11. cvičení z STP

29. dubna 2021

Pro náhodnou veličinu X a pravděpodobnost $\alpha \in (0,1)$ často potřebujeme najít $t \in \mathbb{R}$, že $P(X \le t) = \alpha$, tj. $F_X(t) = \alpha$.

Definice: Mějme náhodnou veličinu X se spojitou distribuční funkcí F_X, která je ostře rostoucí na otevřeném intervalu $I \subseteq \mathbb{R}$ takovém, že $F_X(I) = (0,1)$ (tj. F_X je bijekcí intervalu I a intervalu $(0,1)$). Pak existuje inverzní funkce $q_X := (F_X)^{-1} : (0,1) \to I$ a nazývá se *kvantilová funkce veličiny X*. V tom případě pak pro všechna $\alpha \in (0,1)$ a $t \in I$ platí:

$$P(X \le t) = \alpha \;\Leftrightarrow\; q_X(\alpha) = t$$

Odsud pak ihned plyne, např. že

- $P\Big(X \le q_X(\alpha)\Big) = \alpha$
- $P\Big(q_X(\alpha) \le X\Big) = 1 - \alpha \qquad \text{a} \qquad P\Big(q_X(1-\alpha) \le X\Big) = \alpha$
- $P\Big(q_X(\frac{\alpha}{2}) \le X \le q_X(1-\frac{\alpha}{2})\Big) = 1 - \alpha \qquad \text{a přitom je} \qquad P\Big(X < q_X(\frac{\alpha}{2})\Big) = \frac{\alpha}{2} = P\Big(q_X(1-\frac{\alpha}{2}) < X\Big)$

Pokud navíc $I = \mathbb{R}$ a X má hustotu, která je jako funkce sudá, pak platí:

- $F_X(t) + F_X(-t) = 1$ pro každé $t \in \mathbb{R}$,
- $q_X(\alpha) = -q_X(1-\alpha)$ pro každé $\alpha \in (0,1)$.

Pro hodnotu $q_X(\alpha)$ budeme ve speciálních případech používat toto značení:

- u_α pro $X \sim N(0,1)$,
- $t_{\alpha;\,k}$ pro X s t-rozdělením s k stupni volnosti,
- $\chi^2_{\alpha;\,k}$ pro X s χ^2-rozdělením s k stupni volnosti.

Poznámka: Na následujícím příkladu si ukážeme, v čem spočívá hledání intervalu spolehlivosti pro nějaký parametr. Odvodíme si oboustranný symetrický interval o spolehlivosti $1-\alpha$ pro střední hodnotu μ normálního rozdělení $N(\mu, \sigma^2)$ při *neznámém* rozptylu σ^2 (při známém rozptylu by výsledek vypadal jinak a jednodušeji):

Mějme realizaci náhodného výběru $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ o rozsahu n (pro nezávislé náhodné veličiny $X_i \;\sim\; N(\mu, \sigma^2)$). Hledáme teď nějaké funkce $h_1, h_2 : \mathbb{R}^n \to \mathbb{R}$ (nezávislé na volbě μ i σ) takové, že

$$P\Big(h_1(X_1, \ldots, X_n) \;\le\; \mu \;\le\; h_2(X_1, \ldots, X_n)\Big) = 1 - \alpha$$

(to je ta oboustrannost a $1-\alpha$ spolehlivost) a současně chceme, aby pro zbylé případy ještě platilo, že

$$P\Big(\mu \;<\; h_1(X_1, \ldots, X_n)\Big) = \frac{\alpha}{2} = P\Big(h_2(X_1, \ldots, X_n) \;<\; \mu\Big)$$

(to je ta symetričnost - tj. symetričnost nikoliv ve "vzdálenosti", ale v pravděpodobnosti).

Při dané realizaci $\boldsymbol{x} = (x_1, \ldots, x_n)$ pak jako hledaný interval spolehlivosti $1-\alpha$ pro μ chápeme (číselný) interval tvaru:

$$\Big\langle h_1(x_1, \ldots, x_n),\; h_2(x_1, \ldots, x_n)\Big\rangle \; (\subseteq \mathbb{R})$$

Je ještě dobré poznamenat, že

- *pro parametr μ žádné rozdělení pravděpodobnosti nemáme!*
- *daný interval spolehlivosti pro μ vzniká čistě na základě naměřených hodnot $\boldsymbol{x} = (x_1, \ldots, x_n)$ a také se společně s nimi MĚNÍ! Jeho smysl je ten, že skutečná hodnota $\mu = E(X)$ (která se NEMĚNÍ!) bude obsažena v těchto (obecně proměnných intervalech) s pravděpodobností $1 - \alpha$.*

  Ovšem problémem zůstává, že při neznalosti skutečné hodnoty μ nejsme schopni zjistit, které konkrétní naměřené intervaly μ obsahují a které naopak ne. Víme jen, že těch druhých je jen 5%. V tom je rozdíl oproti např. střílení do terče, kdy před pokusem víme, že se trefíme s pravděpodobností q a po uskutečněném pokusu umíme zjistit, který z výsledků nastal, takže si tuto pravděpodobnost můžeme i ověřit.
- *Pro konkrétní pokus se pak na vyčíslený interval můžeme dívat i takto: Dejme tomu, že pro parametr μ a $1 - \alpha = 95\%$ nám vyjde (při daných měřeních) interval jako $\langle 59.93,\ 62.07\rangle$. Pak 95% vyjadřuje poměr, se kterým se budeme ochotni vsadit, že skutečná hodnota μ (kterou může mít třeba někdo někde přesně zjištěnou), bude obsažena v intervalu $\langle 59.93,\ 62.07\rangle$.*

