MV011 Statistika I

9. Testování statistických hypotéz

Jan Koláček (kolacek@math.muni.cz)

Ústav matematiky a statistiky, Přírodovědecká fakulta, Masarykova univerzita, Brno

Motivační příklad

Příklad 1

Házíme opakovaně mincí. Ze 100 náhodných pokusů: 56 × „hlava“ a 44 × „orel“.
Otázka *: Je důvod se domnívat, že hlava nepadá stejně často jako orel?*

Realizace $\mathbf{x} = (x_1, \ldots, x_{100}) = (1, 1, 0, 1, 0, \ldots, 0, 1)$ náhodného výběru z alternativního rozdělení $A(\theta)$, $\theta \in (0, 1)$ je pravděpodobnost „úspěchu“ (hlava).
Cíl: Na základě $\mathbf{x}$ sestrojit test, který odpoví na otázku.

Řešení. Testujeme **nulovou** hypotézu $H_0: \ \theta = 0{,}5$ proti **alternativní** hypotéze $H_1: \ \theta \neq 0{,}5$.
Např. sestrojíme 95 % interval spolehlivosti pro θ (viz minulá přednáška)

$$\langle D, H\rangle = \overline{X} \pm u_{0{,}975}\sqrt{\frac{\overline{X}(1-\overline{X})}{n}} = 0{,}56 \pm 1{,}96\frac{\sqrt{0{,}56 \cdot 0{,}44}}{10} = \langle 0{,}463; 0{,}657\rangle.$$

Závěr: $0{,}5 \in \langle 0{,}463; 0{,}657\rangle \Rightarrow H_0$ **nezamítáme** (není důvod se domnívat ...)

Testování statistických hypotéz

Mějme náhodný výběr $\mathbf{X} = (X_1, \ldots, X_n)'$ rozsahu n z rozdělení o distribuční funkci $F(x; \boldsymbol{\theta})$, kde $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)' \in \boldsymbol{\Theta} \subset \mathbb{R}^m$. Množina $\boldsymbol{\Theta}$ nechť je neprázdná a otevřená.

Předpokládejme, že o parametru $\boldsymbol{\theta}$ existují dvě konkurující si hypotézy:

H_0: $\boldsymbol{\theta} \in \boldsymbol{\Theta}_0 \subset \boldsymbol{\Theta}$

H_1: $\boldsymbol{\theta} \in \boldsymbol{\Theta}_1 = \boldsymbol{\Theta} - \boldsymbol{\Theta}_0$

Tvrzení $\boxed{H_0}$ se nazývá **nulovou** **hypotézou** (**null hypothesis**).

Tvrzení $\boxed{H_1}$ se nazývá **alternativní** **hypotézou** (**alternative hypothesis**).

O platnosti této hypotézy se má rozhodnout na základě náhodného výběru $\mathbf{X} = (X_1, \ldots, X_n)'$, a to tak, že $\nearrow$ **zamítneme** (**reject**) *nebo* $\searrow$ **nezamítneme**(**accept**) platnost hypotézy H_0.

Testování statistických hypotéz

Na testování použijeme statistiku $T_n = T(\mathbf{X})$, kterou nazýváme **testovací statistikou**. Množinu hodnot, které může testovací statistika nabýt, rozdělíme na dvě disjunktní oblasti. Jednu označíme W_α , a nazveme ji **kritickou oblastí** (nebo také *oblastí zamítnutí hypotézy* (**region of rejection, critical region**)) a druhá je doplňkovou oblastí (*oblast nezamítnutí testované hypotézy*).

Na základě realizace náhodného výběru $\mathbf{x} = (x_1, \dots, x_n)'$ vypočítáme hodnotu testovací statistiky $t_n = T(\mathbf{x})$.

- Pokud hodnota testovací statistiky t_n nabude hodnoty z kritické oblasti, tj. $\boxed{t_n = T(\mathbf{x}) \in W_\alpha}$, pak **nulovou hypotézu zamítáme**.
- Pokud hodnota testovací statistiky nabude hodnoty z oblasti nezamítnutí, tj. $\boxed{t_n = T(\mathbf{x}) \notin W_\alpha}$, tak **nulovou hypotézu nezamítáme**.

