Intervalové odhady (intervaly spolehlivosti)

MODEL: $X_1, \dots, X_n$ je náhodný výběr z rozdělení s distr. fcí $F(x, \theta)$, $\theta \in \Theta$ je neznámý parametr.

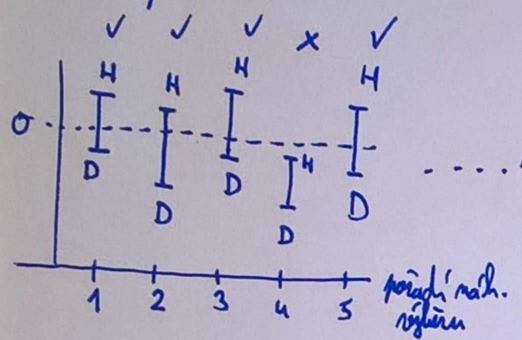
Chceme najít interval, který s danou pravděpodobností překryje skutečnou hodnotu parametru θ.

Interval $\langle D, H \rangle = \langle D(X_1, \dots, X_n), H(X_1, \dots, X_n) \rangle$ nazveme $100(1-\alpha)\%$ interval spolehlivosti pro parametr θ,

jestliže $P\big(D(X_1, \dots, X_n) \le \theta \le H(X_1, \dots, X_n)\big) = 1 - \alpha$, $\forall \theta \in \Theta$.

Pozn.

$1-\alpha$ ... spolehlivost



.... relativní četnost intervalů obsahující $\theta \doteq 1-\alpha$.

1.) Rychlost letadla byla měřena při 5 nezávislých zkouškách. Výsledky byly následující:

867,5 870 874 868,5 871,5 (m/s). Přesnost měřícího přístroje dána směrodatnou odchylkou měření je 2,1 m/s. Najděte 95% interval spolehlivosti (IS) pro skutečnou rychlost.

$\sigma = 2{,}1 \Rightarrow \sigma^2 = 2{,}1^2$

X_i = rychlost při i-té zkoušce

$X_1, \dots, X_5$ náh. výběr z $N(\mu; 2{,}1^2)$

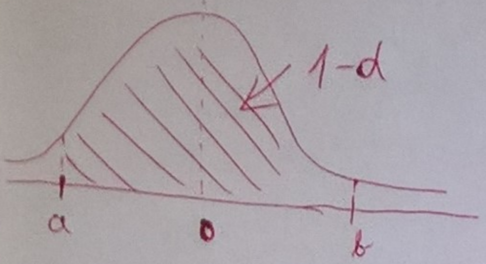
- bodový odhad μ

- metoda momentů, metoda max. věrohodnosti, věta 1.6 $\Rightarrow \hat{\mu} = \overline{X}$.

toto rozdělení nezávisí na μ

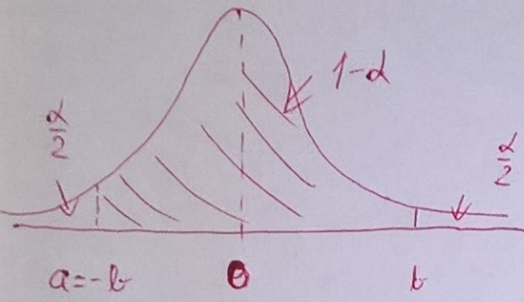
$$\overline{X} \sim N\left(\mu, \frac{2{,}1^2}{n}\right) \overset{\text{standardizace}}{\Rightarrow} \frac{\overline{X} - \mu}{2{,}1}\sqrt{n} \sim N(0,1)$$

pivotová statistika = její rozdělení nezávisí na neznámém parametru úvaha: Nechť $Y \sim N(0,1)$. Najděte $a, b \in \mathbb{R}$ tak, aby $P(a \le Y \le b) = 1-\alpha$, kde $\alpha > 0$ známé.



