11. cvičení z PST

2. prosince 2020

11.1 (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:

$$\mu \in \left\langle \overline{\boldsymbol{x}} - \frac{s_{\boldsymbol{x}}}{\sqrt{n}}\, q_{\mathrm{t}(n-1)}\left(1 - \tfrac{\alpha}{2}\right),\quad \overline{\boldsymbol{x}} + \frac{s_{\boldsymbol{x}}}{\sqrt{n}}\, q_{\mathrm{t}(n-1)}\left(1 - \tfrac{\alpha}{2}\right) \right\rangle .$$

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

- realizace výběrového průměru:

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

- realizace výběrového rozptylu

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

- realizace směrodatné odchylky

$$s_{\boldsymbol{x}} = \sqrt{s_{\boldsymbol{x}}^2} \doteq 2.76\cdot 10^{-2}$$

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

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

11.2 (intervalový odhad pro střední hodnotu a rozptyl)
V terénu jsme naměřili 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

(a) horní a dolní 95% interval spolehlivosti pro střední hodnotu μ;

(b) oboustranný symetrický 90% interval spolehlivosti pro rozptyl σ^2.

Řešení:
(a) K určení intervalového odhadu opět použijeme statistiku

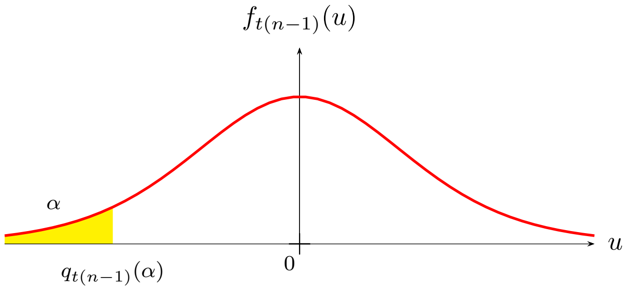
$$T = \frac{\overline{\boldsymbol{X}} - \mu}{S_{\boldsymbol{X}}}\sqrt{n}$$

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).

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

$$P\Big(\underbrace{q_{\mathrm{t}(n-1)}(\alpha)}_{-q_{\mathrm{t}(n-1)}(1-\alpha)} \ \le\ T\Big) = 1-\alpha$$

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



Pro realizaci $t = \frac{\overline{\boldsymbol{x}}-\mu}{s_{\boldsymbol{x}}}\sqrt{n}$ tedy nastává případ

$$q_{\mathrm{t}(n-1)}\,(1-\alpha)\ \le\ \tfrac{\overline{\boldsymbol{x}}-\mu}{s_{\boldsymbol{x}}}\sqrt{n}\ .$$

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

$$\mu\ \le\ \overline{\boldsymbol{x}} + \frac{s_{\boldsymbol{x}}}{\sqrt{n}}\,q_{\mathrm{t}(5)}(0.95)\ .$$

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

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

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

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

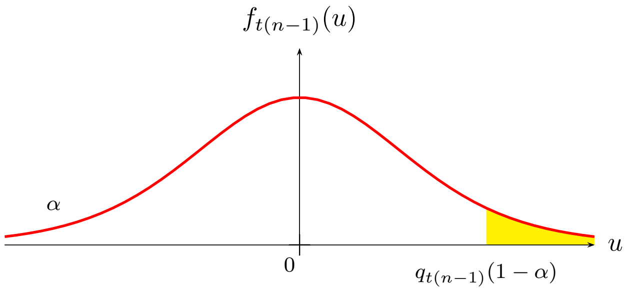
$$\mu \;\le\; \underbrace{72.5 + \frac{8.781}{\sqrt{6}}2.02}_{\doteq 79.74}\ .$$

neboli

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

(a2) 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)

$$P\Big(T \;\le\; q_{\mathrm{t}(n-1)}(1-\alpha)\Big) = 1-\alpha$$



tedy

$$\frac{\overline{\boldsymbol{x}} - \mu}{s_{\boldsymbol{x}}}\sqrt{n} \;\le\; q_{\mathrm{t}(n-1)}\,(1-\alpha)\,.$$

