

11. cvičení z PST

4. prosince 2019

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 $\mathbf{x} = (x_1, \dots, x_n)$ o rozsahu n (pro nezávislé náhodné veličiny $X_i \sim N(\mu, \sigma^2)$). Hledáme teď nějaké funkce $h_1, h_2 : \mathbb{R}^n \rightarrow \mathbb{R}$ (nezávislé na volbě μ i σ) takové, že

$$P\left(h_1(X_1, \dots, X_n) \leq \mu \leq h_2(X_1, \dots, X_n)\right) = 1 - \alpha$$

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

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

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

Při dané realizaci $\mathbf{x} = (x_1, \dots, x_n)$ pak jako hledaný interval spolehlivost chápeme (číselný) interval tvaru:

$$\mu \in \left\langle h_1(x_1, \dots, x_n), h_2(x_1, \dots, x_n) \right\rangle (\subseteq \mathbb{R})$$

Je ještě dobré poznamenat, že

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

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

$$T = \frac{\bar{X} - \mu}{S_{\mathbf{X}}} \sqrt{n},$$

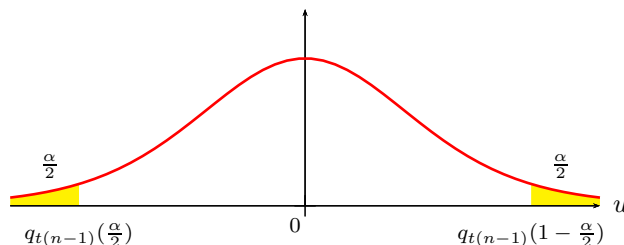
kde

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

a

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

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



Pak pro kvantil $q_{t(n-1)}$ platí, že

$$P\left(\underbrace{q_{t(n-1)}\left(\frac{\alpha}{2}\right)}_{-q_{t(n-1)}\left(1-\frac{\alpha}{2}\right)} \leq T \leq q_{t(n-1)}\left(1-\frac{\alpha}{2}\right)\right) = 1 - \alpha$$

a

$$P\left(T < -q_{t(n-1)}\left(1-\frac{\alpha}{2}\right)\right) = \frac{\alpha}{2} = P\left(q_{t(n-1)}\left(1-\frac{\alpha}{2}\right) < T\right).$$

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

$$-q_{t(n-1)}\left(1-\frac{\alpha}{2}\right) \leq \frac{\bar{X} - \mu}{S_X} \sqrt{n} \leq q_{t(n-1)}\left(1-\frac{\alpha}{2}\right)$$

$$\underbrace{\bar{X} - \frac{S_X}{\sqrt{n}} \cdot q_{t(n-1)}\left(1-\frac{\alpha}{2}\right)}_{h_1(X_1, \dots, X_n)} \leq \mu \leq \underbrace{\bar{X} + \frac{S_X}{\sqrt{n}} \cdot q_{t(n-1)}\left(1-\frac{\alpha}{2}\right)}_{h_2(X_1, \dots, X_n)}$$

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

$$\mu \in \left\langle \bar{\mathbf{x}} - \frac{s_{\mathbf{x}}}{\sqrt{n}} q_{t(n-1)}\left(1-\frac{\alpha}{2}\right), \bar{\mathbf{x}} + \frac{s_{\mathbf{x}}}{\sqrt{n}} q_{t(n-1)}\left(1-\frac{\alpha}{2}\right) \right\rangle$$

Pro názornost uveďme funkce h_1 a h_2 více rozepsané:

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

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

11.1 (intervalový odhad pro střední hodnotu)

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

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

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:

$$\mu \in \left\langle \bar{\mathbf{x}} - \frac{s_{\mathbf{x}}}{\sqrt{n}} q_{t(n-1)}\left(1-\frac{\alpha}{2}\right), \bar{\mathbf{x}} + \frac{s_{\mathbf{x}}}{\sqrt{n}} q_{t(n-1)}\left(1-\frac{\alpha}{2}\right) \right\rangle.$$

Pro jeho vyčíslení potřebujeme znát realizaci výběrového průměru $\bar{\mathbf{x}}$ a výběrového rozptylu $s_{\mathbf{x}}^2$ z realizace $\mathbf{x} = (x_1, \dots, x_n)$ rozsahu $n = 9$:

