13. cvičení z PRA

13. - 17. května 2024

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 $(0,1)$).

Pro pravděpodobnost $\alpha \in (0,1)$ budeme potřebovat najít $t \in \mathbb{R}$, že $P(X \le t) = \alpha$, tj. $F_X(t) = \alpha$.

Kvantilová funkce veličiny X je v tomto případě inverzní funkce $q_X := (F_X)^{-1} : (0,1) \to I$. Pro všechna $\alpha \in (0,1)$ a $t \in I$ pak platí:

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

Odsud pak ihned plyne, např. že

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

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 $\boldsymbol{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ 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

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

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

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

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

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

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

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

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

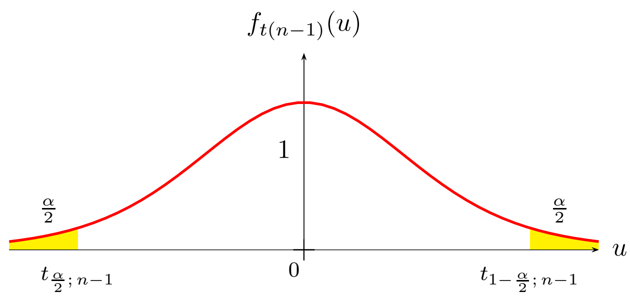
kde

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

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

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

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

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

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 (pokud možno) nezáviselo. (Je to určitý typ "normování").

Shrnutí:

(a) Intervalové odhady i testy hypotéz vždy děláme s co největší přesností, konkrétně:

Pokud můžeme předpokládat, že veličina X má normální rozdělení $N(\mu, \sigma^2)$, pak veličina $T = \frac{\overline{\boldsymbol{X}}-\mu}{S_{\boldsymbol{X}}}\sqrt{n}$ má t-rozdělení s $n-1$ stupni volnosti a tedy používáme při řešení úlohy kvantily t-rozdělení.

Pokud rozdělení X je jiné (např. alternativní, Poissonovo) používáme tomu odpovídající intervalové odhady a testování, případně přejdeme (zejména u neznámých rozdělení) k asymptotickým hodnotám kvantilů, tj. pak u veličiny T pracujeme s přibližně normálním rozdělením $N(0,1)$.

(b) Rozhodování o zamítání nulové hypotézy se dá popsat stručně asi takto:

Jestliže pozorujeme situaci, která by na základě našich předpokladů měla být výjimečná a málo pravděpodobná, pak bychom měli přehodnotit naše předpoklady (které nejspíše nebudou správně). A pokud se výjimečné situace dějí dokonce opakovaně, pak je důvod k přehodnocení o to silnější.

Toto pravidlo je vhodné uplatňovat i v jiných oblastech života.

(c) Při testování hypotéz se *vždy testuje nulová hypotéza*, která vyjadřuje předpoklad o daném rozdělení. Jako nulovou hypotézu si tedy volíme takovou situaci, kdy nám rozdělení dané veličiny dává nějaké vodítko jak, rozhodovat - např. *můžeme testovat nezávislost* veličin (neboť pak něco platí o určité jiné veličině, tzv. testovací statistice), ale *neumíme testovat závislost* těchto veličin (protože pak nemáme žádné podobné vodítko).

Příklad 13.1 (asymptotický test střední hodnoty)

U 64 *praktických lékařů byl naměřen výběrový průměr počtu pacientů za den* 23*, výběrový rozptyl pak byl roven* 36*. Rozdělení počtu pacientů není známé.*

(a) Sestrojte (asymptotický) oboustranný interval pro střední hodnotu počtu pacientů o spolehlivosti 95%*.*

(b) Otestujte na hladině 5%*, zda skutečná střední hodnota počtu pacientů za den může být považována za rovnu* 25*.*

Řešení:

Máme veličiny

$$X_i = \text{"}\textit{počet pacientů u } i\textit{-tého lékaře za den}\text{"}$$

pro $i = 1, \ldots, n$, kde $n = 64$, které budeme pokládat za nezávislé. Jejich rozdělení není známé.

(a) Asymptotický oboustranný interval pro střední hodnotu o spolehlivosti $1-\alpha$ se bude podobat tomu, který se odvozuje, pokud X_i mají normální rozdělení. Rozdíl bude jen v tom, že kvantil $t_{1-\frac{\alpha}{2};\, n-1}$ pro Studentovo rozdělení s $n-1$ stupni volnosti (který používáme, když X_i mají normální rozdělení) se nahradí asymptotickou hodnotou když $n \to \infty$, která odpovídá kvantilu $u_{1-\frac{\alpha}{2}} = \Phi^{-1}\left(1-\frac{\alpha}{2}\right)$ pro rozdělení $N(0,1)$.