a po úpravě

$$\overline{\boldsymbol{x}} - \frac{s_{\boldsymbol{x}}}{\sqrt{n}}q_{\mathrm{t}(5)}(0.95) \;\le\; \mu$$

a dosazení

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

máme

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

(b) K určení intervalového odhadu použijeme statistiku

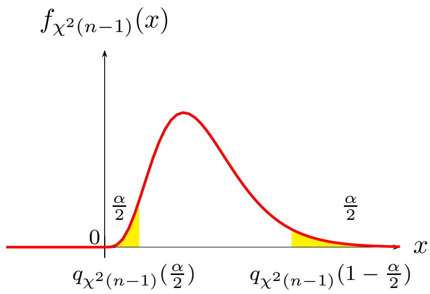
$$T = \frac{(n-1)S_{\boldsymbol{X}}^2}{\sigma^2}$$

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):

$$P\Big(q_{\chi^2(n-1)}\left(\tfrac{\alpha}{2}\right) \;\le\; T \;\le\; q_{\chi^2(n-1)}\left(1-\tfrac{\alpha}{2}\right)\Big) = 1-\alpha$$

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



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

$$q_{\chi^2(n-1)}\left(\frac{\alpha}{2}\right) \;\le\; \tfrac{(n-1)s_{\boldsymbol{x}}^2}{\sigma^2} \;\le\; q_{\chi^2(n-1)}\left(1-\frac{\alpha}{2}\right)$$

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

$$\frac{(n-1)\,s_{\boldsymbol{x}}^2}{q_{\chi^2(n-1)}\left(1-\frac{\alpha}{2}\right)} \;\le\; \sigma^2 \;\le\; \frac{(n-1)\,s_{\boldsymbol{x}}^2}{q_{\chi^2(n-1)}\left(\frac{\alpha}{2}\right)}$$

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

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

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

11.3 (intervalový odhad pro rozptyl)

Deset opakovaných měření obsahu cukru ve vzorku má průměrnou hodnotu $\overline{\boldsymbol{x}} = 0.15$ procenta a směrodatnou odchylku $s_{\boldsymbol{x}} = 0.01$ procenta. Jakou hodnotu σ_0 překročí chyba metody (t.j. směrodatná odchylka) s pravděpodobností nejvýše $\alpha = 1\,\%$? Uveďte použité předpoklady.

Řešení:
U veličiny

Y = *"naměřený obsah cukru (v procentech)"*

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

X = *"chyba měření obsahu cukru (v procentech)"*

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 $(0, \sigma_0^2\rangle$.

Pro statistiku

$$T = \frac{(n-1)\,S^2_{\boldsymbol{X}}}{\sigma^2}$$

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

$$q_{\chi^2(n-1)}(\alpha) \le t = \tfrac{(n-1)\,s^2_{\boldsymbol{x}}}{\sigma^2}$$

takže

$$\sigma^2 \le \tfrac{(n-1)\,s^2_{\boldsymbol{x}}}{q_{\chi^2(n-1)}(\alpha)} =: \sigma_0^2$$

a hledaná hranice tak je

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

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 $\sigma = 3°$, jsme provedli 30 měření stejné teploty. Průměrný výsledek byl 101°. Otestujte na hladině významnosti 5 %, zda teplota nepřesahuje 100°.

Řešení:
Naše veličina

X = *"naměřená teplota"*

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

Podle zadání máme otestovat hypotézu o střední hodnotě

$$\mathbf{H}_0 : \mu \le \mu_0 (= 100°)$$

proti alternativní hypotéze:

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

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 $T = \frac{\overline{\boldsymbol{X}}-\mu_0}{\sigma}\sqrt{n}$ pro $n = 30$