A teď pro náš konkrétní případ: Vezmeme si vhodnou veličinu, jejíž rozdělení známe. V našem případě veličinu

$$T = \frac{\overline{\boldsymbol{X}} - \mu}{S_{\boldsymbol{X}}}\sqrt{n}\,,$$

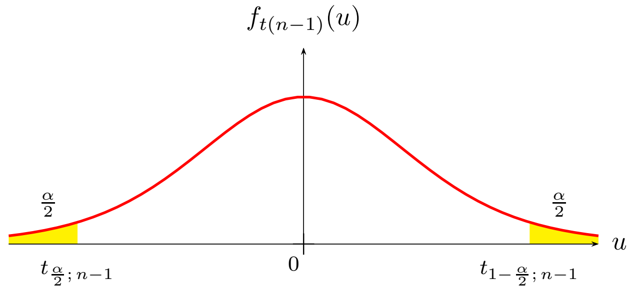
kde

$$\overline{\boldsymbol{X}} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

a

$$(S_{\boldsymbol{X}})^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n}(X_i - \overline{\boldsymbol{X}})^2\right) = \frac{1}{n-1}\left(\sum_{i=1}^{n} X_i^2 - n\cdot(\overline{\boldsymbol{X}})^2\right)\ .$$

Veličina T má tzv. t-rozdělení, tj. Studentovo rozdělení, s $n-1$ stupni volnosti. Hustota $f_{t(n-1)}$ veličiny T je:



Pak pro kvantily platí, že

$$P\Big(\underbrace{t_{\frac{\alpha}{2};n-1}}_{-t_{1-\frac{\alpha}{2};n-1}} \ \le\ T\ \le\ t_{1-\frac{\alpha}{2};n-1}\Big) = 1 - \alpha$$

a

$$P\Big(\ T\ < -t_{1-\frac{\alpha}{2};n-1}\Big) = \tfrac{\alpha}{2} = P\Big(t_{1-\frac{\alpha}{2};n-1}\ <\ T\ \Big)\ .$$

Výrazy uvnitř pravděpodobnosti si teď jen přepíšeme a budeme mít hledané funkce h_1 a h_2:

$$-t_{1-\frac{\alpha}{2};n-1}\ \le\ \frac{\overline{\boldsymbol{X}} - \mu}{S_{\boldsymbol{X}}}\sqrt{n}\ \le\ t_{1-\frac{\alpha}{2};n-1}$$

$$\underbrace{\overline{\boldsymbol{X}} - \frac{S_{\boldsymbol{X}}}{\sqrt{n}}\cdot t_{1-\frac{\alpha}{2};n-1}}_{h_1(X_1,\ldots,X_n)}\ \le\ \mu\ \le\ \underbrace{\overline{\boldsymbol{X}} + \frac{S_{\boldsymbol{X}}}{\sqrt{n}}\cdot t_{1-\frac{\alpha}{2};n-1}}_{h_2(X_1,\ldots,X_n)}$$

Po dosazení konkrétní realizace $\boldsymbol{x}$ vektoru $\boldsymbol{X}$ pak dostaneme výše uvedený interval spolehlivosti pro μ ve tvaru

$$\langle \mu_L,\ \mu_U \rangle = \left\langle \overline{\boldsymbol{x}} - \frac{s_{\boldsymbol{x}}}{\sqrt{n}}\, t_{1-\frac{\alpha}{2};\, n-1}, \quad \overline{\boldsymbol{x}} + \frac{s_{\boldsymbol{x}}}{\sqrt{n}}\, t_{1-\frac{\alpha}{2};\, n-1} \right\rangle$$

Pro názornost uveďme funkce h_1 a h_2 více rozepsané:

$$h_1(z_1,\ldots,z_n) := \frac{1}{n}\left(\sum_{i=1}^{n} z_i\right) - \frac{t_{1-\frac{\alpha}{2};\,n-1}}{\sqrt{n}} \cdot \frac{1}{n-1}\left(\sum_{i=1}^{n} z_i^2 - n\cdot\left(\frac{1}{n}\sum_{i=1}^{n} z_i\right)^2\right)$$

$$h_2(z_1,\ldots,z_n) := \frac{1}{n}\left(\sum_{i=1}^{n} z_i\right) + \frac{t_{1-\frac{\alpha}{2};\,n-1}}{\sqrt{n}} \cdot \frac{1}{n-1}\left(\sum_{i=1}^{n} z_i^2 - n\cdot\left(\frac{1}{n}\sum_{i=1}^{n} z_i\right)^2\right)$$

Příklad 11.1 (oboustranný intervalový odhad pro střední hodnotu)

Opakovaná měření stejné koncentrace látky (tj. náhodné veličiny X) vedla k následujícím výsledkům:

$$\boldsymbol{x} = (0.2,\ 0.23,\ 0.21,\ 0.16,\ 0.18,\ 0.19,\ 0.14,\ 0.18,\ 0.21)\ .$$

Najděte oboustranný symetrický 90% interval spolehlivosti pro střední hodnotu μ.