Ad Příklad 1: $T_n = \frac{\overline{X}-0{,}5}{\sqrt{\overline{X}(1-\overline{X})}}\sqrt{n} = \frac{0{,}06}{\sqrt{0{,}56\cdot 0{,}44}}10 = 1{,}2087$

$W_\alpha = (-\infty, -u_{0{,}975}) \cup (u_{0{,}975}, \infty) = (-\infty; -1{,}96) \cup (1{,}96; \infty)$

$1{,}2087 \notin W_\alpha \Rightarrow H_0$ nezamítáme.

Testování statistických hypotéz

H_0	PLATÍ	NEPLATÍ
ZAMÍTÁME $t_n = T(\mathbf{x}) \in W_\alpha$	**chyba 1. druhu** (**type I error**) (α_0 je **hladina testu**) $\alpha_0 = \sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}_0} P_{\boldsymbol{\theta}}(T(\mathbf{X}) \in W_\alpha \mid H_0) \leq \alpha$	O.K. (tzv. **síla testu** (**power of a test**) či **silofunkce**) $1-\beta(\boldsymbol{\theta}) = P_{\boldsymbol{\theta}}(T(\mathbf{X}) \in W_\alpha \mid H_1)$ pro $\boldsymbol{\theta} \in \boldsymbol{\Theta}_1$
NEZAMÍTÁME $t_n = T(\mathbf{x}) \notin W_\alpha$	O.K.	**chyba 2. druhu** (**type II error**) $\beta(\boldsymbol{\theta}) = P_{\boldsymbol{\theta}}(T(\mathbf{X}) \notin W_\alpha \mid H_1)$ pro $\boldsymbol{\theta} \in \boldsymbol{\Theta}_1$

Chyby

Definice 1

Chybu, která spočívá **v nesprávném zamítnutí nulové hypotézy, i když je správná**, budeme nazývat **chybou prvého druhu**, pravděpodobnost

$$\alpha_0 = \sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}_0} P_{\boldsymbol{\theta}}(T(\mathbf{X}) \in W_\alpha | H_0)$$

nazveme **hladinou významnosti** (též **hladinou testu** (significance level of a test).

Chybu, která spočívá **v nesprávném přijetí nulové hypotézy, i když neplatí**, budeme nazývat **chybou druhého druhu** a její pravděpodobnost pro $\forall \boldsymbol{\theta} \in \boldsymbol{\Theta}_1$ označíme

$$\beta(\boldsymbol{\theta}) = P_{\boldsymbol{\theta}}(T(\mathbf{X}) \notin W_\alpha | H_1) \, .$$

Pravděpodobnost $1 - \beta(\boldsymbol{\theta})$ nazýváme **silou testu** (též **silou kritické oblasti** W_α) a jakožto funkci $\boldsymbol{\theta} \in \boldsymbol{\Theta}_1$ ji také nazveme **silofunkcí testu**.

Vztah mezi testy a intervalovými odhady

Mějme náhodný výběr $\mathbf{X} = (X_1, \ldots, X_n)'$ rozsahu n z rozdělení, které závisí na parametru $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)' \in \boldsymbol{\Theta}$ a parametrickou funkci $\gamma(\boldsymbol{\theta})$.

A Hypotéza $H_0 : \gamma(\boldsymbol{\theta}) = \gamma(\boldsymbol{\theta}_0)$ proti (tzv. *oboustranné*) alternativě $H_1 : \gamma(\boldsymbol{\theta}) \neq \gamma(\boldsymbol{\theta}_0)$:

Mějme **intervalový odhad** $(D_n(\mathbf{X}), H_n(\mathbf{X}))$ parametrické funkce $\gamma(\boldsymbol{\theta})$ o spolehlivosti $1 - \alpha$. Pokud platí nulová hypotéza, pak

$$1 - \alpha = P_{\boldsymbol{\theta}} \left(D_n(\mathbf{X}) \leq \gamma(\boldsymbol{\theta}_0) \leq H_n(\mathbf{X}) \right),$$

takže **kritický obor** tohoto testu má tvar:

$$W_\alpha = \{ \mathbf{X} \in \mathbb{R}^n : \gamma(\boldsymbol{\theta}_0) \notin (D_n(\mathbf{X}), H_n(\mathbf{X})) \}$$

.

Vztah mezi testy a intervalovými odhady

Zjistíme-li v konkrétní situaci, že

$$\gamma(\boldsymbol{\theta}_0) \notin (d_n(\mathbf{x}), h_n(\mathbf{x}))$$

tj. realizace $\mathbf{x} \in W_\alpha$

potom

- buď nastal jev, který má pravděpodobnost α (volí se blízká nule),
- nebo neplatí nulová hypotéza.