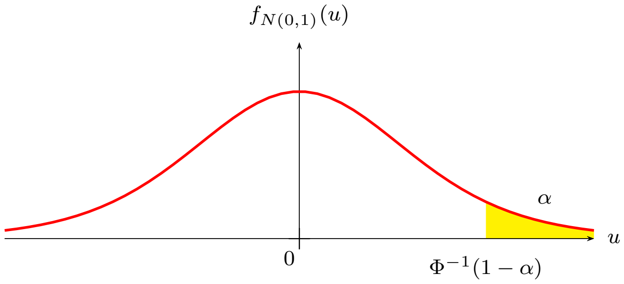
$$t = \frac{\overline{\boldsymbol{x}} - \mu_0}{\sigma}\sqrt{n} = \frac{101 - 100}{3}\sqrt{30} \doteq 1.8257$$

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 $(-\infty, 0\rangle$. 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 rozdělení $N(0,1)$. Chyba 1. druhu s pravděpodobností α zde tedy bude soustředěna jen na jedné straně a je určena právě kvantilem rozdělení $N(0,1)$, tj. funkcí Φ^{-1}:



Pomocí intervalu spolehlivosti:

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

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

se dá ekvivalentně přepsat jako

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

neboli

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

což je tvar zamítacího kritéria s použitím interval spolehlivosti.

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

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

Protože máme $\mu_0 = 100 \notin \langle 100.1,\ \infty) = \langle \mu_L,\ \infty)$, (tj. kriterium 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 $\langle \mu_L,\ \infty)$ 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% pravděpodobnosti), je to důvod k zamítnutí.

11.5 (test střední hodnoty normálního rozdělení při neznámém rozptylu)

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

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

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

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 opět

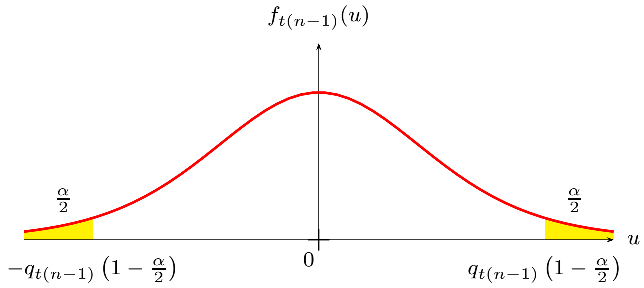
- veličina $\overline{\boldsymbol{X}} = \frac{1}{n} \sum_{i=1}^{n} X_i$ je *výběrový průměr* a
- veličina $S_{\boldsymbol{X}}^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{\boldsymbol{X}})^2$ je *výběrový rozptyl.*

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

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

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

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

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

$$|t| = 1.75 \ngtr 2.011 \doteq q_{t(48)}(0.975) = q_{t(n-1)}\left(1 - \frac{\alpha}{2}\right)\ ,$$

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 $\mathbf{H}_0$ také **NEZAMÍTÁME**.
(Pro úplnost si ale stejně ještě vyjádříme příslušnou podmínku: $|t| = 1.75 \ngtr 2.682 \doteq q_{t(48)}(0.995)$.)

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 $\mathbf{H}_0$ na hladině α

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

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

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

což nám dává oboustranný interval spolehlivosti (pro μ).

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

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

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

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

proti alternativní hypotéze:

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

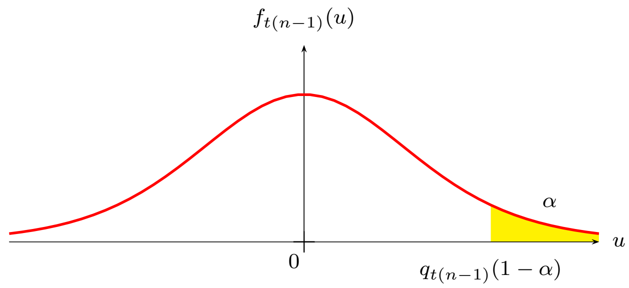
Pomocí testovací statistiky:

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

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

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). POZOR, pro $E(X) = \mu < \mu_0$ už T zase NEMÁ t-rozdělení!!