Příslušný asymptotický interval o spolehlivosti $1-\alpha = 95\%$ tedy je:

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

kde $\overline{\boldsymbol{x}} = 23$ je výběrový průměr a $s_{\boldsymbol{x}}^2 = 36$ je výběrový rozptyl. Pro $\alpha = 5\%$ máme hodnotu kvantilu $u_{1-\frac{\alpha}{2}} = u_{0.975} = \Phi^{-1}(0.975) \doteq 1.96$. Po dosazení máme tedy

$$\langle \mu_L,\ \mu_U \rangle := \left\langle\ 23 - \frac{\sqrt{36}}{\sqrt{64}} \cdot 1.96 \quad , \quad 23 + \frac{\sqrt{36}}{\sqrt{64}} \cdot 1.96\ \right\rangle \doteq \langle\ 22.28,\ 23.72\ \rangle$$

Tento interval se pochopitelně MĚNÍ s každým měřením (protože je závislý na naměřených vstupech), a jeho smysl je ten, že skutečná hodnota $\mu = E(X)$ (která se NEMĚNÍ!) bude obsažena v tomto (obecně proměnném intervalu) s pravděpodobností 95%.

(b) Podle zadání máme na hladině $\alpha = 5\%$ otestovat hypotézu o střední hodnotě $\mu = E(X)$ tvaru

$\mathbf{H}_0 : \mu = \mu_0$

proti alternativní hypotéze:

$\mathbf{H}_A : \mu \neq \mu_0$.

kde $\mu_0 = 25$.

Pomocí intervalového odhadu:

Využijeme už spočítaného asymptotického oboustranného intervalu $\langle \mu_L,\ \mu_U \rangle$ pro střední hodnotu o spolehlivosti 95%. Podle toho, co jsme uvedli výše v části (a), je pravděpodobnost, že střední hodnota $\mu = E(X)$ bude obsažena v (proměnném) intervalu $\langle \mu_L,\ \mu_U \rangle$, rovna 95%. Tedy mimo tento interval se ocitne jen v 5% případů.

Jestliže předpokládáme, že $\mu = \mu_0$ (tj. hypotézu $\mathbf{H}_0$), bude kritérium pro její zamítnutí na hladině α přirozeně tvaru:

$$\text{zamítáme } \mathbf{H}_0 \text{ (na hladině } \alpha) \quad \Leftrightarrow \quad \mu_0 \notin \langle \mu_L,\ \mu_U \rangle\ .$$

A protože skutečně nakonec máme, že $\mu_0 = 25 \notin \langle\ 22.28,\ 23.72\ \rangle = \langle \mu_L,\ \mu_U \rangle$, tak hypotézu $\mathbf{H}_0$ **ZAMÍTÁME** na hladině 5%.

Pomocí testovací statistiky:

Podmínka pro zamítnutí $\mathbf{H}_0$ na hladině α se dá z formy pro interval spolehlivosti

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

ekvivalentně přepsat jako

$$|t| > u_{1-\frac{\alpha}{2}}$$

kde $t = \frac{\overline{\boldsymbol{x}} - \mu_0}{s_{\boldsymbol{x}}}\sqrt{n}$, což je (podobně jako v **Příkladu 11.4**) hodnota testovací veličiny (tzv. *statistiky*):

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

Protože však neznáme rozdělení veličin X_i, neznáme ani *přesné* rozdělení této statistiky T. To ale na druhou stranu nevadí, protože pro dost velká n nakonec bude mít veličina T přibližné rozdělení $N(0,1)$ (bez ohledu na počáteční rozdělení rozdělení veličin X_i). To je tedy důvod, proč se pak v zamítacím kritériu objevují kvantily pro norm. rozdělení. Co přesně znamená "dost velká n", závisí pochopitelně na tom, jak "divoké" je rozdělení veličin X_i.

Shrňme si to tedy tak, že kritérium pro zamítnutí $\mathbf{H}_0$ (na hladině α) je tvaru

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

Při konkrétním dosazení máme

$$t \;=\; \frac{\overline{\boldsymbol{x}} - \mu_0}{s_{\boldsymbol{x}}}\sqrt{n} \;=\; \frac{23-25}{\sqrt{36}}\sqrt{64} \;=\; -\frac{16}{3} \;\doteq\; -5.33$$

a tudíž

$$|t| \doteq |5.33| > 1.96 \doteq u_{0.975}$$

což znamená, že hypotézu $\mathbf{H}_0$ (opět) **ZAMÍTÁME** na hladině 5%.

(Výsledek musel samozřejmě dopadnout stejně jako pomoci intervalu spolehlivosti, protože je to ekvivalentní princip.)

Příklad 13.2 (test střední hodnoty normálního rozdělení při **neznámém** rozptylu)

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

Testujte na hladině 5%, *zda měl výrobce pravdu. 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.