- realizace výběrového průměru:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{0.2 + 0.23 + \dots + 0.21}{9} = \frac{1.7}{9} \doteq 0.189$$

- realizace výběrového rozptylu

$$\begin{aligned} s_x^2 &= \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 \right) - \frac{n}{n-1} \bar{x}^2 = \\ &= \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 \right) - \frac{1}{n(n-1)} \left(\sum_{i=1}^n x_i \right)^2 = \\ &= \underbrace{\frac{0.2^2 + 0.23^2 + \dots + 0.21^2}{8}}_{= \frac{2.9448}{8} = 0.3681} - \frac{1.7^2}{9 \cdot 8} \doteq 7.6 \cdot 10^{-4}, \end{aligned}$$

- realizace směrodatné odchylky

$$s_x = \sqrt{s_x^2} \doteq 2.76 \cdot 10^{-2}$$

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

$$\begin{aligned} \mu &\in \left\langle \bar{x} - \frac{s_x}{\sqrt{n}} q_{t(n-1)} \left(1 - \frac{\alpha}{2}\right), \bar{x} + \frac{s_x}{\sqrt{n}} q_{t(n-1)} \left(1 - \frac{\alpha}{2}\right) \right\rangle \doteq \\ &\doteq \left\langle 0.189 - \frac{2.76 \cdot 10^{-2}}{3} \underbrace{q_{t(8)}(0.95)}_{1.86}, 0.189 + \frac{2.76 \cdot 10^{-2}}{3} q_{t(8)}(0.95) \right\rangle \doteq \\ &\doteq \langle 0.172, 0.206 \rangle. \end{aligned}$$

11.2 (intervalový odhad pro střední hodnotu a rozptyl)

Soubor dat

$$\mathbf{x} = (75, 85, 58, 72, 70, 75)$$

je náhodným výběrem z normálního rozdělení $N(\mu, \sigma^2)$. Stanovte

- horní a dolní 95% interval spolehlivosti pro střední hodnotu μ ;
- oboustranný symetrický 90% interval spolehlivosti pro rozptyl σ^2 .

Řešení:

- K určení intervalového odhadu opět použijeme statistiku

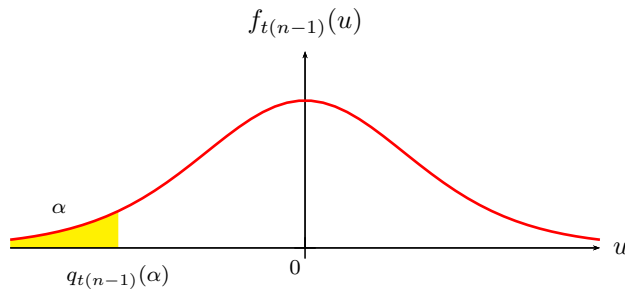
$$T = \frac{\bar{X} - \mu}{S_X} \sqrt{n}$$

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).

(a1) **Horní** interval spolehlivosti pro μ (tj. μ bude omezené *shora*) dostaneme ze vztahu

$$P\left(\underbrace{q_{t(n-1)}(\alpha)}_{-q_{t(n-1)}(1-\alpha)} \leq T\right) = 1 - \alpha$$

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



Pro realizaci $t = \frac{\bar{x} - \mu}{s_x} \sqrt{n}$ tedy nastává případ

$$q_{t(n-1)}(1 - \alpha) \leq \frac{\bar{x} - \mu}{s_x} \sqrt{n}.$$

s pravděpodobností $1 - \alpha = 95\%$. Po úpravě máme horní interval spolehlivosti pro μ ve tvaru:

$$\mu \leq \bar{x} + \frac{s_x}{\sqrt{n}} q_{t(5)}(0.95).$$

Pro jeho vyčíslení potřebujeme znát realizaci výběrového průměru \bar{x} a výběrového rozptylu s_x^2

$$\bar{x} = \frac{1}{6} \sum_{i=1}^6 x_i = \frac{435}{6} = 72.5$$

$$s_x^2 = \frac{1}{5} \left(\sum_{i=1}^6 x_i^2 - 6 \cdot (\bar{x})^2 \right) = \frac{385.5}{5} \doteq 77.1, \quad s_x \doteq 8.781.$$