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



Podobně jako předtím máme

$$P_{(\widetilde{H}_0 \text{ platí})}\Big(\text{nastává chyba 1. druhu}\Big) = P_{(\widetilde{H}_0 \text{ platí})}\Big(\text{zamítáme } \widetilde{\mathbf{H}}_0\Big) =$$

$$= P_{(\widetilde{H}_0 \text{ platí})}\Big(\ T > q_{t(n-1)}(1-\alpha)\ \Big) = \alpha$$

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 $\widetilde{\mathbf{H}}_0$ **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

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

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

Pomocí intervalového odhadu:

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

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

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

$$\mu_0 \notin \Big\langle \ \overline{\boldsymbol{x}} - \frac{s_{\boldsymbol{x}}}{\sqrt{n}} \cdot q_{t(n-1)}(1-\alpha) \ , \quad +\infty \ \Big) \quad =: \langle \mu_L, \ +\infty)$$

což nám dává dolní interval spolehlivosti (pro μ).

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

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

Protože máme $\mu_0 = 8 \notin \langle 8.017, \ +\infty) = \langle \mu_L, \ +\infty)$, hypotézu $\widetilde{\mathbf{H}}_0$ **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 $\widetilde{\mathbf{H}}_0$: $\mu \leq \mu_0$ se snažíme vyhnout tomu, že bychom omylem poškodili *výrobce*. Zde jsme na hladině významnosti $\alpha = 5\%$ hypotézu $\widetilde{\mathbf{H}}_0$ 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:

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

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:

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

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

Řešení:

Zadání si vyjádříme tak, že u veličiny

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

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:

$$T = \frac{(n-1)S_{\boldsymbol{X}}}{\sigma_0^2}$$

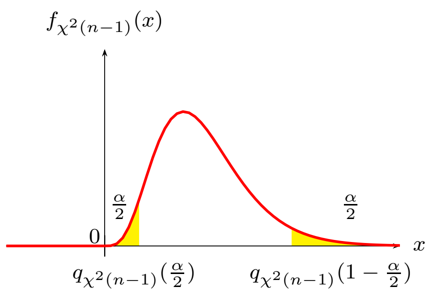
která za předpokladu, že $\sigma^2 = \sigma_0^2$ (tj. když platí $\mathbf{H}_0$) má $\chi^2(n-1)$-rozdělení.

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

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

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

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

Vyčíslení statistiky:

- rozsah souboru je $n = 10$,
- $\sum_i x_i = 3.483$,

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

Realizace testovací statistiky je

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

Protože

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

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

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

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

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

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

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

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

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

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.)

11.7 (test rozptylu normálního rozdělení)
V $n = 5$ stejných vzorcích se zjišťoval obsah příměsi X (v procentech). Výsledkem byla realizace

$$\boldsymbol{x} = (0.8, \quad 1, \quad 0.6, \quad 1.4, \quad 0.9)\ .$$

Posuďte na hladině významnosti $\alpha = 0.05$, zda směrodatná odchylka měření je nejvýše $\sigma_0 = 0.1$ procenta. Předpokládejte, že obsah příměsi X má normální rozdělení a jednotlivá měření jsou nezávislá.

Řešení:
Naše veličina X, udávající obsah příměsi v procentech, má normální rozdělení $N(\mu, \sigma^2)$. Místo testu směrodatné odchylky σ budeme (ekvivalentně) testovat rozptyl σ^2 a sice nulovou hypotézu tvaru

$\mathbf{H}_0 : \sigma^2 \leq (0.1)^2\ (= \sigma_0)^2$

proti alternativní hypotéze:

$$\mathbf{H}_1 : \sigma^2 > (0.1)^2$$

na hladině významnosti $\alpha = 0.05$.