Řešení:
Pro veličinu X budeme předpokládat normální rozdělení $N(\mu, \sigma^2)$ (to jednak přibližně platí a především pouze za tohoto předpokladu můžeme používat známé vzorce). Dále budeme předpokládat, že měření byla nezávislá.

Intervalový odhad střední hodnoty μ (pro spolehlivost $0.9 = 1-\alpha$) je:

$$\langle \mu_L,\ \mu_U \rangle = \left\langle \overline{\boldsymbol{x}} - \frac{s_{\boldsymbol{x}}}{\sqrt{n}}\, t_{1-\frac{\alpha}{2};\, n-1}, \quad \overline{\boldsymbol{x}} + \frac{s_{\boldsymbol{x}}}{\sqrt{n}}\, t_{1-\frac{\alpha}{2};\, n-1} \right\rangle .$$

Pro jeho vyčíslení potřebujeme znát realizaci výběrového průměru $\overline{\boldsymbol{x}}$ a výběrového rozptylu $s^2_{\boldsymbol{x}}$ z realizace $\boldsymbol{x} = (x_1,\ldots,x_n)$ rozsahu $n = 9$:

- realizace výběrového průměru:

$$\overline{\boldsymbol{x}} = \tfrac{1}{n}\sum_{i=1}^{n} x_i = \frac{0.2 + 0.23 + \cdots + 0.21}{9} = \frac{1.7}{9} \doteq \mathbf{0.189}$$

- realizace výběrového rozptylu

$$\begin{aligned} s^2_{\boldsymbol{x}} &= \tfrac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{\boldsymbol{x}})^2 = \tfrac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2\right) - \tfrac{n}{n-1}\overline{\boldsymbol{x}}^2 = \\ &= \tfrac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2\right) - \tfrac{1}{n(n-1)}\left(\sum_{i=1}^{n} x_i\right)^2 = \\ &= \underbrace{\frac{0.2^2 + 0.23^2 + \cdots + 0.21^2}{8}}_{=\frac{2{,}9448}{8}=0.3681} - \frac{1.7^2}{9\cdot 8} \doteq \mathbf{7.6\cdot 10^{-4}}\ , \end{aligned}$$

- realizace směrodatné odchylky
$$s_{\boldsymbol{x}} = \sqrt{s^2_{\boldsymbol{x}}} \doteq \mathbf{2.76 \cdot 10^{-2}}$$

Intervalový odhad střední hodnoty μ (pro spolehlivost $0.9 = 1 - \alpha$) nyní je:

$$\begin{aligned}\langle \mu_L,\ \mu_U \rangle &= \left\langle \overline{\boldsymbol{x}} - \frac{s_{\boldsymbol{x}}}{\sqrt{n}}\, t_{1-\frac{\alpha}{2};\, n-1},\quad \overline{\boldsymbol{x}} + \frac{s_{\boldsymbol{x}}}{\sqrt{n}}\, t_{1-\frac{\alpha}{2};\, n-1} \right\rangle \doteq \\ &\doteq \left\langle 0.189 - \frac{2.76 \cdot 10^{-2}}{3}\, \underbrace{t_{0.95;\, 8}}_{1.86},\quad 0.189 + \frac{2.76 \cdot 10^{-2}}{3}\, t_{0.95;\, 8} \right\rangle \doteq \\ &\doteq \langle \mathbf{0.172},\ \mathbf{0.206} \rangle \ .\end{aligned}$$

Příklad 11.2 (jednostranné intervalové odhady pro střední hodnotu)
V terénu jsme naměřeli tyto výšky rostlin daného druhu (v centimetrech)

$$(75,\ 85,\ 58,\ 72,\ 70,\ 75)\ .$$

Předpokládejme, že výška rostliny X má normálního rozdělení $\mathrm{N}(\mu,\, \sigma^2)$. *Stanovte horní a dolní* 95% *interval spolehlivosti pro střední hodnotu μ.*

Řešení:
K určení intervalového odhadu opět použijeme statistiku

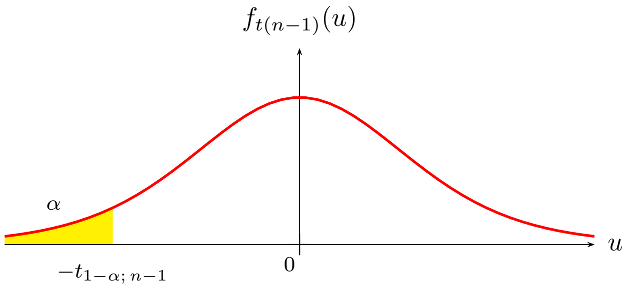
$$T = \frac{\overline{\boldsymbol{X}} - \mu}{S_{\boldsymbol{X}}} \sqrt{n}$$

která má Studentovo rozdělení $t(n-1)$, kde $n = 6$ je rozsah souboru. Poznamenejme, že zatímco centrální limitní větu používáme pro velká n, protože obvykle pro X máme nějaké "obecné" rozdělení, tak v případě, kdy X má *právě* normální rozdělení, známe rozdělení veličiny T také přesně a to pro jakákoliv n (tj. i malá).

