnú odchýlku $\sigma$ skupinového rozdelenia a následne nájdite interval spoľahlivosti pre strednú hodnotu a rozptyl s hladinou významnosti $\alpha = 5\%$.

**a)** Najskôr určíme variačné rozpätie

$$R = x_{\max} - x_{\min} = 2.63 - 2.00 = 0.63,$$

odtiaľ pre šírku intervalu (triedy) bude zrejme platiť

$$h = \frac{R}{k} = \frac{0.63}{7} = 0.09 \doteq 0.1.$$

A tabuľka skupinového rozdlenia početnosti má tvar

| | $x_j$ | $n_j$ | $p_j$ | $N_j = \sum_{i=1}^{j} n_j$ | $M_j$ |
|---|---|---|---|---|---|
| $[2.00, 2.10)$ | 2.05 | 7 | $\frac{7}{60}$ | 7 | $\frac{7}{60}$ |
| $[2.10, 2.20)$ | 2.15 | 11 | $\frac{11}{60}$ | 18 | $\frac{18}{60}$ |
| $[2.20, 2.30)$ | 2.25 | 10 | $\frac{10}{60}$ | 28 | $\frac{28}{60}$ |
| $[2.30, 2.40)$ | 2.35 | 12 | $\frac{12}{60}$ | 40 | $\frac{40}{60}$ |
| $[2.40, 2.50)$ | 2.45 | 9 | $\frac{9}{60}$ | 49 | $\frac{49}{60}$ |
| $[2.50, 2.60)$ | 2.55 | 8 | $\frac{8}{60}$ | 57 | $\frac{57}{60}$ |
| $[2.60, 2.70)$ | 2.65 | 3 | $\frac{3}{60}$ | 60 | $\frac{60}{60} = 1$ |
| | | $\sum = 60$ | | | |

**b)** Najlepším bodovým odhadom strednej hodnoty $E(\xi) = \mu$ je výberový priemer

$$
\begin{aligned}
\overline{x} &= \frac{1}{n} \sum_{j=1}^{k} x_j \cdot n_j \\
&= \frac{1}{60} (2.05 \cdot 7 + 2.15 \cdot 11 + 2.25 \cdot 10 + 2.35 \cdot 12 + 2.45 \cdot 9 + 2.55 \cdot 8 + 2.65 \cdot 3) \\
&\doteq 2.318.
\end{aligned}
$$

Najlepší bodový odhad rozptylu $D(\xi) = \sigma^2$ je výberový rozptyl

$$
\begin{aligned}
S_x^2 &= \frac{1}{n-1} \sum_{j=1}^{k} (x_j - \overline{x})^2 \cdot n_j \\
&= \frac{1}{59} ((2.05 - 2.318)^2 \cdot 7 + (2.15 - 2.318)^2 \cdot 11 \\
&\quad + (2.25 - 2.318)^2 \cdot 10 + (2.35 - 2.318)^2 \cdot 12 \\
&\quad + (2.45 - 2.318)^2 \cdot 9 + (2.55 - 2.318)^2 \cdot 8 \\
&\quad + (2.65 - 2.318)^2 \cdot 3 \\
&\doteq 0.0303.
\end{aligned}
$$

A bodovým odhadom smerodajnej odchýlky $\sigma$ je výberová smerodajná odchýlka

$$S_x = \sqrt{S_x^2} = \sqrt{0.0303} \doteq 0.174.$$

Interval spoľahlivosti pre strednú hodnotu na hladine významnosti $\alpha = 5\% =$

0.05 je pre veľký súbor ($n = 60 > 30$)

$$\begin{aligned}
\overline{x} - \frac{S_x}{\sqrt{n}} \cdot u_{1-\frac{\alpha}{2}} &\leq \mu \leq \overline{x} + \frac{S_x}{\sqrt{n}} \cdot u_{1-\frac{\alpha}{2}}, \\
2.318 - \frac{0.174}{\sqrt{60}} \cdot u_{0.975} &\leq \mu \leq 2.318 + \frac{0.174}{\sqrt{60}} \cdot u_{0.975}, \\
2.318 - \frac{0.174}{\sqrt{60}} \cdot 1.96 &\leq \mu \leq 2.318 + \frac{0.174}{\sqrt{60}} \cdot 1.96, \\
2.318 - 0.044 &\leq \mu \leq 2.318 + 0.044, \\
2.274 &\leq \mu \leq 2.362.
\end{aligned}$$