... nekonečně mnoho řešení

Přidejme podmínku, aby interval $\langle a, b \rangle$ byl co nejkratší.



$$\Rightarrow a = -b$$

$$P(a \le Y \le b) = P(-b \le Y \le b) = \Phi(b) - \underbrace{\Phi(-b)}_{1-\Phi(b)} =$$

$$= -1 + 2\Phi(b) = 1 - \alpha$$

$$\Phi(b) = 1 - \frac{\alpha}{2}$$

$$\boxed{b = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) = u_{1-\frac{\alpha}{2}}}$$

$$P\left(-u_{1-\frac{\alpha}{2}} \le Y \le u_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$$ + pivotová statistika $\frac{\overline{X}-\mu}{2{,}1}\sqrt{n} \sim N(0,1)$

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

ekvivalentní úpravy $\rightsquigarrow P(\ldots \le \mu \le \ldots) = 1-\alpha$

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

$$P\left(-\overline{X} - \frac{2{,}1}{\sqrt{n}} u_{1-\frac{\alpha}{2}} \le -\mu \le -\overline{X} + \frac{2{,}1}{\sqrt{n}} u_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

$$P\Big(\underbrace{\overline{X} - \frac{2{,}1}{\sqrt{n}} u_{1-\frac{\alpha}{2}}}_{D(X_1,\ldots,X_n)} \le \mu \le \underbrace{\overline{X} + \frac{2{,}1}{\sqrt{n}} u_{1-\frac{\alpha}{2}}}_{H(X_1,\ldots,X_n)}\Big) = 1 - \alpha, \ \forall \mu$$

Najděme interval pro naše data:

$$1-\alpha = 0{,}95 \Rightarrow \alpha = 0{,}05 \Rightarrow 1-\frac{\alpha}{2} = 0{,}975$$

$$u_{0,975} = u_{1-\frac{\alpha}{2}} = 1{,}96$$

$$\overline{X} = \frac{1}{5}(867{,}5 + 870 + 874 + 868{,}5 + 871{,}5) = 870{,}3$$

$$n = 5$$

$$\langle D, H \rangle = \left\langle \overline{X} - \frac{2{,}1}{\sqrt{n}} u_{1-\frac{\alpha}{2}},\ \overline{X} + \frac{2{,}1}{\sqrt{n}} u_{1-\frac{\alpha}{2}} \right\rangle$$

$$\underline{\langle 868{,}5 ; 872{,}1 \rangle}.$$

2.) Měříme pH vody. Odebrali jsme 9 vzorků a dostali následující výsledky:

7,20 7,09 6,92 7,18 6,79 6,98 6,86 7,18 7,16.

Najděte 90% IS pro pH vody.

X_i = pH naměřené na i-tém vzorku

$X_1, \dots, X_9$ náh. výběr z $N(\mu, \sigma^2)$ ← neznáme (rušivý parametr)

• bodový odhad μ

– metoda momentů, max. věrohodnosti, věta 1.6 $\Rightarrow \hat{\mu} = \overline{X}$.

$$\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \overset{\text{stand.}}{\Longrightarrow} \underbrace{\frac{\overline{X}-\mu}{\sigma}\sqrt{n}}_{\text{závisí na rušivém parametru } \sigma^2 \text{ (není pivotová statistika)}} \sim \overbrace{N(0,1)}^{\text{nezávisí na } \mu}$$

σ^2 odhadneme pomocí $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$ (výběrový rozptyl)

$$\underbrace{\underbrace{\frac{\overline{X}-\mu}{S}\sqrt{n}}_{\text{pivotová statistika}} \sim t(n-1) \qquad + \quad T \sim t(n-1) \quad P\left(-t_{1-\frac{\alpha}{2}}(n-1) \le T \le t_{1-\frac{\alpha}{2}}(n-1)\right) = 1-\alpha}$$

$$P\left(-t_{1-\frac{\alpha}{2}}(n-1) \le \frac{\overline{X}-\mu}{S}\sqrt{n} \le t_{1-\frac{\alpha}{2}}(n-1)\right) = 1-\alpha, \quad \forall \mu.$$

Ekvivalentní úpravy . . .

$$P\Big(\underbrace{\overline{X} - \frac{S}{\sqrt{n}}\, t_{1-\frac{\alpha}{2}}(n-1)}_{D(X_1,\dots,X_n)} \le \mu \le \underbrace{\overline{X} + \frac{S}{\sqrt{n}}\, t_{1-\frac{\alpha}{2}}(n-1)}_{H(X_1,\dots,X_n)}\Big) = 1-\alpha\,, \quad \forall \mu$$

pro naše data:

$$\left.\begin{aligned}
&1-\alpha = 0{,}9 \Rightarrow \alpha = 0{,}1 \Rightarrow 1-\tfrac{\alpha}{2} = 0{,}95 \\
&t_{0{,}95}(8) = 1{,}86 \\
&\overline{x} = \tfrac{1}{9}(7{,}20 + \dots + 7{,}16) = 7{,}04 \\
&s^2 = \tfrac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x})^2 = \tfrac{1}{n-1}\sum_{i=1}^{n} x_i^2 - \tfrac{n}{n-1}(\overline{x})^2 = 0{,}02433
\end{aligned}\right\} \langle D, H\rangle = \langle 6{,}94 ; 7{,}14\rangle$$