- Horní interval spolehlivosti pro μ (tj. μ bude omezené *seshora*) dostaneme ze vztahu

$$P\Big(\underbrace{t_{\alpha;\, n-1}}_{-t_{1-\alpha;\, n-1}} \ \leq\ T \Big) = 1 - \alpha$$

který vyjadřuje dolní $1 - \alpha = 95\%$ intervalový odhad pro veličinu T (viz obrázek):



Případ

$$-t_{1-\alpha;\, n-1} \ \leq\ \frac{\overline{\boldsymbol{X}} - \mu}{S_{\boldsymbol{X}}} \sqrt{n}\ .$$

tedy nastává s pravděpodobností $1-\alpha = 95\%$. Po úpravě dostaneme pro horní interval spolehlivosti parametru μ tvar:

$$\mu \;\le\; \overline{\boldsymbol{X}} + \frac{S_{\boldsymbol{X}}}{\sqrt{n}}\, t_{0.95;\,5}\ .$$

Pro jeho vyčíslení potřebujeme znát realizaci výběrového průměru $\overline{\boldsymbol{x}}$ a výběrového rozptylu $s^2_{\boldsymbol{x}}$

$$\overline{\boldsymbol{x}} = \frac{1}{6}\sum_{i=1}^{6} x_i = \frac{435}{6} = \mathbf{72.5}$$

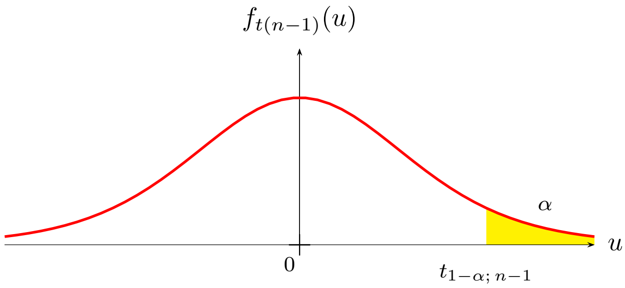
$$s^2_{\boldsymbol{x}} = \frac{1}{5}\left(\sum_{i=1}^{6} x_i^2 - 6\cdot(\overline{\boldsymbol{x}})^2\right) = \frac{385.5}{5} \doteq \mathbf{77.1}\,, \qquad s_{\boldsymbol{x}} \doteq \mathbf{8.781}\,.$$

Z tabulek kvantilů Studentova rozdělení dostaneme $t_{0.95;\,5} \doteq 2.02$, a hledaný interval je tedy

$$(-\infty,\ \mu_H\rangle = \left(-\infty, \quad \overline{\boldsymbol{x}} + \frac{s_{\boldsymbol{x}}}{\sqrt{n}}\cdot t_{0.95;\,5}\right\rangle = \left(-\infty, \quad 72.5 + \frac{8.781}{\sqrt{6}}\cdot 2.02\right\rangle =$$
$$= (-\infty,\ \mathbf{79.74}\rangle\ .$$

• Podobně dostaneme dolní interval spolehlivosti pro μ (tj. μ bude omezené *zezdola*) ze vztahu pro horní $1-\alpha = 95\%$ intervalový odhad veličiny T (viz obrázek)

$$P\Big(T \;\le\; t_{1-\alpha;\,n-1}\Big) = 1-\alpha$$



tedy

$$\frac{\overline{\boldsymbol{X}} - \mu}{S_{\boldsymbol{X}}}\sqrt{n} \;\le\; t_{1-\alpha;\,n-1}.$$

a po úpravě

$$\overline{\boldsymbol{X}} - \frac{S_{\boldsymbol{X}}}{\sqrt{n}}\, t_{0.95;\,5} \;\le\; \mu$$

a dosazení máme

$$\langle\mu_D,\ \infty) = \left\langle\overline{\boldsymbol{x}} - \frac{s_{\boldsymbol{x}}}{\sqrt{n}}\cdot t_{0.95;\,5}\ , \quad \infty\right) = \left\langle 72.5 - \frac{8.781}{\sqrt{6}}\cdot 2.02\ , \quad \infty\right) =$$
$$= \langle\mathbf{65.26},\ \infty)\ .$$

K testování hypotéz viz “Poznámky”. Pozor, kritéria pro zamítnutí nulové hypotézy a jejich ekvivalentní tvary používají často opakované negace různých podmínek (což někdy činí problémy s tím se v dané situaci zorientovat).

Podstatné je také to, že statistiky (tj. speciálně volené náhodné veličiny), které používáme při testech daného parametru, jsou sestaveny tak, aby jejich rozdělení na tomto parametru nezáviselo. (Je to určitý typ "normování").

Příklad 11.3 (test střední hodnoty normálního rozdělení při **známém** rozptylu)

Teploměrem, o jehož chybě předpokládáme, že má normální rozdělení s nulovou střední hodnotou a směrodatnou odchylkou $\sigma = 3°$, *jsme provedli* 30 *měření teploty roztoku. Průměrný výsledek byl* $61°$. *Otestujte na hladině významnosti* 5 %, *zda teplota roztoku je* $60°$.