Interval spoľahlivosti pre rozptyl (resp. smerodajnú odchýlka)

$$\begin{aligned}
\frac{(n-1) \cdot S_x^2}{\chi^2_{\frac{\alpha}{2}}} &\leq \sigma^2 \leq \frac{(n-1) \cdot S_x^2}{\chi^2_{1-\frac{\alpha}{2}}}, \\
\frac{59 \cdot 0.0303}{\chi^2_{0.025}} &\leq \sigma^2 \leq \frac{59 \cdot 0.0303}{\chi^2_{0.975}}, \\
\frac{59 \cdot 0.0303}{83.30} &\leq \sigma^2 \leq \frac{59 \cdot 0.0303}{40.48}, \\
0.0215 &\leq \sigma^2 \leq 0.0442
\end{aligned}$$

odkiaľ

$$\begin{aligned}
\sqrt{0.0215} &\leq \sigma \leq \sqrt{0.0442}, \\
0.147 &\leq \sigma \leq 0.210\,.
\end{aligned}$$

**Príklad 3** *Meraním tvrdosti legovanej ocele sme získali hodnoty:*

13.1; 12.8; 11.9; 12.4; 13.5; 13.7; 12.0; 13.8; 10.6; 12.4; 13.5; 11.7; 13.9; 11.5; 12.5; 11.9;
12.1; 12.7; 12.9; 11.5; 11.9; 12.8; 11.4; 11.7; 11.4; 13.4; 12.5; 11.5; 12.8; 12.3; 11.7; 11.4.

**a)** *Roztriedte do 5 tried a zapíšte do tabuľky rozdelenia početnosti,*

**b)** *bodovo odhadnúť strednú hodnotu, rozptyl (resp. smerodajnú odchýlku),*

**c)** *nájdite intervalový odhad na hladine spoľahlivosti $\alpha = 5\%$ pre strednú hodnotu a rozptyl (resp. smerodajnú odchýlku).*

**a)** *Variačné rozpätie je*

$$R = x_{\max} - x_{\min} = 13.9 - 10.6 = 3.3,$$

*a následne vypočítame šírku intervalu*

$$h = \frac{R}{k} = \frac{3.3}{5} = 0.66 \doteq 0.7,$$

*odkiaľ získame* ***tabuľku rozdelenia početností***

| | $x_j$ | $n_j$ | $p_j = \frac{n_j}{N}$ | $N_j = \sum_{i=1}^{j} n_j$ | $M_j = \frac{N_j}{N}$ |
|---|---|---|---|---|---|
| $[10.6, 11.3)$ | 10.95 | 1 | $\frac{1}{32}$ | 1 | $\frac{1}{32}$ |
| $[11.3, 12.0)$ | 11.65 | 12 | $\frac{12}{32}$ | 13 | $\frac{13}{32}$ |
| $[12.0, 12.7)$ | 12.35 | 7 | $\frac{7}{32}$ | 20 | $\frac{20}{32}$ |
| $[12.7, 13.4)$ | 13.05 | 6 | $\frac{6}{32}$ | 26 | $\frac{26}{32}$ |
| $[13.4, 14.1)$ | 13.75 | 6 | $\frac{6}{32}$ | 32 | $\frac{32}{32} = 1$ |
| | | $\sum = 32$ | | | |