Z tabulek kvantilů Studentova rozdělení dostaneme $q_{t(5)}(0.95) \doteq 2.02$, a hledaný interval je tedy

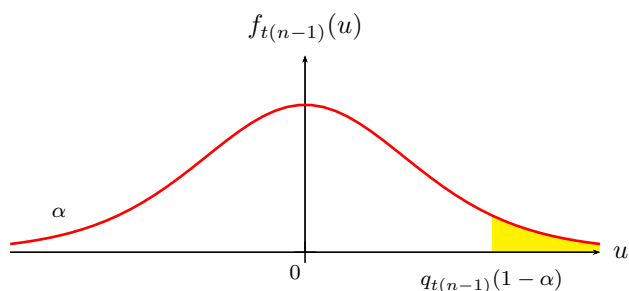
$$\mu \leq \underbrace{72.5 + \frac{8.781}{\sqrt{6}} 2.02}_{\doteq 79.74}$$

neboli

$$\mu \in (-\infty, 79.74).$$

(a2) Podobně dostaneme **dolní** interval spolehlivosti pro μ (tj. μ bude omezené *zdeola*) ze vztahu pro **horní** $1 - \alpha = 95\%$ intervalový odhad veličiny T (viz obrázek)

$$P\left(T \leq q_{t(n-1)}(1 - \alpha)\right) = 1 - \alpha$$



tedy

$$\frac{\bar{x} - \mu}{s_x} \sqrt{n} \leq q_{t(n-1)}(1 - \alpha).$$

a po úpravě

$$\bar{x} - \frac{s_x}{\sqrt{n}} q_{t(5)}(0.95) \leq \mu$$

a dosazení

$$\underbrace{72.5 - \frac{8.781}{\sqrt{6}} 2.02}_{=65.26} \leq \mu$$

máme

$$\mu \in (65.26, \infty).$$

(b) K určení intervalového odhadu použijeme statistiku

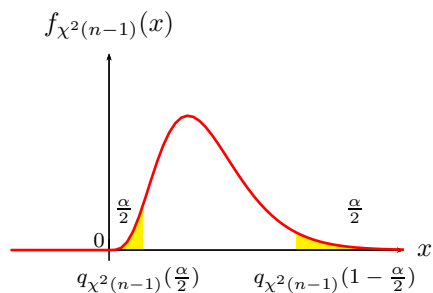
$$T = \frac{(n-1)S_{\mathbf{X}}^2}{\sigma^2}$$

která má tzv. chi-kvadrát rozdělení $\chi^2(n-1)$ s $n-1$ stupni volnosti, kde $n=6$ je opět rozsah souboru (ten může být i malé číslo).

Oboustranný symetrický interval spolehlivosti $1-\alpha=90\%$ pro σ^2 dostaneme zase analogicky z oboustranného symetrického intervalového odhadu veličiny T (viz obrázek):

$$P\left(q_{\chi^2(n-1)}\left(\frac{\alpha}{2}\right) \leq T \leq q_{\chi^2(n-1)}\left(1-\frac{\alpha}{2}\right)\right) = 1-\alpha$$

Graf zde představuje opět hustotu $f_{\chi^2(n-1)}$ veličiny T .



Pro realizaci $t = \frac{(n-1)s_x^2}{\sigma^2}$ tedy nastává případ

$$q_{\chi^2(n-1)}\left(\frac{\alpha}{2}\right) \leq \frac{(n-1)s_x^2}{\sigma^2} \leq q_{\chi^2(n-1)}\left(1-\frac{\alpha}{2}\right)$$

s pravděpodobností $1 - \alpha = 90\%$. Po úpravě máme

$$\frac{(n-1) s_{\mathbf{x}}^2}{q_{\chi^2(n-1)}\left(1 - \frac{\alpha}{2}\right)} \leq \sigma^2 \leq \frac{(n-1) s_{\mathbf{x}}^2}{q_{\chi^2(n-1)}\left(\frac{\alpha}{2}\right)}$$

