Intervalové odhady II

1.) V následující tabulce jsou uvedeny výsledky měření odporu (v Ω) dvou drátů A a B:

drát A	0,140	0,138	0,143	0,142	0,144	0,137
drát B	0,135	0,140	0,142	0,136	0,138	

K měření byly použity měřicí přístroje s rozptyly $\sigma_1^2 = 4\cdot 10^{-6}$, resp. $\sigma_2^2 = 9\cdot 10^{-6}$. Stanovte 95% dolní (levostranný) IS pro rozdíl středních hodnot odporu drátů.

X_i = výsledek i-tého měření drátu A

Y_i = výsledek -||- B

známe

$X_1, \dots, X_6$ náh. výběr z $N(\mu_1, \sigma_1^2)$

$Y_1, \dots, Y_5$ náh. výběr z $N(\mu_2, \sigma_2^2)$

vzájemně nezávislé

známe

• bodový odhad $\mu_1 - \mu_2$ ... $\overline{X} - \overline{Y}$

$$\left.\begin{aligned} \overline{X} &\sim N\left(\mu_1, \frac{\sigma_1^2}{6}\right) \\ \overline{Y} &\sim N\left(\mu_2, \frac{\sigma_2^2}{5}\right) \end{aligned}\right\} \quad \overline{X} - \overline{Y} \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{6} + \frac{\sigma_2^2}{5}\right)$$

$$\underbrace{\frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{6} + \frac{\sigma_2^2}{5}}}}_{\text{pivotová statistika}} \sim N(0,1)$$

$$P\left(\frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{6} + \frac{\sigma_2^2}{5}}} \le u_{1-\alpha}\right) = 1 - \alpha \quad , \forall \mu_1, \mu_2.$$

$$P\Bigg(\overbrace{\overline{X} - \overline{Y} - u_{1-\alpha}\sqrt{\frac{\sigma_1^2}{6} + \frac{\sigma_2^2}{5}}}^{D(X_1, \dots, X_6, Y_1, \dots, Y_5)} \le \mu_1 - \mu_2\Bigg) = 1 - \alpha \quad , \forall \mu_1, \mu_2$$

- - - -

pro naše data. Hledaný IS je $\langle -0{,}000113 ; \infty)$.

2.) Zkoumáme výšku prvňáčků. Vybrali jsme 10 dívek a 15 chlapců, změřili jejich výšku a získali násl. hodnoty:

průměrná výška dívek je 120,56 a příslušný výběrový rozptyl 9,14 ,

chlapců 124,13 8,95.

Určete 90 % IS pro rozdíl středních hodnot výšek dívek a chlapců. Najděte i 95 % horní (pravostranný) IS.

X_i = výška i-té dívky

Y_i = výška i-tého chlapce

označ. $n_1 = 10$, $n_2 = 15$

neznámé, ale stejné v obou výběrech

$X_1, \dots, X_{10}$ náh. výběr z $N(\mu_1, \sigma^2)$

$Y_1, \dots, Y_{15}$ náh. výběr z $N(\mu_2, \sigma^2)$

} vzájemně nezávislé

- bodový odhad $\mu_1 - \mu_2$ je opět $\overline{X} - \overline{Y}$

$$\overline{X} - \overline{Y} \sim N\left(\mu_1 - \mu_2 ; \frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}\right)$$

$$\frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{\sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim N(0,1)$$

neznáme, nahradíme jeho odhadem $S_*^2 = \dfrac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$, kde S_1^2 je výběrový rozptyl n. v. $X_1, \dots, X_{n_1}$

S_2^2 $Y_1, \dots, Y_{n_2}$

$$\underbrace{\frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{S_*^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}}_{\text{pivotová statistika}} \sim t(n_1 + n_2 - 2)$$

OBOUSTRANNÝ:

$$P\left(-t_{1-\frac{\alpha}{2}}(n_1+n_2-2) \leq \frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{S_* \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \leq t_{1-\frac{\alpha}{2}}(n_1+n_2-2)\right) = 1-\alpha, \quad \forall \mu_1, \mu_2$$

PRAVOSTRANNÝ:

$$P\left(\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{S_* \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \geq -t_{1-\alpha}(n_1+n_2-2)\right) = 1-\alpha, \quad \forall \mu_1, \mu_2$$

oboustranný IS:

$$P\left(\underbrace{\overline{X}-\overline{Y}-t_{1-\frac{\alpha}{2}}(n_1+n_2-2) S_* \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}_{D} \leq \mu_1-\mu_2 \leq \underbrace{\overline{X}-\overline{Y}+t_{1-\frac{\alpha}{2}}(n_1+n_2-2) S_* \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}_{H}\right) = 1-\alpha$$

pravostranný (horní) IS:

$$P\left(\mu_1-\mu_2 \leq \underbrace{\overline{X}-\overline{Y}+t_{1-\alpha}(n_1+n_2-2) S_* \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}_{H}\right) = 1-\alpha, \quad \forall \mu_1, \mu_2.$$

data:

$\overline{x} = 120{,}56$ $\quad \overline{y} = 124{,}13$ $\quad t_{0,95}(23) \doteq 1{,}71$

$s_1^2 = 9{,}14$ $\quad s_2^2 = 8{,}95$

$n_1 = 10$ $\quad n_2 = 15$

$$s_*^2 = \frac{9 \cdot 9{,}14 + 14 \cdot 8{,}95}{23} = 9{,}024$$

90% IS je $\langle -5{,}67; -1{,}47 \rangle$.

