11. cvičení z PST

30. listopadu 2022

Připomenutí: Pro náhodnou veličinu X se spojitou ostře rostoucí distribuční funkcí F_X označujeme jako *kvantilovou funkci veličiny* X funkci $q_X := (F_X)^{-1} : (0,1) \to \mathbb{R}$. Pro všechna $\alpha \in (0,1)$ a $t \in \mathbb{R}$ pak platí:

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

Odsud pak ihned plyne, např. že

- $P\Big(X \leq q_X(\alpha)\Big) = \alpha$
- $P\Big(q_X(\alpha) \leq X\Big) = 1 - \alpha$ a $P\Big(q_X(1-\alpha) \leq X\Big) = \alpha$
- $P\Big(q_X(\frac{\alpha}{2}) \leq X \leq q_X(1-\frac{\alpha}{2})\Big) = 1 - \alpha$ 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 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.

K testování hypotéz viz "Poznámky k testování hypotéz".

Příklad 11.1 (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* $s^2_{\boldsymbol{x}} = 2.56\ (\ell/100\,\text{km})^2$*.*

(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

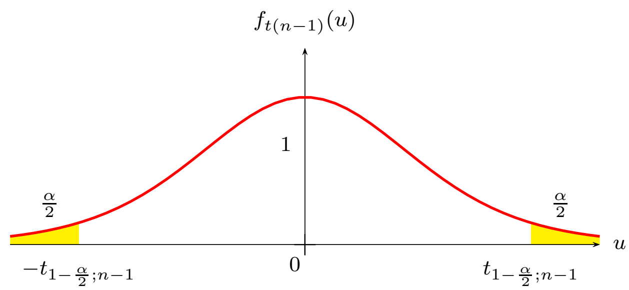
- veličina $\overline{\boldsymbol{X}} = \frac{1}{n} \sum_{i=1}^{n} X_i$ je *výběrový průměr* a
- veličina $S^2_{\boldsymbol{X}} = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{\boldsymbol{X}})^2$ je *výběrový rozptyl.*

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

$$|t| = 1.75 \ngtr 2.011 \doteq t_{0.975;48} = t_{1-\frac{\alpha}{2};n-1} \ ,$$

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

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

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

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

proti alternativní hypotéze:

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

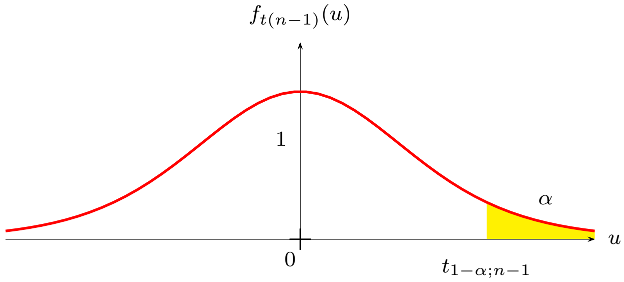
Pomocí testovací statistiky:

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

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

(Obecně tu máme jen nerovnost (pro případy $\mu < \mu_0$). Pro případ $\mu = \mu_0$ nastává rovnost.)

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 11.2 (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 (pomocí intervalu spolehlivosti) na hladině 5%, zda skutečná střední hodnota počtu pacientů za den může být považována za rovnu 25.

Řešení:

Máme veličiny

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

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. Když si teď připustíme, že X_i by mohly mít Poissonovo rozdělení, tak i relativně malé hodnoty n (my máme $n = 64$) mohou být dostatečné pro použití asymptotiky.

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.)

Poznámky k testu nezávislosti: Máme veličiny

- X s (různými) hodnotami $\{a_1, \ldots, a_k\}$ a
- Y s (různými) hodnotami $\{b_1, \ldots, b_\ell\}$

a chceme otestovat (na hladině α), hypotézu

$\mathbf{H}_0$: rozdělení veličin X a Y jsou *nezávislá*

proti alternativní hypotéze:

$\mathbf{H}_A$: rozdělení veličin X a Y jsou *závislá*

Při n pokusech s náhodným vektorem (X, Y) si pro $i = 1, \ldots, k$ označme veličiny

$$N_{ij} = \text{“počet výskytů případu } (X,Y) = (a_i, b_j) \text{ při } n \text{ pokusech”} \; .$$

a které tvoří náhodný vektor

$$\boldsymbol{N} = (N_{11}, \ldots, N_{k\ell})$$

s multinomickým rozdělením. Marginální rozdělení jednotlivých veličin $N_{i,j}$ jsou binomická a *za předpokladu nezávislosti* X *a* Y mají střední hodnotu

$$E(N_{ij}) = n \cdot P(X = a_i, Y = b_j) \overset{\text{(nezáv.)}}{=} n \cdot P(X = a_i) \cdot P(Y = b_j) \; .$$