3) Laborant provedl 10 měření jednoho předmětu s následujícími výsledky:

42 48 60 36 50 52 38 56 43 45 .

Sestrojte 95% IS pro rozptyl jeho měření.

X_i = výsledek i-tého měření

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

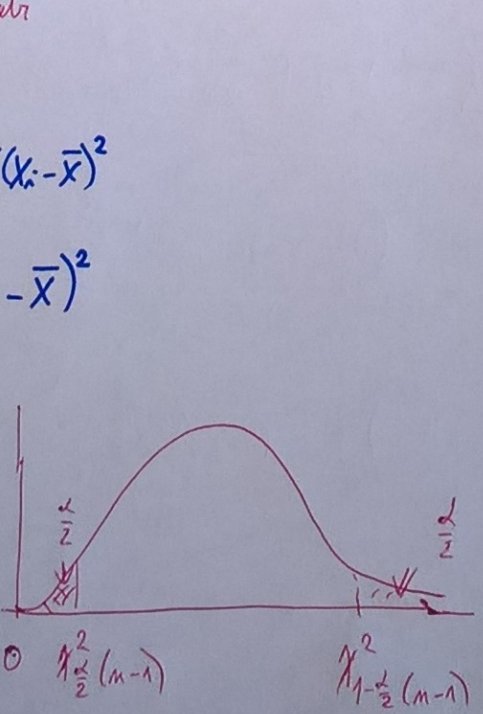
rušivý parametr

- bodový odhad σ^2

- metoda momentů, max. věrohodnosti: $\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$

- věta 1.7 : $\hat{\sigma}^2 = S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$

$$\underbrace{\frac{(n-1)S^2}{\sigma^2}}_{\text{pivotová statistika}} \sim \chi^2(n-1)$$



$$P\left(\chi^2_{\frac{\alpha}{2}}(n-1) \le \frac{(n-1)S^2}{\sigma^2} \le \chi^2_{1-\frac{\alpha}{2}}(n-1)\right) = 1-\alpha, \ \forall \sigma^2$$

$$P\left(\frac{\chi^2_{\frac{\alpha}{2}}(n-1)}{(n-1)S^2} \le \frac{1}{\sigma^2} \le \frac{\chi^2_{1-\frac{\alpha}{2}}(n-1)}{(n-1)S^2}\right) = 1-\alpha$$

$$P\Bigg(\underbrace{\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}}_{D(X_1,\dots,X_n)} \le \sigma^2 \le \underbrace{\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}}_{H(X_1,\dots,X_n)}\Bigg) = 1-\alpha, \ \forall \sigma^2.$$

data:

$$\left.\begin{aligned} \chi^2_{0{,}025}(9) &= 2{,}7 \\ \chi^2_{0{,}975}(9) &= 19{,}023 \\ \bar{x} &= 47 \\ s^2 &= 59{,}11 \end{aligned}\right\} \langle D, H\rangle = \langle 28; 197\rangle$$

Jednostranné IS, intervaly typu $(-\infty, H\rangle$, $\langle D, \infty)$:

$D = D(X_1,\dots,X_n)$ je dolní (levostranný) $100(1-\alpha)\%$ IS pro parametr θ, jestliže:

$$P(D(X_1,\dots,X_n) \le \theta) = 1-\alpha, \ \forall \theta.$$

$H = H(X_1,\dots,X_n)$ je horní (pravostranný) $100(1-\alpha)\%$ IS pro parametr θ, jestliže:

$$P(\theta \le H(X_1,\dots,X_n)) = 1-\alpha, \ \forall \theta.$$

Cvičení:

Sestrojte dolní i horní 95% IS pro rozptyl z předchozího příkladu:

horní ... $\langle 0; 160\rangle$ dolní ... $\langle 31{,}4; \infty)$