Z tabulek kvantilů χ^2 -rozdělení pak dostaneme, že hledaný interval je pro rozptyl σ^2 (o spolehlivosti $0.9 = 1 - \alpha$) je:

$$\begin{aligned} \sigma^2 &\in \left\langle \frac{(n-1) s_{\mathbf{x}}^2}{q_{\chi^2(n-1)}\left(1 - \frac{\alpha}{2}\right)}, \frac{(n-1) s_{\mathbf{x}}^2}{q_{\chi^2(n-1)}\left(\frac{\alpha}{2}\right)} \right\rangle \doteq \\ &\doteq \left\langle \frac{(n-1) s_{\mathbf{x}}^2}{q_{\chi^2(5)}(0.95)}, \frac{(n-1) s_{\mathbf{x}}^2}{q_{\chi^2(5)}(0.05)} \right\rangle \doteq \\ &\doteq \left\langle \frac{5 \cdot 77.1}{11.07}, \frac{5 \cdot 77.1}{1.145} \right\rangle \doteq \\ &\doteq \langle 34.82, 336.68 \rangle . \end{aligned}$$

Všimněte si, že intervalový odhad výběrového rozptylu nemá střed ve svém bodovém odhadu $s_{\mathbf{x}}^2 \doteq 77.1$. Velké hodnoty rozptylu i intervalu jsou dány tím, že některé naměřené hodnoty (a sice 58 a 85) se dost odchylojí od výběrového průměru $\bar{\mathbf{x}} = 72.5$.

11.3 (intervalový odhad pro rozptyl)

Deset opakovaných měření obsahu alkoholu ve vzorku krve má průměrnou hodnotu $\bar{\mathbf{x}} = 0.15$ promile (alkoholu v krvi) a směrodatnou odchylku $s_{\mathbf{x}} = 0.01$ promile (alkoholu v krvi). Jakou hodnotu σ_0 překročí chyba metody (t.j. směrodatná odchylka) s pravděpodobností nejvýše $\alpha = 1\%$? Uveďte použité předpoklady.

Řešení:

U veličiny

$$Y = \text{“naměřený obsah alkoholu v krvi (v promilích)”}$$

budeme předpokládat normální rozdělení $N(\mu, \sigma^2)$. Naše veličina

$$X = \text{“chyba měření obsahu alkoholu v krvi (v promilích)”}$$

bude tedy určena jako $X = Y - \mu$ s rozdělením $N(0, \sigma^2)$. Jednotlivá měření považujeme za nezávislá. Hledáme **horní** $1 - \alpha = 99\%$ intervalový odhad pro **parametr rozptylu σ^2** , který bude tvaru $(0, \sigma_0^2)$.

Pro statistiku

$$T = \frac{(n-1) S_{\mathbf{X}}^2}{\sigma^2}$$

s χ^2 -rozdělením s $n - 1 = 9$ stupni volnosti máme, že **dolní** $1 - \alpha = 99\%$ interval spolehlivosti pro **realizaci t** veličiny T je:

$$q_{\chi^2(n-1)}(\alpha) \leq t = \frac{(n-1) s_{\mathbf{x}}^2}{\sigma^2}$$

takže

$$\sigma^2 \leq \frac{(n-1) s_{\mathbf{x}}^2}{q_{\chi^2(n-1)}(\alpha)} =: \sigma_0^2$$

a hledaná hranice tak je

$$\sigma_0 = s_{\mathbf{x}} \cdot \sqrt{\frac{n-1}{q_{\chi^2(n-1)}(\alpha)}} = 0.01 \cdot \sqrt{\frac{9}{q_{\chi^2(9)}(0.01)}} \doteq \frac{0.03}{\sqrt{2.0879}} \doteq 0.02076.$$

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).

11.4 (test střední hodnoty normálního rozdělení při známém rozptylu)

Teploměrem, o jehož chybě předpokládáme, že má normální rozdělení se směrodatnou odchylkou $\sigma = 3^\circ$, jsme provedli 30 měření stejné teploty. Průměrný výsledek byl 101° . Otestujte na hladině významnosti 5 %, zda teplota nepřesahuje 100° .

Řešení:

Naše veličina

$$X = \text{“naměřená teplota”}$$

(v jednotkách $^\circ$) má podle předpokladu rozdělení $N(\mu, \sigma^2)$, kde $\sigma = 3^\circ$.