Řešení:

Předpokládejme, že teplota roztoku je konstantní (ale neznámá) hodnota μ. Chyba měření teploty je veličina $Y \sim N(0, \sigma^2)$, kde $\sigma = 3°$. Naše veličina

$$X = \text{"naměřená teplota"}$$

(v jednotkách °) je tedy tvaru $X = \mu + Y$ a má tudíž rozdělení $N(\mu, \sigma^2)$.

Podle zadání máme otestovat hypotézu o střední hodnotě

$$\mathbf{H}_0 : \mu = \mu_0 \ (= 60°)$$

proti alternativní hypotéze:

$$\mathbf{H}_1 : \mu \neq \mu_0 \ (= 60°) \ .$$

Hodnotu rozptylu σ^2 známe. Takže použijeme test střední hodnoty se známým rozptylem (protože tím se dozvíme víc než kdybychom uvažovali neznámý rozptyl a navíc ani neznáme hodnotu výběrové směrodatné odchylky).

Pomocí testovací statistiky:

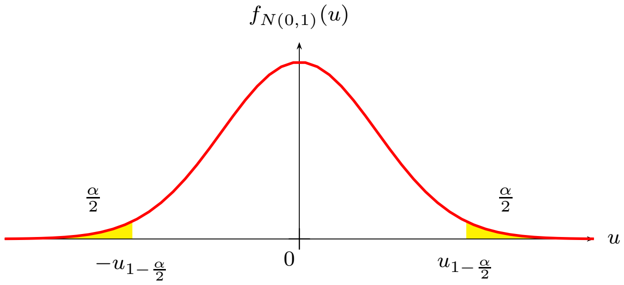
Realizace testovací statistiky $T = \frac{\overline{X} - \mu_0}{\sigma} \sqrt{n}$ pro $n = 30$ je

$$t = \frac{\overline{x} - \mu_0}{\sigma} \sqrt{n} = \frac{61 - 60}{3} \sqrt{30} \doteq 1.83$$

Zamítací kritérium zde máme tvaru

$$\text{zamítáme } \mathbf{H}_0 \text{ (na hladině } \alpha) \quad \Leftrightarrow \quad |t| > \underbrace{u_{1-\frac{\alpha}{2}}}_{\Phi^{-1}\left(1-\frac{\alpha}{2}\right)} \quad .$$

Proč má zamítací kritérium uvedený tvar: Za předpokladu platnosti nulové hypotézy, tj. pokud $E(X) = \mu_0$, bude mít statistika T normální rozdělení $N(0, 1)$. Očekávané hodnoty takovéto statistiky T by se měly pohybovat kolem nuly. Pokud se příliš odchýlí, bude to důvod k zamítnutí nulové hypotézy. Nebudeme přitom preferovat vychýlení na žádnou ze stran - tj. chybu 1. druhu s pravděpodobností α rozdělíme na poloviny $\frac{\alpha}{2}$ na obě strany.



Pak máme

$$P_{(H_0 \text{ platí})}\Big(\text{nastává chyba 1. druhu}\Big) = P_{(H_0 \text{ platí})}\Big(\text{zamítáme } \mathbf{H}_0\Big) =$$
$$= P_{(H_0 \text{ platí})}\Big(\ |T| > u_{1-\frac{\alpha}{2}}\ \Big) = \frac{\alpha}{2} + \frac{\alpha}{2} = \alpha$$

Hodnotu kvantilu pro $\alpha = 0.05$ (tj. pro $1 - \frac{\alpha}{2} = 0.975$) určíme z tabulek jako

$$u_{0.975} = \Phi^{-1}(0.975) \doteq 1.96\ .$$

Protože zamítací kritérium NENÍ splněno, tj. máme:

$$1.96 = u_{0.975} \not< |t| = 1.83$$

tak nulovou hypotézu $\mathbf{H}_0$ NEzamítáme.

Pomocí intervalu spolehlivosti:
Kritérium pro zamítnutí $\mathbf{H}_0$ na hladině α

$$u_{1-\frac{\alpha}{2}} < |t| \quad \left(= \frac{|\overline{\boldsymbol{x}} - \mu_0|}{\sigma}\sqrt{n}\ \right)$$

se dá ekvivalentně přepsat jako

$$|\overline{\boldsymbol{x}} - \mu_0| > \frac{\sigma}{\sqrt{n}} \cdot u_{1-\frac{\alpha}{2}}$$

neboli

$$\mu_0 \notin \left\langle\ \overline{\boldsymbol{x}} - \frac{\sigma}{\sqrt{n}} \cdot u_{1-\frac{\alpha}{2}} \quad , \quad \overline{\boldsymbol{x}} + \frac{\sigma}{\sqrt{n}} \cdot u_{1-\frac{\alpha}{2}}\ \right\rangle \ =: \langle \mu_L,\ \mu_U \rangle$$

což je tvar zamítacího kritéria s použitím intervalu spolehlivosti.