Protože při obvyklé volbě $\alpha = 0.05$ nebo $\alpha = 0.01$ je tento jev „prakticky nemožný", proto nulovou hypotézu H_0 **zamítáme ve prospěch alternativy** H_1.

V opačném případě, tj. pokud

$$\gamma(\boldsymbol{\theta}_0) \in (d_n(\mathbf{x}), h_n(\mathbf{x}))$$

tj. realizace $\mathbf{x} \notin W_\alpha$

nulovou hypotézu H_0 **nezamítáme**.

Vztah mezi testy a intervalovými odhady

B Hypotéza $H_0 : \gamma(\boldsymbol{\theta}) = \gamma(\boldsymbol{\theta}_0)$ proti (tzv. *jednostranné*) alternativě $H_1 : \gamma(\boldsymbol{\theta}) > \gamma(\boldsymbol{\theta}_0)$:

V tomto případě využijeme **dolní odhad** $D_n(\mathbf{X})$ parametrické funkce $\gamma(\boldsymbol{\theta})$ o spolehlivosti $1-\alpha$. Pokud platí nulová hypotéza, pak

$$1-\alpha = P_{\boldsymbol{\theta}}\left(D_n(\mathbf{X}) \leq \gamma(\boldsymbol{\theta}_0)\right),$$

takže **kritický obor** tohoto testu má tvar:

$$W_\alpha = \{\mathbf{X} \in \mathbb{R}^n : D_n(\mathbf{X}) > \gamma(\boldsymbol{\theta}_0)\}.$$

Vztah mezi testy a intervalovými odhady

C Hypotéza $H_0 : \gamma(\boldsymbol{\theta}) = \gamma(\boldsymbol{\theta}_0)$ proti (tzv. *jednostranné*) alternativě $H_1 : \gamma(\boldsymbol{\theta}) < \gamma(\boldsymbol{\theta}_0)$ V tomto případě využijeme **horní odhad** $H_n(\mathbf{X})$ parametrické funkce $\gamma(\boldsymbol{\theta})$ o spolehlivosti $1-\alpha$. Pokud platí nulová hypotéza, pak

$$1-\alpha = P_{\boldsymbol{\theta}}\left(\gamma(\boldsymbol{\theta}_0) \leq H_n(\mathbf{X})\right),$$

takže **kritický obor** tohoto testu má tvar:

$$W_\alpha = \{\mathbf{X} \in \mathbb{R}^n : H_n(\mathbf{X}) < \gamma(\boldsymbol{\theta}_0)\}.$$

Příklad

Příklad 2 (Rychlost letadla)

Rychlost letadla byla určována v 5 nezávislých zkouškách. Výsledky (v m/s) jsou uvedeny v následující tabulce:

1	2	3	4	5
867,2	868,3	873,6	870,5	871,9

Z dlouhodobých výzkumů víme, že rychlost letadla se řídí normálním rozdělením se směrodatnou odchylkou $\sigma = 2,1\ m/s$. Na hladině významnosti $\alpha = 0,05$ testujte hypotézu, že střední hodnota rychlosti letadla je 872 m/s.

Řešení. Jedná se o realizaci náhodného výběru z $N(\mu_0; 2,1^2)$.
Testujeme hypotézu $H_0 : \mu_0 = 872$ proti alternativě $H_1 : \mu_0 \neq 872$.

Řešení

(I) Testování nulové hypotézy pomocí pivotové statistiky $U_{\overline{X}}$ a kritické hodnoty.

Protože kritický obor W_0 lze ekvivalentně vyjádřit i takto

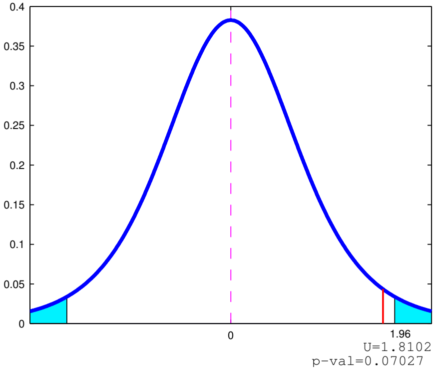
$$\begin{aligned} W_0 &= \left\{ \mathbf{x} \in \mathbb{R}^n : \mu_0 < \bar{x} - \tfrac{\sigma}{\sqrt{n}} u_{1-\frac{\alpha}{2}} \vee \mu_0 > \bar{x} + \tfrac{\sigma}{\sqrt{n}} u_{1-\frac{\alpha}{2}} \right\} \\ &= \left\{ \mathbf{x} \in \mathbb{R}^n : |\bar{x} - \mu_0| > \tfrac{\sigma}{\sqrt{n}} u_{1-\frac{\alpha}{2}} \right\} \\ &= \left\{ \mathbf{x} \in \mathbb{R}^n : |u_{\bar{x}}| = \frac{|\bar{x} - \mu_0|}{\sigma} \sqrt{n} > u_{1-\frac{\alpha}{2}} \right\} \end{aligned}$$