Podle zadání máme otestovat hypotézu o střední hodnotě

$$\mathbf{H}_0 : \mu \leq \mu_0 (= 100^\circ)$$

proti alternativní hypotéze:

$$\mathbf{H}_1 : \mu > \mu_0 (= 100^\circ).$$

Hodnotu rozptylu σ^2 známe. Takže použijeme test střední hodnoty se známým rozptylem (protože tím se dozvíme víc než kdybychom uvažovali neznámý rozptyl).

Pomocí testovací statistiky:

Realizaci testovací statistiky $T = \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n}$ pro $n = 30$

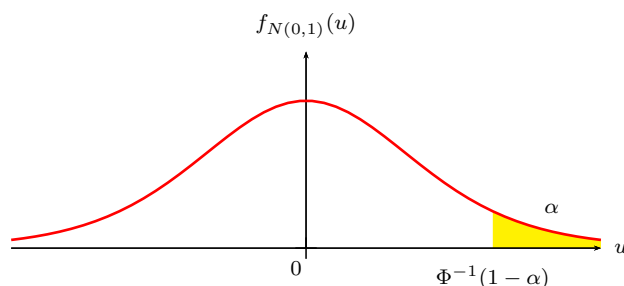
$$t = \frac{\bar{x} - \mu_0}{\sigma} \sqrt{n} = \frac{101 - 100}{3} \sqrt{30} \doteq 1.8257$$

porovnáme s kvantilem $\Phi^{-1}(1 - \alpha) = \Phi^{-1}(0.95) \doteq 1.64$ pro $\alpha = 0.05$. Protože je splněno zamítací kritérium

$$\Phi^{-1}(1 - \alpha) < t$$

tak nulovou hypotézu \mathbf{H}_0 zamítáme.

Proč má zamítací kritérium uvedený tvar: Za předpokladu platnosti nulové hypotézy, tj. pokud $E(X) = \mu \leq \mu_0$, bude mít hustota pro statistiku T svůj vrchol (a i střední hodnotu) v intervalu $(-\infty, 0)$. 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 rozdělení $N(0, 1)$. Chyba 1. druhu s pravděpodobností α zde tedy bude soustředěna jen na jedné straně a je určena prave kvantilem rozdělení $N(0, 1)$, tj. funkcí Φ^{-1} :



Pomocí intervalu spolehlivosti:

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

$$\Phi^{-1}(1 - \alpha) < t \left(= \frac{\bar{x} - \mu_0}{\sigma} \sqrt{n} \right)$$

se dá ekvivalentně přepsat jako

$$\mu_0 < \bar{x} - \frac{\sigma}{\sqrt{n}} \cdot \Phi^{-1}(1 - \alpha)$$

neboli

$$\mu_0 \notin \left\langle \bar{x} - \frac{\sigma}{\sqrt{n}} \cdot \Phi^{-1}(1 - \alpha) , \infty \right\rangle =: \langle \mu_L, \infty \rangle$$

což je tvar zamítacího kritéria s použitím interval spolehlivosti.

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

$$\langle \mu_L, \infty \rangle = \left\langle 101 - \frac{3}{\sqrt{30}} \cdot 1.64 , \infty \right\rangle = \langle 100.1, \infty \rangle$$

Protože máme $\mu_0 = 100 \notin \langle 100.1, \infty \rangle = \langle \mu_L, \infty \rangle$, (tj. kritérium pro zamítnutí je splněno) hypotézu \mathbf{H}_0 **ZAMÍTÁME** na hladině 5%.

(Výsledek musel samozřejmě dopadnout stejně jako při testovací statistice, protože je to ekvivalentní princip.)

Tvar intervalu spolehlivosti si můžeme intuitivně zapamatovat takto:

Při pravdivosti \mathbf{H}_0 je $\mu \leq \mu_0$. Protože $\langle \mu_L, \infty \rangle$ 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.

11.5 (test střední hodnoty normálního rozdělení při neznámém rozptylu)

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

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

Jak dopadne testování těchto hypotéz na hladině 1%?