Při vyčíslení pro $\alpha = 5\%$ tedy dostaneme

$$\langle \mu_L,\ \mu_U \rangle = \left\langle\ 61 - \underbrace{\frac{3}{\sqrt{30}} \cdot 1.96}_{\doteq 1.07} \quad , \quad 61 + \frac{3}{\sqrt{30}} \cdot 1.96\ \right\rangle = \langle 59.93,\ 62.07 \rangle$$

Protože máme

$$\mu_0 = 60 \in \langle 59.93,\ 62.07 \rangle = \langle \mu_L,\ \mu_U \rangle$$

(tj. kritérium pro zamítnutí není splněno) hypotézu $\mathbf{H}_0$ **NEZAMÍTÁME** na hladině 5%.
(Výsledek musel samozřejmě dopadnout stejně jako při testovací statistice, protože je to ekvivalentní princip.)

Intuitivně to můžeme chápat takto:

(1) μ se má nacházet s 95% pravděpodobností v intervalu spolehlivosti;

(2) my předpokládáme, že $\mu = \mu_0$;

(3) zjistili jsme, že μ_0 v intervalu je.

Tedy μ se skutečně nachází v intervalu spolehlivosti a my tak nemáme žádný důvod k tomu hypotézu zamítnout.

(A kdybychom to naopak nezjistili, pak bychom hypotézu zamítli, ale spletli bychom se přitom jen s 5% pravděpodobností.)

Příklad 11.4 (test střední hodnoty normálního rozdělení při **neznámém** rozptylu)

Výrobce tvrdí, že spotřeba jím vyráběného automobilu je $\mu_0 = 8\ \ell/100\,\mathrm{km}$. *Průměrná spotřeba u* $n = 49$ *uživatelů ale byla* $\overline{\boldsymbol{x}} = 8.4\ \ell/100\,\mathrm{km}$. *Naměřen byl dále výběrový rozptyl* $s_{\boldsymbol{x}}^2 = 2.56\ (\ell/100\,\mathrm{km})^2$.

(a) Testujte na hladině 5%, *zda měl výrobce pravdu (tj. zda spotřeba je rovna* $8\ \ell/100\,\mathrm{km}$*).*

(b) Testujte na hladině 5%, *zda je spotřeba* ***nejvýše*** *rovna* $8\ \ell/100\,\mathrm{km}$.

Jak dopadne testovaní těchto hypotéz na hladině 1%*?*

Řešení:
U veličiny

$$X = \text{“}\textit{spotřeba automobilu}\text{”}$$

budeme předpokládat normální rozdělení $N(\mu, \sigma^2)$. Jednotlivá měření X_i, pro $i = 1, \ldots, 49$, jsou nezávislá. Oba parametry jsou neznámé a my chceme testovat střední hodnotu μ.

(a) Podle zadání máme na hladině $\alpha = 5\%$ (příp. 1%) otestovat hypotézu o střední hodnotě

$\mathbf{H}_0 : \mu = \mu_0 (= 8)$

proti alternativní hypotéze:

$\mathbf{H}_1 : \mu \neq \mu_0 (= 8)$.

Pomocí testovací statistiky:

Protože hodnotu rozptylu neznáme, provedeme t-test s testovací veličinou (tzv. *statistikou*):

$$T = \frac{\overline{\boldsymbol{X}} - \mu_0}{S_{\boldsymbol{X}}} \sqrt{n}$$

kde

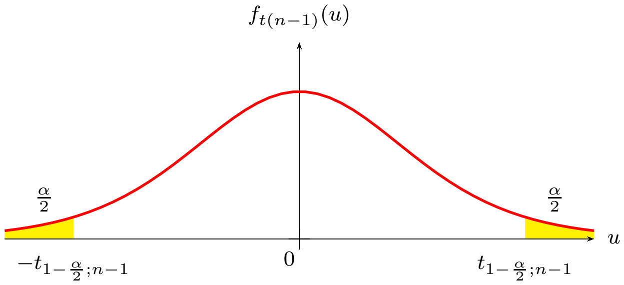
- veličina $\overline{\boldsymbol{X}} = \frac{1}{n} \sum_{i=1}^{n} X_i$ je *výběrový průměr* a
- veličina $S_{\boldsymbol{X}}^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{\boldsymbol{X}})^2$ je *výběrový rozptyl.*

Kritérium pro **ZAMÍTNUTÍ H**$_0$ (na hladině α) je tvaru

$$\text{zamítáme } \mathbf{H}_0 \text{ (na hladině } \alpha) \quad \Leftrightarrow \quad |t| > t_{1-\frac{\alpha}{2};n-1} \ .$$

kde t je hodnota T na základě naměřených dat.