**b)** *Najlepším nevychýleným bodovým odhadom strednej hodnoty* $E(\xi) = \mu$ *je skupinový výberový priemer*

$$\begin{aligned}
\overline{x} &= \frac{1}{N}\sum_{j=1}^{k} x_j \cdot n_j = \frac{1}{32}\sum_{j=1}^{5} x_j \cdot n_j \\
&= \frac{1}{32}\left(10.95 \cdot 1 + 11.65 \cdot 12 + 12.35 \cdot 7 + 13.05 \cdot 6 + 13.75 \cdot 6\right) \\
&\doteq 12.44,
\end{aligned}$$

*podobne najlepším nevychýleným bodovým odhadom rozptylu* $D(\xi) = \sigma^2$ *je skupinový výberový rozptyl*

$$\begin{aligned}
S_x^2 &= \frac{1}{N-1}\sum_{j=1}^{k} \left(x_j - \overline{x}\right)^2 \cdot n_j = \frac{1}{31}\sum_{j=1}^{5} \left(x_j - 12.44\right)^2 \cdot n_j \\
&= \frac{1}{31}\Big((10.95 - 12.44)^2 \cdot 1 + (11.65 - 12.44)^2 \cdot 12 \\
&\quad + (12.35 - 12.44)^2 \cdot 7 + (13.05 - 12.44)^2 \cdot 6 \\
&\quad + (13.75 - 12.44)^2 \cdot 6\Big) \\
&\doteq 0.72,
\end{aligned}$$

*a taktiež najlepším nevychýleným bodovým odhadom smerodajnej odchýlky* $\sigma$ *je skupinová výberová odchýlka*

$$S_x = \sqrt{S_x^2} = \sqrt{0.72} \doteq 0.85.$$

**c)** *Interval spoľahlivosti pre strednú hodnotu s hladinou významnosti* $\alpha = 5\%$

*má, nakoľko súbor je veľký $n = 32 > 30$, tvar*

$$\begin{aligned}
\overline{x} - \frac{S_x}{\sqrt{n}} \cdot u_{1-\frac{\alpha}{2}} &\leq \mu \leq \overline{x} + \frac{S_x}{\sqrt{n}} \cdot u_{1-\frac{\alpha}{2}}, \\
12.\,44 - \frac{0.85}{\sqrt{32}} \cdot u_{1-\frac{5\%}{2}} &\leq \mu \leq 12.\,44 + \frac{0.85}{\sqrt{32}} \cdot u_{1-\frac{0.05}{2}}, \\
12.\,44 - \frac{0.85}{\sqrt{32}} \cdot u_{1-0.025} &\leq \mu \leq 12.\,44 + \frac{0.85}{\sqrt{32}} \cdot u_{0.975}, \\
12.\,44 - \frac{0.85}{\sqrt{32}} \cdot 1.96 &\leq \mu \leq 12.\,44 + \frac{0.85}{\sqrt{32}} \cdot 1.96, \\
12.\,44 - 0.30 &\leq \mu \leq 12.\,44 + 0.30, \\
12.13 &\leq \mu \leq 12.73.
\end{aligned}$$

*Podobne interval spoľahlivosti pre rozptyl (smerodajnú odchýlku) s hladinou významnosti $\alpha = 5\%$ má tvar*

$$\begin{aligned}
\frac{(n-1) \cdot S_x^2}{\chi^2_{\frac{\alpha}{2}}} &\leq \sigma^2 \leq \frac{(n-1) \cdot S_x^2}{\chi^2_{1-\frac{\alpha}{2}}}, \\
\frac{31 \cdot 0.72}{\chi^2_{0.025}} &\leq \sigma^2 \leq \frac{31 \cdot 0.72}{\chi^2_{0.975}}, \\
\frac{31 \cdot 0.72}{46.98} &\leq \sigma^2 \leq \frac{31 \cdot 0.72}{16.79}, \\
0.48 &\leq \sigma^2 \leq 1.33,
\end{aligned}$$

*odkiaľ zrejme platí*

$$\begin{aligned}
\sqrt{0.48} &\leq \sigma \leq \sqrt{1.33}, \\
0.69 &\leq \sigma \leq 1.\,15.
\end{aligned}$$