Řešení:

U veličiny

$$X = \text{“spotřeba automobilu”}$$

budeme předpokládat normální rozdělení $N(\mu, \sigma^2)$. Jednotlivá měření X_i , pro $i = 1, \dots, 49$, 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*):

$$T = \frac{\bar{X} - \mu_0}{S_X} \sqrt{n}$$

kde opět

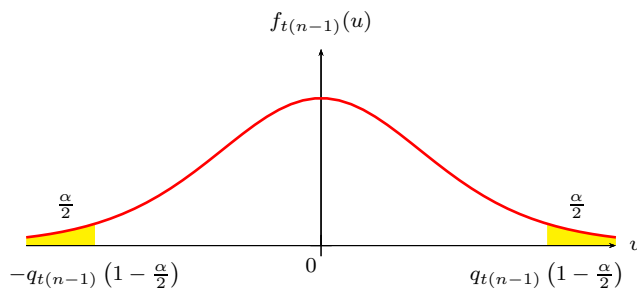
- veličina $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ je *výběrový průměr* a
- veličina $S_X^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ je *výběrový rozptyl*.

Kritérium pro **ZAMÍTNUTÍ H_0** (na hladině α) je tvaru

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

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

$$\begin{aligned} P_{(H_0 \text{ platí})}(\text{nastává chyba 1. druhu}) &= P_{(H_0 \text{ platí})}(\text{zamítáme } \mathbf{H}_0) = \\ &= P_{(H_0 \text{ platí})}(|T| > q_{t(n-1)}(1 - \frac{\alpha}{2})) = \frac{\alpha}{2} + \frac{\alpha}{2} = \alpha \end{aligned}$$

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

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

Protože pro $\alpha = 0.05$ je

$$|t| = 1.75 \not> 2.011 \doteq q_{t(48)}(0.975) = q_{t(n-1)} \left(1 - \frac{\alpha}{2}\right) ,$$

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

Naše měření tak nejsou dostačující na to, abychom mohli zamítnout tvrzení výrobce na hladině významnosti 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 H_0 také **NEZAMÍTÁME**.

(Pro úplnost si ale stejně ještě vyjádříme příslušnou podmínku: $|t| = 1.75 \not> 2.682 \doteq q_{t(48)}(0.995)$.)

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 H_0 na hladině α

$$|t| > q_{t(n-1)}\left(1 - \frac{\alpha}{2}\right)$$

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

$$\mu_0 \notin \left\langle \bar{x} - \frac{s_x}{\sqrt{n}} \cdot q_{t(n-1)}\left(1 - \frac{\alpha}{2}\right), \quad \bar{x} + \frac{s_x}{\sqrt{n}} \cdot q_{t(n-1)}\left(1 - \frac{\alpha}{2}\right) \right\rangle =: \langle \mu_L, \mu_U \rangle$$

což nám dává oboustranný interval spolehlivosti (pro μ).

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

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

Protože máme $\mu_0 = 8 \in \langle 7.94, 8.86 \rangle = \langle \mu_L, \mu_U \rangle$, hypotézu 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

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

proti alternativní hypotéze:

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

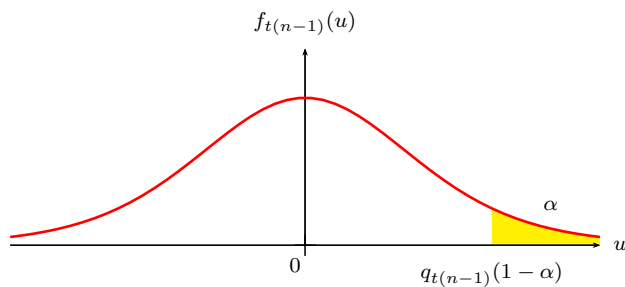
Pomocí testovací statistiky:

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

$$t > q_{t(n-1)}(1 - \alpha) \Leftrightarrow \text{zamítáme } \tilde{H}_0 \text{ (na hladině } \alpha \text{)} .$$

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 $(-\infty, 0)$. 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. "Nejhorsší" 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). **POZOR**, pro $E(X) = \mu < \mu_0$ už T zase NEMA t -rozdělení!!