Proč má zamítací kritérium uvedený tvar: Za předpokladu platnosti nulové hypotézy, tj. pokud $E(X) = \mu_0$, bude mít statistika T tzv. Studentovo t-rozdělení s $n-1$ stupni volnosti a hustotou $f_{t(n-1)}$ (která má podobný, ale ne stejný, průběh jako u $N(0,1)$):



Očekávané hodnoty takovéto statistiky T by se měly pohybovat blízko nuly. Pokud se příliš odchýlí, bude to důvod k zamítnutí nulové hypotézy. Nebudeme přitom preferovat vychýlení na žádnou ze stran - tj. chybu 1. druhu s pravděpodobností α rozdělíme na poloviny $\frac{\alpha}{2}$ na obě strany. Pak máme

$$P_{(H_0 \text{ platí})}\Big(\text{nastává chyba 1. druhu}\Big) = P_{(H_0 \text{ platí})}\Big(\text{zamítáme } \mathbf{H}_0\Big) =$$
$$= P_{(H_0 \text{ platí})}\Big(\ |T| > t_{1-\frac{\alpha}{2};n-1}\ \Big) = \frac{\alpha}{2} + \frac{\alpha}{2} = \alpha$$

Teď už tedy dosadíme konkrétní naměřené hodnoty (které pro jednotlivé veličiny značíme pro odlišení malými písmeny, tj. $\overline{\boldsymbol{x}}$, $s_{\boldsymbol{x}}^2$ a t). Realizace testovací statistiky je

$$t = \frac{\overline{\boldsymbol{x}} - \mu_0}{s_{\boldsymbol{x}}}\sqrt{n} = \frac{8.4 - 8}{\sqrt{2.56}}\sqrt{49} = \frac{0.4}{1.6} \cdot 7 = 1.75\ .$$

Protože pro $\alpha = 0.05$ je

$$|t| = 1.75 \not> 2.011 \doteq t_{0.975;48} = t_{1-\frac{\alpha}{2};n-1}\ ,$$

nulovou hypotézu **NEZAMÍTÁME** na hladině 5%.

Protože při snížení hladiny se zmenšuje i kritický obor W (je to vidět i na obrázku, kde žlutá plocha bude menší), tak na hladině 1% hypotézu $\mathbf{H}_0$ také **NEZAMÍTÁME**.
(Pro úplnost si ale stejně ještě vyjádříme příslušnou podmínku: $|t| = 1.75 \not> 2.682 \doteq t_{0.995;48}$.)

Obecněji tedy:

snižujeme hladinu chyby 1. druhu (tj. chceme si být více jistí) $\Rightarrow$ musíme tolerovat více "prohřešků" $\Rightarrow$ častěji nezamítáme

Pomocí intervalu spolehlivosti:
Kritérium pro zamítnutí $\mathbf{H}_0$ na hladině α

$$|t| > t_{1-\frac{\alpha}{2};n-1}$$

se dá ekvivalentně přepsat (při vyjádření $t = \frac{\overline{\boldsymbol{x}}-\mu_0}{s_{\boldsymbol{x}}}\sqrt{n}$) jako

$$\mu_0 \notin \Big\langle\ \overline{\boldsymbol{x}} - \frac{s_{\boldsymbol{x}}}{\sqrt{n}} \cdot t_{1-\frac{\alpha}{2};n-1}\quad ,\quad \overline{\boldsymbol{x}} + \frac{s_{\boldsymbol{x}}}{\sqrt{n}} \cdot t_{1-\frac{\alpha}{2};n-1}\ \Big\rangle =: \langle \mu_L,\ \mu_U \rangle$$

což je hledaný interval spolehlivosti.

Při vyčíslení pro $\alpha = 5\%$, tj. $t_{1-\frac{\alpha}{2};n-1} = t_{0.975;\,48} \doteq 2.011$, tedy dostaneme

$$\langle \mu_L,\ \mu_U \rangle = \Big\langle\ 8.4 - \frac{1.6}{\sqrt{49}} \cdot 2.011\quad ,\quad 8.4 + \frac{1.6}{\sqrt{49}} \cdot 2.011\ \Big\rangle = \langle 7.94,\ 8.86 \rangle$$

Protože máme $\mu_0 = 8 \in \langle 7.94,\ 8.86 \rangle = \langle \mu_L,\ \mu_U \rangle$, hypotézu $\mathbf{H}_0$ **NEZAMÍTÁME** na hladině 5%.
(Výsledek musel samozřejmě dopadnout stejně jako při testovací statistice, protože je to ekvivalentní princip.)

- -

(b) V tomto případě budeme na hladině $\alpha = 5\%$ (příp. 1%) testovat hypotézu o střední hodnotě

$\widetilde{\mathbf{H}}_0 : \mu \leq \mu_0 (= 8)$

proti alternativní hypotéze:

$$\widetilde{\mathbf{H}}_1 : \mu > \mu_0 (= 8) \ .$$

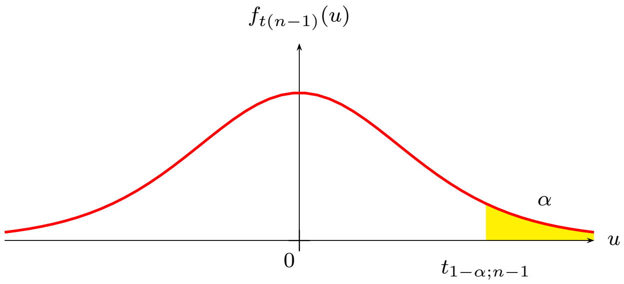
Pomocí testovací statistiky:

Statistika T bude mít stejný tvar jako v předešlém případě (a). Kritérium pro **ZAMÍTNUTÍ $\widetilde{\mathbf{H}}_0$** (na hladině α) bude ale teď jiné, a sice

$$\text{zamítáme } \widetilde{\mathbf{H}}_0 \text{ (na hladině } \alpha) \quad \Leftrightarrow \quad t > t_{1-\alpha;n-1} \ .$$

Proč má zamítací kritérium uvedený tvar: Za předpokladu platnosti nulové hypotézy, tj. pokud $E(X) \leq \mu_0$, bude mít hustota pro statistiku T svůj vrchol v intervalu $(-\infty, 0\rangle$. Očekávané hodnoty takovéto statistiky T by se měly pohybovat spíše v záporných až nulových hodnotách. Pokud se příliš odchýlí do kladných hodnot, bude to důvod k zamítnutí nulové hypotézy. "Nejhorší" z tohoto hlediska je krajní případ $E(X) = \mu_0$, pro který má T opět Studentovo t-rozdělení s $n-1$ stupni volnosti (viz obrázek). (V ostatních případech $\mu < \mu_0$ uz ovšem statistika T Studentovo rozdělení *nemá*! Pro zájemce je více podrobností na konci tohoto dokumentu.)

Chyba 1. druhu s pravděpodobností α zde tedy bude soustředěna jen na jedné straně:



Podobně jako předtím máme

$$P_{(\tilde{H}_0 \text{ platí})}\Big(\text{nastává chyba 1. druhu}\Big) = P_{(\tilde{H}_0 \text{ platí})}\Big(\text{zamítáme } \widetilde{\mathbf{H}}_0\Big) =$$
$$= P_{(\tilde{H}_0 \text{ platí})}\Big(\ T > t_{1-\alpha;n-1}\ \Big) = \alpha$$

Hodnota statistiky T zůstane stejná jako předtím, tedy $t = 1.75$, a protože pro $\alpha = 0.05$ máme splněno zamítací kritérium

$$t = 1.75 > 1.677 \doteq t_{0.95;\,48} = t_{1-\alpha;\,n-1} \ ,$$

hypotézu $\widetilde{\mathbf{H}}_0$ **ZAMÍTNEME**.

(Pozor, jde o jednostranný test, takže kvantil je jiný! Veškerou chybu jsme spotřebovali jen na kladné hodnoty. A toto malé zvětšení, oproti oboustrannému testu, už stačilo na zamítnutí.)

Pro $\alpha = 1\%$ pak máme

$$t = 1.75 \not> 2.407 \doteq t_{0.99;\ 48} \ ,$$

takže při této hladině hypotézu $\widetilde{\mathbf{H}}_0$ naopak **NEZAMÍTNEME**.

Pomocí intervalového odhadu:

Kritérium pro zamítnutí $\widetilde{\mathbf{H}}_0$ na hladině α

$$t > t_{1-\alpha;\,n-1}$$

se dá ekvivalentně přepsat (opět při vyjádření $t = \frac{\overline{\boldsymbol{x}} - \mu_0}{s_{\boldsymbol{x}}} \sqrt{n}$) jako

$$\mu_0 \notin \Big\langle\ \overline{\boldsymbol{x}} - \frac{s_{\boldsymbol{x}}}{\sqrt{n}} \cdot t_{1-\alpha;\,n-1}\ ,\quad +\infty\ \Big) \quad =: \langle \mu_D,\ +\infty)$$

což je hledaný dolní interval spolehlivosti.

Při vyčíslení pro $\alpha = 5\%$, tj. $t_{1-\alpha;\,n-1} = t_{0.95;\,48} \doteq 1.677$, tedy dostaneme

$$\langle \mu_D,\ +\infty) = \Big\langle\ 8.4 - \frac{1.6}{\sqrt{49}} \cdot 1.677\ ,\ \ +\infty\Big) = \langle 8.017,\ +\infty)$$

Protože máme $\mu_0 = 8 \notin \langle 8.017,\ +\infty) = \langle \mu_D,\ +\infty)$, hypotézu $\widetilde{\mathbf{H}}_0$ **ZAMÍTÁME** na hladině 5%. (Výsledek opět dopadne stejně jako při testovací statistice, protože je to ekvivalentní princip.)

Tvar intervalu spolehlivosti si můžeme intuitivně zapamatovat takto:

Při pravdivosti $\widetilde{\mathbf{H}}_0$ je $\mu \leq \mu_0$. Protože $\langle \mu_D,\ \infty)$ představuje dolní interval spolehlivosti 95% pro μ, musí být s touto pravděpodobností v tomto intervalu i μ_0. Pokud není (což nastane jen s 5% pravděpodobností), je to důvod k zamítnutí.

Důležitá poznámka: Všimněme si, že jsme došli k těmto (zdánlivě protichůdným výsledkům): na hladině $\alpha = 5\%$ jsme

- hypotézu $\mu = \mu_0$ nezamítli
- hypotézu $\mu \leq \mu_0$ zamítli

přestože nezamítnutý případ je podpřípadem zamítnutého! To vypadá sice jako rozpor, ale ve skutečnosti v každém z případů testujeme hypotézy jiným způsobem. Jak už bylo napsáno výše, chyba se v případě oboustranného testu rozloží symetricky na obě strany, zatímco u jednostranného testu je nahromaděna jen na jednom konci.