95% horní IS je $(-\infty, -1{,}47\rangle$.

3) Bylo zkoumáno množství SiO_2 ve jednom vzorku pomocí 2 různých metod.

analyticky 20,1 19,6 20,0 19,9 20,1

fotokolorimetricky 20,9 20,1 20,6 20,5 20,7 20,5

Sestrojte 90% IS pro podíl rozptylů jednotlivých typů měření.

X_i = výsledek i-tého měření anal.

Y_i = -||- fotokol.

$$\left.\begin{array}{l} X_1, \dots, X_5 \text{ náh. výběr z } N(\mu_1, \sigma_1^2) \\ Y_1, \dots, Y_6 \text{ náh. výběr z } N(\mu_2, \sigma_2^2) \end{array}\right\} \text{vzájemně nezávislé}$$

- bodový odhad $\frac{\sigma_1^2}{\sigma_2^2}$ je $\frac{S_1^2}{S_2^2}$, kde $S_1^2 = \frac{1}{n_1-1} \sum_{i=1}^{n_1} (X_i - \bar{X})^2$ a $S_2^2 = \frac{1}{n_2-1} \sum_{i=1}^{n_2} (Y_i - \bar{Y})^2$.

pivotová statistika

$$\frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \sim F(n_1-1, n_2-1)$$

$$P\left(F_{\frac{\alpha}{2}}(n_1-1, n_2-1) \leq \frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \leq F_{1-\frac{\alpha}{2}}(n_1-1, n_2-1) \right) = 1-\alpha, \quad \forall \sigma_1^2 > 0, \sigma_2^2 > 0.$$

$$P\left(\underbrace{\frac{S_1^2}{S_2^2} \cdot \frac{1}{F_{1-\frac{\alpha}{2}}(n_1-1, n_2-1)}}_{D} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \underbrace{\frac{S_1^2}{S_2^2} \cdot \frac{1}{F_{\frac{\alpha}{2}}(n_1-1, n_2-1)}}_{H} \right) = 1-\alpha.$$

data:

$$\left.\begin{array}{l} s_1^2 = 0{,}043 \\ s_2^2 = 0{,}071 \\ F_{0,05}(4,5) = 0{,}160 \\ F_{0,95}(4,5) = 5{,}19 \end{array}\right\} 90\% \text{ IS je } \langle 0{,}117; 3{,}79 \rangle$$

Platí:

$$F_{1-\alpha}(n, m) = \frac{1}{F_\alpha(m, n)}.$$

4) Ve firmě pracuje 8 000 zaměstnanců. 160 z nich bylo náhodně vybráno a dotázáno, zda do práce cestují vlakem. 48 uvedlo, že ano. Najděte 95% IS pro podíl a počet zaměstnanců cestujících vlakem.

$$X_i \begin{cases} 1 & \dots \text{ i-tý zaměstnanec cestuje vlakem,} \\ 0 & \dots \text{ jinak.} \end{cases}$$

$X_1, \dots, X_{160}$ náh. výběr z $a(p)$, kde $p = P(\text{i-tý zaměstnanec cestuje vlakem})$

- bodový odhad p je $\hat{p} = \overline{X}$.

podle CLV $\overline{X} \approx N\left(p; \frac{p(1-p)}{n}\right)$

$$\underbrace{\frac{\overline{X}-p}{\sqrt{\frac{p(1-p)}{n}}}}_{\substack{\text{lze použít jako pivotovou statistiku} \\ \text{je to však výpočetně náročnější!}}} \approx N(0,1)$$

$p(1-p)$ nahradíme jeho odhadem $\overline{X}(1-\overline{X})$

$$\frac{\overline{X}-p}{\underbrace{\sqrt{\frac{\overline{X}(1-\overline{X})}{n}}}_{\text{pivotová statistika}}} \approx N(0,1)$$

$$P\left(-u_{1-\frac{\alpha}{2}} \le \frac{\overline{X}-p}{\sqrt{\frac{\overline{X}(1-\overline{X})}{n}}} \le u_{1-\frac{\alpha}{2}}\right) \doteq 1-\alpha, \ \forall p$$

$$P\Bigg(\underbrace{\overline{X} - u_{1-\frac{\alpha}{2}}\sqrt{\frac{\overline{X}(1-\overline{X})}{n}}}_{D} \le p \le \underbrace{\overline{X} + u_{1-\frac{\alpha}{2}}\sqrt{\frac{\overline{X}(1-\overline{X})}{n}}}_{H}\Bigg) \doteq 1-\alpha$$

IS pro počet získáme vynásobením 8000, tj.
$8000 \cdot \langle 0{,}229; 0{,}371\rangle = \langle 1832; 2968\rangle$.

dosaz. $\overline{x} = \frac{48}{160} = 0{,}3$

$u_{0{,}975} = 1{,}96$

$n = 160$

$\Big\}$ 95% IS pro podíl je $\langle 0{,}229; 0{,}371\rangle$.