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



Podobně jako předtím máme

$$P_{(\tilde{H}_0 \text{ platí})}(\text{nastává chyba 1. druhu}) = P_{(\tilde{H}_0 \text{ platí})}(\text{zamítáme } \tilde{H}_0) = \\ = P_{(\tilde{H}_0 \text{ platí})}(T > q_{t(n-1)}(1 - \alpha)) = \alpha$$

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

$$t = 1.75 > 1.677 \doteq q_{t(48)}(0.95) = q_{t(n-1)}(1 - \alpha),$$

hypotézu \tilde{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

$$t = 1.75 \not> 2.407 \doteq q_{t(48)}(0.99),$$

takže při této hladině hypotézu \tilde{H}_0 naopak **NEZAMÍTNEME**.

Pomocí intervalového odhadu:

Kritérium pro zamítnutí \tilde{H}_0 na hladině α

$$q_{t(n-1)}(1 - \alpha) < t$$

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

$$\mu_0 \notin \left\langle \bar{x} - \frac{s_x}{\sqrt{n}} \cdot q_{t(n-1)}(1 - \alpha), +\infty \right\rangle =: \langle \mu_L, +\infty \rangle$$

což nám dává dolní interval spolehlivosti (pro μ).

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

$$\langle \mu_L, +\infty \rangle = \left\langle 8.4 - \frac{1.6}{\sqrt{49}} \cdot 1.677, +\infty \right\rangle = \langle 8.017, +\infty \rangle$$

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

+++++

Poznámka: Podle zadání jsme uvažovali případ, kde se ptáme na rovnost (tj. $\mu = \mu_0$). V této situaci máme jedinou možnost, jak zvolit nulovou hypotézu - a sice výše uvedeným způsobem. Jako nulovou hypotézu není možné zvolit případ $\mu \neq \mu_0$, protože množina $\{\mu \in \mathbb{R} \mid \mu \neq \mu_0\}$ není uzavřená.

Při testu hypotézy $\tilde{\mathbf{H}}_0: \mu \leq \mu_0$ se snažíme vyhnout tomu, že bychom omylem poškodili *výrobce*. Zde jsme na hladině významnosti $\alpha = 5\%$ hypotézu $\tilde{\mathbf{H}}_0$ zamítli.

Můžeme si ještě zkusit otestovat hypotézu $\mathbf{H}'_0: \mu \geq \mu_0$, při níž se naopak snažíme vyhnout tomu, že bychom omylem poškodili *uživatele*. Na hladině významnosti $\alpha = 5\%$ pak dostaneme:

$$t = 1.75 \not\leq -1.6772 \doteq -q_{t(48)}(0.95) = q_{t(n-1)}(\alpha)$$

takže hypotézu \mathbf{H}'_0 uživatelů **NEZAMÍTNEME**.

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.

11.6 (test rozptylu normálního rozdělení)

Generátor náhodných čísel s normovaným normálním $N(0, 1)$ rozdělením dal následující výsledky:

$$\mathbf{x} = (-0.503, 0.811, 1.078, -0.501, 0.562, -1.032, 0.152, 0.859, -0.156, 2.213)$$

Posuďte na hladině významnosti 5 %, zda data odpovídají předpokládanému rozptylu.

Řešení:

Zadání si vyjádříme tak, že u veličiny

$$X = \text{“hodnota náhodného čísla”}$$

budeme předpokládat normální rozdělení $N(\mu, \sigma^2)$ a hypotéza bude o parametru rozptylu σ^2 .

Na hladině $\alpha = 5\%$ budeme testovat hypotézu o střední hodnotě

$$\mathbf{H}_0 : \sigma^2 = \sigma_0^2 (:= 1)$$

proti alternativní hypotéze:

$$\mathbf{H}_1 : \sigma^2 \neq \sigma_0^2 (:= 1) .$$

Pomocí testovací statistiky:

Použijeme statistiku:

$$T = \frac{(n-1)S_X}{\sigma_0^2}$$

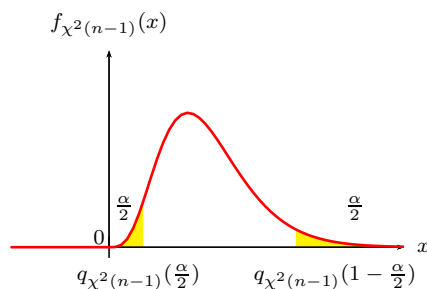
která za předpokladu, že $\sigma^2 = \sigma_0^2$ (tj. když platí \mathbf{H}_0) má $\chi^2(n-1)$ -rozdělení.

