Confidence Intervals

- We measure the blood pressure of 20 persons having diabetes (1st number in mmHg)

  we get

  $$X_1 = 134 \; ; \quad X_2 = 141 \; ; \; \ldots \; , \quad X_{20} = 137$$

  with

  $$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \underset{n=20}{\overset{\uparrow}{=}} 135{,}6$$

  and

  $$S_x^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2 = 12{,}6$$

- Model: X_i are observations of $\mathcal{N}(\mu ; \sigma^2)$

  with μ and σ^2 unknown.

- Estimate μ! $\rightsquigarrow$ $\mu^* = \overline{X} = 135{,}6$ //

  But how confident can we be about this estimate?

- Confidence intervals: A confidence interval with level of significance $1-\alpha$ for μ is a random interval I_μ such that

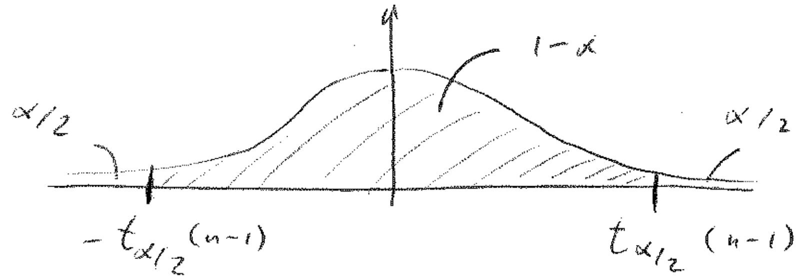
  $$P(\mu \in I_\mu) \geq 1 - \alpha.$$

  [common values for α: 0,05 ; 0,01]

- we calculate a 95 % - CI for μ:

Formel sampling: $$\frac{\bar{X} - \mu}{s/\sqrt{n}} \sim t(n-1).$$

density of $t(n-1)$



we have

$$1-\alpha = \mathbb{P}\left(-t_{\alpha/2}(n-1) \le \frac{\bar{X}-\mu}{s/\sqrt{n}} \le t_{\alpha/2}(n-1)\right)$$

$$= \mathbb{P}\left(-t_{\alpha/2}(n-1)\cdot \frac{s}{\sqrt{n}} \le \bar{X} - \mu \le t_{\alpha/2}(n-1)\cdot \frac{s}{\sqrt{n}}\right)$$

$$= \mathbb{P}\left(\bar{X} - t_{\alpha/2}(n-1)\cdot s/\sqrt{n} \le \mu \le \bar{X} + t_{\alpha/2}(n-1)\cdot s/\sqrt{n}\right),$$

i.e. a $1-\alpha$ - CI for μ is

$$I_\mu = \left[\bar{X} - t_{\alpha/2}(n-1)\cdot s/\sqrt{n} \,,\, \bar{X} + t_{\alpha/2}(n-1)\cdot s/\sqrt{n} \right]$$

In our case: $\bar{X} = 135{,}6$; $s_X = \sqrt{12{,}6}$; $n = 20$;

$\alpha = 0{,}05$; $t_{0{,}025}(19) \underset{\substack{\uparrow \\ \text{table 6}}}{=} 2{,}093$

↳ a 95 % - confidence interval for μ is

$$I_\mu = [133{,}92 \,;\, 137{,}25],$$

i.e. $\mu \in I_\mu$ with probability 95 %

- Was it a high value? We compare with a control group. We test 32 random persons and get

$$Y_1 = 116, \; Y_2 = 119, \; \ldots, \; Y_{32} = 124$$

with

$$\bar{Y} = 118{,}3 \quad \text{and} \quad s_Y^2 = 14{,}9.$$

- Model: $Y_1; \ldots; Y_{32}$ are observations of $N(\mu'; \sigma^2)$ with μ' and σ^2 unknown but same σ^2 as before.

- Question: $\mu = \mu'$ or equivalent $\mu - \mu' = 0$?