Nabízí se nyní dvě možnosti, jak úlohu chápat:

(a) Budeme testovat hypotézu, zda spotřeba je (přesně) rovna $8\ \ell/100\,\text{km}$.

(b) Budene testovat hypotézu, zda spotřeba je **nejvýše** rovna $8\ \ell/100\,\text{km}$ (protože to by spíše spotřebitele zajímalo).

- -

(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

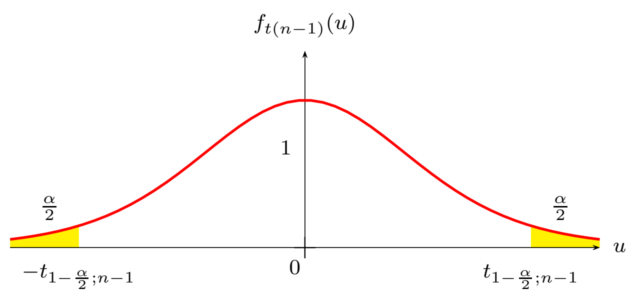
- veličina $\overline{\boldsymbol{X}} = \frac{1}{n} \sum_{i=1}^{n} X_i$ je *výběrový průměr* a
- veličina $S^2_{\boldsymbol{X}} = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{\boldsymbol{X}})^2$ je *výběrový rozptyl.*

Kritérium pro **ZAMÍTNUTÍ** $\mathbf{H}_0$ (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. $\overline{\boldsymbol{x}}$, $s^2_{\boldsymbol{x}}$ 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

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

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

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

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

$\widetilde{\mathbf{H}}_0 : \mu \le \mu_0$

proti alternativní hypotéze:

$\widetilde{\mathbf{H}}_1 : \mu > \mu_0$.

Tato situace se také často zapisuje v podobě nulové hypotézy $\widetilde{\mathbf{H}}_0 : \mu = \mu_0$ při stejné alternativní hypotéze $\widetilde{\mathbf{H}}_1 : \mu > \mu_0$. Postup je v obou těchto případech stejný i když právě tato shoda není ihned zřejmá a je potřeba ji dokázat. Volba hypotézy ve tvaru $\widetilde{\mathbf{H}}_0 : \mu \le \mu_0$ na druhou stranu více vysvětluje tvar zamítacího kritéria.

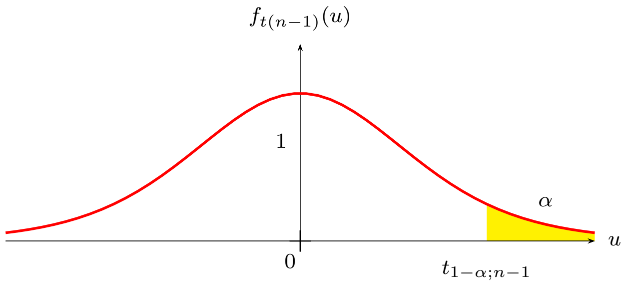
Pomocí testovací statistiky:
Statistika T bude mít stejný tvar jako v předešlém případě (a). Kritérium pro **ZAMÍTNUTÍ** $\widetilde{\mathbf{H}}_0$ (na hladině α) bude ale teď jiné, a sice

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

Proč má zamítací kritérium uvedený tvar: Za předpokladu platnosti nulové hypotézy, tj. pokud $E(X) \le \mu_0$, 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 $\mu < \mu_0$ už 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

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

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.

Příklad 13.3 (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

$$X = \text{“}\textit{naměřená teplota}\text{”}$$

(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:

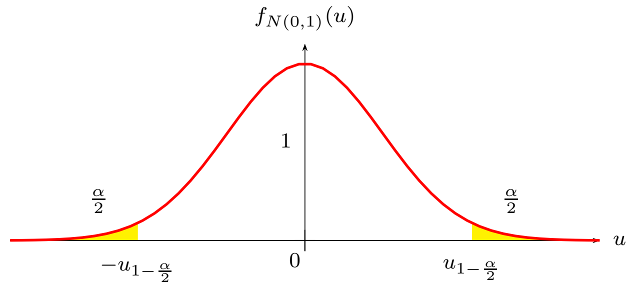
Testovací statistika $T = \frac{\overline{\boldsymbol{X}} - \mu_0}{\sigma}\sqrt{n}$ pro $n = 30$ má nyní **normální rozdělení** $N(0,1)$**, nikoliv** t**-rozdělení** (protože rozptyl je známý). Realizace T je

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

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:

$$1.96 = u_{0.975} \nless |t| = 1.83$$

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.)