Kritérium pro **ZAMÍTNUTÍ** \mathbf{H}_0 (na hladině α) bude tvaru

$$\left(t < q_{\chi^2(n-1)}\left(\frac{\alpha}{2}\right) \vee q_{\chi^2(n-1)}\left(1 - \frac{\alpha}{2}\right) < t \right) \Leftrightarrow \text{zamítáme } \mathbf{H}_0 \text{ (na hladině } \alpha \text{)} .$$

kde t je hodnota T na základě naměřených dat.

Proč má zamítací kritérium uvedený tvar: Za předpokladu platnosti nulové hypotézy, tj. pokud $D(X) = \sigma_0^2$, bude mít statistika T rozdělení χ^2 s $n - 1$ stupni volnosti:



Očekávané hodnoty takovéto statistiky T by se měly pohybovat kolem střední hodnoty (která je $E(T) = n - 1$ pro $n - 1$ stupňů volnosti). 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

$$\begin{aligned} P_{(\mathbf{H}_0 \text{ platí})}(\text{nastává chyba 1. druhu}) &= P_{(\mathbf{H}_0 \text{ platí})}(\text{zamítáme } \mathbf{H}_0) = \\ &= P_{(\mathbf{H}_0 \text{ platí})}\left(T < q_{\chi^2(n-1)}\left(\frac{\alpha}{2}\right) \vee q_{\chi^2(n-1)}\left(1 - \frac{\alpha}{2}\right) < T\right) = \frac{\alpha}{2} + \frac{\alpha}{2} = \alpha \end{aligned}$$

Vyčíslení statistiky:

- rozsah souboru je $n = 10$,
- $\sum_i x_i = 3.483$,
- $\sum_i x_i^2 = 9.387373$,
- $(n - 1)s_{\mathbf{x}}^2 = \frac{1}{n} \left(n \sum_i x_i^2 - (\sum_i x_i)^2 \right) = \frac{93.87373 - (3.483)^2}{10} \doteq 8.174$.

Realizace testovací statistiky je

$$t = \frac{(n - 1) s_{\mathbf{x}}^2}{\sigma_0^2} \doteq \frac{8.174}{1} = 8.174,$$

Protože

$$t = 8.174 \in \left\langle \underbrace{q_{\chi^2(9)}(0.025)}_{2.7}, \underbrace{q_{\chi^2(9)}(0.975)}_{19.02} \right\rangle$$

nulovou hypotézu **NEZAMÍTÁME** (na hladině $\alpha = 5\%$).

Pomocí intervalu spolehlivosti:

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

$$t \notin \langle q_{\chi^2(9)}(0.025), q_{\chi^2(9)}(0.975) \rangle$$

se dá ekvivalentně přepsat (při vyjádření $t = \frac{(n-1)s_{\mathbf{x}}^2}{\sigma_0^2}$) jako

$$\sigma_0^2 \notin \left\langle \frac{(n-1) s_{\mathbf{x}}^2}{q_{\chi^2(n-1)}\left(1 - \frac{\alpha}{2}\right)}, \frac{(n-1) s_{\mathbf{x}}^2}{q_{\chi^2(n-1)}\left(\frac{\alpha}{2}\right)} \right\rangle =: \langle \sigma_L^2, \sigma_U^2 \rangle$$

což nám dává oboustranný interval spolehlivosti (pro σ^2).

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

$$\langle \sigma_L^2, \sigma_U^2 \rangle = \left\langle \frac{8.174}{19.02}, \frac{8.174}{2.7} \right\rangle = \langle 0.43, 3.03 \rangle$$

Protože máme $\sigma_0^2 = 1 \in \langle 0.43, 3.03 \rangle = \langle \sigma_L^2, \sigma_U^2 \rangle$, (tj. zamítací kritérium není splněno) hypotézu \mathbf{H}_0 **NEZAMÍTÁME** na hladině 5%.

(Výsledek musel samozřejmě dopadnout stejně jako při testovací statistice, protože je to ekvivalentní princip.)