Tentokrát budeme používat statistiku

$$T = \frac{(n-1)\, S^2_{\boldsymbol{X}}}{\sigma_0^2} \ ,$$

která má pro případ $\sigma = \sigma_0$ tzv. χ^2-rozdělení s $n-1$ stupni volnosti. Obecněji, teprve veličina $\frac{\sigma_0^2}{\sigma^2} \cdot T$ bude mít χ^2-rozdělení. Za předpokladu nulové hypotézy, tj. pro $0 \leq \sigma \leq \sigma_0$, budou očekávané hodnoty statistiky T především v intervalu $(-\infty, 1\rangle$ (ve skutečnosti to bude jen interval $\langle 0, 1\rangle$, protože T je nezáporná veličina). Kritický obor tak bude

$$W : \ \Big(q_{\chi^2(n-1)}(1-\alpha), \quad \infty \Big)$$

a kritérium pro **ZAMÍTNUTÍ** proto bude tvaru

$$\text{zamítáme } \mathbf{H}_0 \text{ (na dané hladině } \alpha) \quad \Leftrightarrow \quad t > q_{\chi^2(n-1)}\,(1-\alpha)\,.$$

Dosadíme opět konkrétní hodnoty:

$$\overline{\boldsymbol{x}} = \frac{0.8+1+0.6+1.4+0.9}{5} = \frac{4.7}{5} = 0.94 \ ,$$

$$s^2_{\boldsymbol{x}} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{\boldsymbol{x}})^2 = \frac{0.14^2+0.06^2+0.34^2+0.46^2+0.04^2}{4} = \frac{0.352}{4} = 0.088 \ .$$

Realizace testovací statistiky je

$$t = \frac{(n-1)\, s^2_{\boldsymbol{x}}}{\sigma_0^2} = \frac{4 \cdot 0.088}{(0.1)^2} = 35.2$$

a hodnota kvantilu je

$$q_{\chi^2(n-1)}\,(1-\alpha) = q_{\chi^2(4)}(0.95) \doteq 9.49 \ .$$

Protože

$$t \doteq 35.2 \not> 9.49 \doteq q_{\chi^2(4)}(0.95) \ ,$$

nulovou hypotézu **ZAMÍTÁME**.

Zdůvodnění tvaru kritického oboru: Vyznačme si závislost X a T na parametru σ jako

$$T_\sigma = \frac{(n-1) S^2_{\boldsymbol{X}_\sigma}}{(\sigma_0)^2} \ .$$

Kritický obor má být tvaru

$$W : \ (u_1, \infty) \ ,$$

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

$$(\forall\ 0 \leq \sigma \leq \sigma_0) \quad P(T_\sigma \in W) = P(u_1 < T_\sigma) \leq \alpha \ .$$

Opět případ $\sigma = \sigma_0$ je za předpokladu $\mathbf{H}_0$ ten “nejhorší” možný, jak je vidět z následujícího:

$$\sigma \leq \sigma_0 \quad \Rightarrow \quad T_\sigma \ = \ \underbrace{\frac{\sigma^2}{(\sigma_0)^2}}_{\leq 1} \cdot \frac{(n-1) S^2_{\boldsymbol{X}_\sigma}}{\sigma^2} \ \leq \ \underbrace{\frac{(n-1) S^2_{\boldsymbol{X}_\sigma}}{\sigma^2}}_{\chi^2-\text{rozdělení}} \quad \Rightarrow$$

$$\Rightarrow \quad P\Big(u_1 < T_\sigma\Big) \leq P\left(u_1 < \frac{(n-1)S^2_{\boldsymbol{X}\sigma}}{\sigma^2}\right) = 1 - F_{\chi^2(n-1)}(u_1) = P\Big(u_1 < T_{\sigma_0}\Big)$$

Vidíme tedy, že $P\Big(u_1 < T_\sigma\Big) \leq P\Big(u_1 < T_{\sigma_0}\Big)$ a hledané u_1 tak musí splňovat

$$P\Big(u_1 < T_{\sigma_0}\Big) = \alpha$$

tedy

$$u_1 = q_{\chi^2(n-1)}(1-\alpha) \ ,$$

a kritický obor je tak skutečně tvaru

$$W : \quad \Big(q_{\chi^2(n-1)}(1-\alpha), \quad \infty\Big) \ .$$