Podrobnější zdůvodnění tvaru zamítacího kritéria pro test nulové hypotézy $\mu \leq \mu_0$ s neznámým rozptylem:

Nejdříve si pro zjednodušení uvědomme následující věc:
Jestliže máme dvě veličiny X, Y takové, že $X \leq Y$ (tj. $X(\omega) \leq Y(\omega)$ pro všechna $\omega \in \Omega$), pak platí, že

- $E(X) \leq E(Y)$ pokud střední hodnoty existují,
- $P(c < X) \leq P(c < Y)$ pro všechna $c \in \mathbb{R}$.

A nyní to zdůvodnění: Abychom více zdůraznili závislost veličiny $X \sim N(\mu, \sigma^2)$ na parametru μ, budeme ji vyznačovat jako X_μ a podobně to budeme psát u statistiky $T_\mu = \frac{\overline{\boldsymbol{X}}_\mu - \mu_0}{S_{\boldsymbol{X}_\mu}}\sqrt{n}$. Je nutné zdůraznit, že za předpokladu nulové hypotézy (tj. že $\mu \leq \mu_0$) statistika T_μ obecně **NEMÁ** Studentovo t-rozdělení (t-rozdělení se objeví právě jen pokud $\mu = \mu_0$).

Nyní předpokládejme platnost nulové hypotézy $\mathbf{H}_0: \ \mu \leq \mu_0$. Pak máme:

$$\mu \leq \mu_0 \quad \Rightarrow \quad T_\mu = \underbrace{\frac{\overline{\boldsymbol{X}}_\mu - \mu_0}{S_{\boldsymbol{X}_\mu}}\sqrt{n}}_{\text{složitější rozdělení}} \quad = \quad \frac{\overline{\boldsymbol{X}}_\mu - \mu}{S_{\boldsymbol{X}_\mu}}\sqrt{n} + \underbrace{\frac{\mu - \mu_0}{S_{\boldsymbol{X}_\mu}}\sqrt{n}}_{\leq 0} \quad \leq \quad \underbrace{\frac{\overline{\boldsymbol{X}}_\mu - \mu}{S_{\boldsymbol{X}_\mu}}\sqrt{n}}_{t_{(n-1)}-\text{rozdělení}} \ =: U_\mu$$

Tedy dostali jsme, že $T_\mu \leq U_\mu$ a U_μ má $t_{(n-1)}$-rozdělení s $n-1$ stupni volnosti. Z toho máme, že:

$$E(T_\mu) \leq E(U_\mu) = 0$$

Můžeme tak očekávat, že hodnoty statistiky T_μ budou především v intervalu $(-\infty, 0\rangle$. Jako obor $\widetilde{W} \subseteq \mathbb{R}$ takových hodnot veličiny T, kdy už budeme zamítat, si proto zvolíme

$$\widetilde{W} : (u_1, \infty) \ ,$$

kde požadujeme, aby $u_1 \in \mathbb{R}$ bylo nejmenší takové, aby chyba 1. druhu byla nejvýše α, tj.

$$(\forall \mu \leq \mu_0) \quad \alpha \geq P_{(\mathbf{H}_0 \text{ platí})}\Big(\text{nastává chyba 1. druhu}\Big) = P_{(\mathbf{H}_0 \text{ platí})}\Big(\text{zamítáme } \mathbf{H}_0\Big) =$$
$$= P_{(\mathbf{H}_0 \text{ platí})}\Big(\ T_\mu \in \widetilde{W} \ \Big) = P_{(\mathbf{H}_0 \text{ platí})}\Big(\ u_1 < T_\mu \ \Big)$$

A teď si jen určíme, kolik musí být u_1 a současně ukážeme, že největší možná chyba nastává pro případ $\mu = \mu_0$:

Z toho, že $T_\mu \leq U_\mu$ a veličiny U_μ a T_{μ_0} mají obě $t_{(n-1)}$-rozdělení s $n-1$ stupni volnosti ihned dostaneme, že

$$P_{(\mathbf{H}_0 \text{ platí})}\Big(u_1 < T_\mu\Big) \leq P_{(\mathbf{H}_0 \text{ platí})}\Big(u_1 < U_\mu\Big) = 1 - F_{t_{(n-1)}}(u_1) = P_{(\mathbf{H}_0 \text{ platí})}\Big(u_1 < T_{\mu_0}\Big)$$

Vidíme tedy, že $P_{(\mathbf{H}_0 \text{ platí})}\Big(u_1 < T_\mu\Big) \leq P_{(\mathbf{H}_0 \text{ platí})}\Big(u_1 < T_{\mu_0}\Big) \leq \alpha$ a hledané nejmenší u_1 tak musí splňovat, že

$$P_{(\mathbf{H}_0 \text{ platí})}\Big(u_1 < T_{\mu_0}\Big) = \alpha$$

a protože T_{μ_0} má Studentovo rozdělení, je tudíž

$$u_1 = t_{1-\alpha;\, n-1}$$

a kritérium pro **ZAMÍTNUTÍ** je tak skutečně tvaru

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