počítejme $|u_{\bar{x}}| = \frac{|870{,}3-872|}{2{,}1}\sqrt{5} = 1,81$. Protože $|u_{\bar{x}}| = 1,81$ nepřekračuje

kritickou hodnotu $u_{1-\frac{\alpha}{2}} = u_{0{,}975} = 1,96$, nulovou hypotézu na 5% hladině **nezamítneme**.

Řešení

(II) Testování nulové hypotézy pomocí p-hodnoty



Dosažená hladina odpovídající testové statistice (tj. tzv. p-hodnota, **P-value, significance value**), což je nejmenší hladina testu, při které bychom ještě hypotézu H_0 zamítli, je rovna $0,0702$, takže například při $\alpha = 10\%$ by již dosažený výsledek byl statisticky významný.

Protože p-hodnota je větší než zvolená hladina významnosti $\alpha = 0,05$, hypotézu **nezamítáme**.

Řešení

(III) Testování nulové hypotézy pomocí intervalu spolehlivosti $\langle D, H \rangle$

Protože jde o oboustranný test, použijeme **interval spolehlivosti** pro střední hodnotu μ při známém σ

$$\langle d, h \rangle = \bar{x} \pm \frac{\sigma}{\sqrt{n}} u_{1-\frac{\alpha}{2}} = 180,3 - \frac{2{,}1}{\sqrt{5}} 1,96 = \langle 868,46; 872,14 \rangle$$

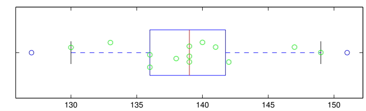
Protože interval spolehlivosti $\langle 868,46; 872,14 \rangle$ pokrývá hodnotu 872, proto nulovou hypotézu na hladině významnosti $\alpha = 0,05$ **nezamítáme**.

Příklad

Příklad 3 (Výška desetiletých chlapců)

V roce 1961 byla u 15 náhodně vybraných chlapců z populace všech desetiletých chlapců žijících v Československu zjištěna výška:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
130	140	136	141	139	133	149	151	139	136	138	142	127	139	147



Je známo, že každá následující generace je v průměru o něco vyšší než generace předcházející. Můžeme se tedy ptát, zda průměr $\bar{x} = 139.133$ *zjištěný v náhodném výběru rozsahu* $n = 15$ *znamená, že na 5% hladině máme zamítnout nulovou hypotézu* $H_0 : \mu = 136,1$ *(zjištění z roku 1951) ve prospěch alternativní hypotézy* $H_1 : \mu > 136,1$.

Rozptyl $\sigma^2 = 6,4^2$ *cm*^{2}*, zjištěný v roce 1951 (kdy se provádělo rozsáhlé šetření), můžeme považovat za známý, neboť variabilita výšek zůstává (na rozdíl od střední výšky) téměř nezměněná.*

Řešení

Řešení (I) Testování nulové hypotézy pomocí pivotové statistiky $U_{\overline{X}}$ a kritické hodnoty.

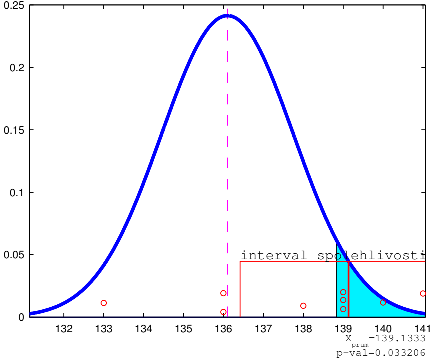
Protože kritický obor W_0 lze ekvivalentně vyjádřit i takto