- We calculate a 95%-confidence interval for $\mu - \mu'$:

Formelsamling:

$$\frac{\bar{X} - \bar{Y} - (\mu - \mu')}{s_p \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$$

$$\text{with} \quad s_p^2 = \frac{(n_1 - 1)\, s_X^2 + (n_2 - 1)\, s_Y^2}{n_1 + n_2 - 2}.$$

$\xrightarrow{\text{as before}}$ a $1-\alpha$ %-confidence interval for $\mu - \mu'$ is

$$I_{\mu - \mu'} = \left[\bar{X} - \bar{Y} \mp t_{\alpha/2}(n_1 + n_2 - 2) \cdot s_p \cdot \sqrt{\tfrac{1}{n_1} + \tfrac{1}{n_2}} \right].$$

In our case: $\bar{X} = 135,6$; $S_X^2 = 12,6$; $n_1 = 20$;

$\bar{Y} = 118,3$; $S_Y = 14,9$, $n_2 = 32$,

$S_p^2 = 14,026$; $\alpha = 0,05$;

$t_{0.025}(50) = 2,0086$

↳ a 95% confidence interval for $\mu - \mu'$

is $I_{\mu-\mu'} = [15,16 ; 19,44]$.

↳ with prob. 95% $\mu - \mu' \in I_{\mu-\mu'}$;

we have $0 \notin I_{\mu-\mu'}$

↳ with prob. 95% $\mu - \mu' \neq 0 \iff$

$\mu = \mu'$

↳ there is a significant difference

between "x - group" and "Y - group".

- "X - group" needs treatement ; we give them medicine and check their blood preasure after one week again. we get

$X_1' = 127$; $X_2' = 137$; ... ; $X_{20}' = 138$.

- Question: Did the medicine have an effect ?

- we have paired observations (stickprov = par), i.e. we are interested in difference $z_i = x_i - x_i'$.

we have

$$z_1 = 7 \; ; \quad z_2 = 4 \; ; \; \dots \; ; \quad z_{20} = -1$$

with $\bar{z} = 4{,}8$ and $s_z^2 = 3{,}26$.

- Model: z_i are observations from $N(\Delta \, ; \, \tilde{\sigma}^2)$

with Δ and $\tilde{\sigma}^2$ unknown.

- we calculate a 95% - confidence interval for Δ:

$$I_\Delta = \left[\, \bar{z} \mp t_{0.025}(19) \cdot s_z / \sqrt{20} \, \right]$$

$$= [\, 3{,}95 \; ; \; 5{,}65 \,]$$

↳ with prob. (at least) 95% $\Delta \in I_\Delta$;

we have $0 \notin I_\Delta$

↳ with prob. (at least) 95% $\Delta \neq 0$

↳ medicine had significant effect //

D. 10

a) Let X be the number of how often A gave a higher value than B.

Then $X \sim Bin(80, p)$ with p unknown.

We observe $x = 52$ and estimate

$$\hat{p} = \frac{52}{80} = 0,65.$$

We have $80 \cdot \hat{p} \cdot (1-\hat{p}) = 18,2 > 10$. Hence an approx. confidence interval for p with level 95% is given by ($\alpha = 5\% = 0,05$)

$$I_p = \left[\hat{p} \mp \lambda_{\alpha/2} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right]$$

$n = 80$;
$\lambda_{0.025} = 1.96$

$$\overset{\downarrow}{=} \left[0,65 \mp 1.96 \cdot \sqrt{\frac{0.65 \cdot 0.35}{80}} \right]$$

$$= \left[0,545 \ ; \ 0,755 \right] //$$

b) We assume that the z_i are observations of a $\mathcal{N}(\Delta, \sigma^2)$ distribution with unknown Δ and σ^2.

(Paired observations / Stickprov i par !).

Then a 95% confidence interval for Δ is

$$I_\Delta = \left[\bar{z} \mp t_{\alpha/2}(n-1) \cdot \frac{s_z}{\sqrt{n}} \right]$$

$n=80$;

$$t_{0.025}(79) \approx 1.99 = \left[0.1 \mp 1.99 \cdot \frac{0.52}{\sqrt{80}} \right]$$

$$= \left[-0{,}0157 \,;\, 0{,}2157 \right]$$