Tyto střední hodnoty představují teoretické (tj. ideální očekávané) četnosti. Ovšem pravděpodobnosti $P(X = a_i)$ a $P(Y = b_j)$ nemáme v hypotéze uvedeny. Proto je odhadneme jako

$$P(X = a_i) \doteq \frac{n_{i,\bullet}}{n} \quad \text{a} \quad P(Y = b_j) \doteq \frac{n_{\bullet,j}}{n}$$

kde $n_{i,\bullet}$ a $n_{\bullet,j}$ jsou naměřené hodnoty veličin

$$N_{i,\bullet} = \sum_{j=1}^{\ell} N_{ij} = \text{“počet výskytů případu } X = a_i \text{ při } n \text{ pokusech”}$$

$$N_{\bullet,j} = \sum_{i=1}^{k} N_{ij} = \text{“počet výskytů případu } Y = b_j \text{ při } n \text{ pokusech”}$$

což jsou tzv. *marginální četnosti*. Celkově tedy hodnotu $E(N_{ij})$ odhadujeme naměřenou hodnotou veličiny

$$\widetilde{N}_{ij} = \frac{N_{i,\bullet} \cdot N_{\bullet,j}}{n} \; .$$

Testovací veličinu pak volíme jako:

$$T = \sum_{i=1}^{k} \sum_{j=1}^{\ell} \frac{\left(N_{ij} - \widetilde{N}_{ij}\right)^2}{\widetilde{N}_{ij}}$$

která má asymptoticky (tj. pro $n \to \infty$) χ^2-rozdělení s $(k-1) \cdot (\ell - 1)$ stupni volnosti. Pro praktické použití této asymptotiky se obvykle požaduje, aby platilo, že

$$\frac{n_{i,\bullet} \cdot n_{\bullet,j}}{n} \geq 5 \;\; \text{pro všechna} \;\; i = 1, \ldots, k \text{ a } j = 1, \ldots, \ell.$$

Kritérium pro **ZAMÍTNUTÍ** $\mathbf{H}_0$ (na hladině α) volíme podobně jako u testu dobré shody a sice

$$t > \chi^2_{1-\alpha;\, (k-1)\cdot(\ell-1)} \quad \Leftrightarrow \quad \text{zamítáme } \mathbf{H}_0 \text{ (na hladině } \alpha) \; .$$

Příklad 11.3 *Na $n = 100$ osobách byla pozorována barva očí a vlasů. Naměřeny byly následující sdružené četnosti:*

Oči \ *Vlasy*	*tmavé*	*světlé*
modré	10	20
šedé	10	10
hnědé	40	10

Jsou barvy očí a vlasů nezávislé? Testujte na hladině 5%.

Řešení:

Označme si veličiny

X= *"barva očí daného člověka"*

Y= *"barva vlasů daného člověka"*

a dál budeme pracovat s náhodným vektorem (X, Y), tj. u daného člověka budeme zjišťovat barvu očí a barvu vlasů.

Budeme testovat hypotézu:

$\mathbf{H}_0$: rozdělení veličin X a Y jsou *nezávislá*

proti alternativní hypotéze:

$\mathbf{H}_1$: rozdělení veličin X a Y jsou *závislá*.

na hladině významnosti $\alpha = 5\%$.

Četnost případu $(X, Y) = (i, j)$ v tabulce označme jako $n_{i,j}$ a marginální četnosti pak budou

$n_{i,\bullet} = \sum_j n_{i,j}$ pro případ $X = i$

$n_{\bullet,j} = \sum_i n_{i,j}$ pro případ $Y = j$.

což jsou součty v řádcích a sloupcích tabulky:

n_{ij} $(X =)\ i$ \ $(Y =)\ j$	tmavé	světlé	$n_{i,\bullet}$
modré	10	20	30
šedé	10	10	20
hnědé	40	10	50
$n_{\bullet,j}$	60	40	

Za předpokladu $\mathbf{H}_0$ pak jako teoretické četnosti budeme chápat hodnoty $\frac{n_{i,\bullet} \cdot n_{\bullet,j}}{n}$ v této tabulce:

$\widetilde{n}_{ij} = \frac{n_{i,\bullet} \cdot n_{\bullet,j}}{n}$ $(X =)\ i$ \ $(Y =)\ j$	tmavé	světlé	$n_{i,\bullet}$
modré	$\frac{30 \cdot 60}{100} = 18$	$\frac{30 \cdot 40}{100} = 12$	30
šedé	$\frac{20 \cdot 60}{100} = 12$	$\frac{20 \cdot 40}{100} = 8$	20
hnědé	$\frac{50 \cdot 60}{100} = 30$	$\frac{50 \cdot 40}{100} = 20$	50
$n_{\bullet,j}$	60	40	

Podmínka na tyto teoretické (tj. očekávané) četnosti ≥ 5 je splněna, takže test nezávislosti můžeme použít. Pro hodnotu testovací statistiky dostaneme

$$t = \sum_{i,j} \frac{\left(n_{ij} - \widetilde{n}_{ij}\right)^2}{\widetilde{n}_{ij}} =$$

$$= \frac{(10-18)^2}{18} + \frac{(10-12)^2}{12} + \frac{(40-30)^2}{30} + \frac{(20-12)^2}{12} + \frac{(10-8)^2}{8} + \frac{(10-20)^2}{20} = 18 + \frac{1}{18} \doteq 18.056 \ .$$

Tuto hodnotu dále porovnáme s hodnotou kvantilu χ^2 pro $(k-1)(\ell-1)$ stupňů volnosti, kde k je počet položek veličiny X a ℓ je počet položek veličiny Y. Tento počet je nyní jiný, než by byl u "obvyklého" testu dobré shody s $k \cdot \ell$ položkami, protože data jsme použili k odhadu marginálních pravděpodobností.

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

$$t > \chi^2_{1-\alpha;\ (k-1)(\ell-1)} \quad \Leftrightarrow \quad \text{zamítáme } \mathbf{H}_0 \text{ (na hladině } \alpha) \ .$$

Hledaný kvantil je

$$\chi^2_{1-\alpha;\ (3-1)(2-1)} = \chi^2_{0.95;\ 2} \doteq 5.992 \ .$$

Protože

$$t \doteq 18.056 > 5.992 \doteq \chi^2_{0.95;\ 2} \ ,$$

hypotézu o nezávislosti **ZAMÍTÁME**.

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:

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

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

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