$$W_0 = \left\{ \mathbf{x} \in \mathbb{R}^n : \bar{x} - \tfrac{\sigma}{\sqrt{n}} u_{1-\alpha} > \mu_0 \right\} = \left\{ \mathbf{x} \in \mathbb{R}^n : u_{\bar{x}} = \tfrac{\bar{x}-\mu_0}{\sigma} \sqrt{n} > u_{1-\alpha} \right\},$$

počítejme $u_{\bar{x}} = \frac{139{,}133-136{,}1}{6{,}4}\sqrt{15} = 1{,}835$. Protože $u_{\bar{x}} = 1{,}835$ překračuje kritickou hodnotu $u_{1-\alpha} = u_{0{,}95} = 1{,}645$, nulovou hypotézu na 5% hladině **zamítneme** ve prospěch alternativní hypotézy, že se střední výška desetiletých hochů zvětšila.

Řešení

(II) Testování nulové hypotézy pomocí p-hodnoty



Dosažená hladina odpovídající testové statistice (tj. tzv. p-hodnota, **P-value, significance value**), což je nejmenší hladina testu, při které bychom ještě hypotézu H_0 zamítli, je rovna $0,033$, takže například při $\alpha = 2,5\%$ by již dosažený výsledek nebyl statisticky významný.

Protože p-hodnota je menší než zvolená hladina významnosti $\alpha = 0.05$, hypotézu **zamítáme**.

Řešení

(III) Testování nulové hypotézy pomocí intervalu spolehlivosti $\langle D, +\infty)$

Protože jde o jednostranný test, použijeme **dolní odhad** střední hodnoty μ

$$d = \bar{x} - \frac{\sigma}{\sqrt{n}} u_{1-\alpha} = 139,133 - \frac{6,4}{\sqrt{15}} 1,645 = 136,415$$

Protože interval spolehlivosti $\langle 136,415, +\infty)$ nepokrývá hodnotu $136,1$, proto nulovou hypotézu na hladině významnosti $\alpha = 0,05$ **zamítáme**.

Příklad

Příklad 4 (Spotřeba auta)

Spotřeba téhož auta byla testována u 11 řidičů s výsledky 8,8; 8,9; 9,0; 8,7; 9,3; 9,0; 8,7; 8,8; 9,4; 8,6; 8,9 ($l/100\,km$).

a) *Můžeme na hladině významnosti 0,05 zamítnout hypotézu, že je pravdivá výrobcem udávaná spotřeba 8,8 $l/100\,km$?*

b) *Můžeme na stejné hladině významnosti popřít tvrzení, že rozptyl spotřeby je 0,1?*

Řešení.

a) Jedná se o realizaci náhodného výběru z $N(\mu_0; \sigma_0^2)$.

Testujeme hypotézu $H_0 : \mu_0 = 8{,}8$ proti alternativě $H_1 : \mu_0 \neq 8{,}8$.

Řešení

(I) Testování nulové hypotézy pomocí pivotové statistiky T a kritické hodnoty.

Protože kritický obor W_0 lze ekvivalentně vyjádřit i takto

$$
\begin{aligned}
W_0 &= \left\{ \mathbf{x} \in \mathbb{R}^n : \mu_0 < \bar{x} - \tfrac{s}{\sqrt{n}} t_{1-\frac{\alpha}{2}}(n-1) \vee \mu_0 > \bar{x} + \tfrac{s}{\sqrt{n}} t_{1-\frac{\alpha}{2}}(n-1) \right\} \\
&= \left\{ \mathbf{x} \in \mathbb{R}^n : |\bar{x} - \mu_0| > \tfrac{s}{\sqrt{n}} t_{1-\frac{\alpha}{2}}(n-1) \right\} \\
&= \left\{ \mathbf{x} \in \mathbb{R}^n : |t| = \frac{|\bar{x} - \mu_0|}{s} \sqrt{n} > t_{1-\frac{\alpha}{2}}(n-1) \right\}
\end{aligned}
$$

počítejme $|t| = \frac{8{,}918-8{,}8}{0{,}248}\sqrt{11} = 1,5788$. Protože $|t| = 1,5788$ nepřekračuje

kritickou hodnotu $t_{1-\frac{\alpha}{2}}(n-1) = t_{0{,}975}(10) = 2,228$, nulovou hypotézu na 5% hladině **nezamítneme**.

Řešení

(II) TESTOVÁNÍ NULOVÉ HYPOTÉZY POMOCÍ p-HODNOTY

Dosažená hladina odpovídající testové statistice (tj. tzv. p-hodnota, **P-value, significance value**), což je nejmenší hladina testu, při které bychom ještě hypotézu H_0 zamítli, je rovna $0,1455$.

Protože p-hodnota je větší než zvolená hladina významnosti $\alpha = 0,05$, hypotézu **nezamítáme**.

Řešení

(III) Testování nulové hypotézy pomocí intervalu spolehlivosti $\langle D, H\rangle$

Protože jde o oboustranný test, použijeme **interval spolehlivosti** pro střední hodnotu μ při neznámém σ

$$\langle d,h\rangle = \bar{x} \pm \tfrac{s}{\sqrt{n}} t_{1-\frac{\alpha}{2}}(n-1) = 8{,}918 - \tfrac{0{,}248}{\sqrt{11}} 2{,}228 = \langle 8{,}751; 9{,}085\rangle$$

Protože interval spolehlivosti $\langle 8{,}751; 9{,}085\rangle$ pokrývá hodnotu $8{,}8$, proto nulovou hypotézu na hladině významnosti $\alpha = 0{,}05$ **nezamítáme**.

Řešení

b) Testujeme hypotézu $H_0: \sigma_0^2 = 0,1$ proti alternativě $H_1: \sigma_0^2 \neq 0,1$.
(I) Testování nulové hypotézy pomocí pivotové statistiky K a kritické hodnoty.

Protože kritický obor W_0 lze ekvivalentně vyjádřit i takto

$$W_0 = \left\{ \mathbf{x} \in \mathbb{R}^n : \sigma_0^2 < \frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)} \vee \sigma_0^2 > \frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}(n-1)} \right\}$$
$$= \left\{ \mathbf{x} \in \mathbb{R}^n : k = \frac{(n-1)s^2}{\sigma_0^2} > \chi^2_{1-\frac{\alpha}{2}}(n-1) \vee k < \chi^2_{\frac{\alpha}{2}}(n-1) \right\}$$

počítejme $k = \frac{10 \cdot 0,248^2}{0,1} = 6,164$. Protože $k = 6,164$ nepřekračuje kritické hodnoty

$\chi^2_{0,025}(10) = 3,247$ ani $\chi^2_{0,975}(10) = 20,483$, nulovou hypotézu na 5% hladině **nezamítneme**.

Řešení

(II) TESTOVÁNÍ NULOVÉ HYPOTÉZY POMOCÍ p-HODNOTY

Dosažená hladina odpovídající testové statistice (tj. tzv. p-hodnota, **P-value, significance value**), což je nejmenší hladina testu, při které bychom ještě hypotézu H_0 zamítli, je rovna $0,397$.

Protože p-hodnota je větší než zvolená hladina významnosti $\alpha = 0,05$, hypotézu **nezamítáme**.

Řešení

(III) TESTOVÁNÍ NULOVÉ HYPOTÉZY POMOCÍ INTERVALU SPOLEHLIVOSTI $\langle D, H\rangle$

Protože jde o oboustranný test, použijeme **interval spolehlivosti** pro σ^2

$$\langle d, h\rangle = \left\langle \frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}; \frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}(n-1)} \right\rangle = \langle 0.03009; 0,1898\rangle$$

Protože interval spolehlivosti $\langle 0.03009; 0,1898\rangle$ pokrývá hodnotu $0,1$, proto nulovou hypotézu na hladině významnosti $\alpha = 0,05$ **nezamítáme**.

Testy o parametrech normálního rozdělení

H_0	H_1	Hypotézu H_0 zamítáme, pokud $\mathbf{X} \in W_\alpha$, tj.	Předpoklady
$\mu = \mu_0$	$\mu \neq \mu_0$	$\lvert\bar{X} - \mu_0\rvert\sqrt{n} \geq \sigma u_{1-\frac{\alpha}{2}}$	σ^2 známé
$\mu = \mu_0$	$\mu > \mu_0$	$(\bar{X} - \mu_0)\sqrt{n} \geq \sigma u_{1-\alpha}$	σ^2 známé
$\mu = \mu_0$	$\mu < \mu_0$	$(\bar{X} - \mu_0)\sqrt{n} \leq -\sigma u_{1-\alpha}$	σ^2 známé
$\mu = \mu_0$	$\mu \neq \mu_0$	$\lvert\bar{X} - \mu_0\rvert\sqrt{n} \geq S t_{1-\frac{\alpha}{2}}(n-1)$	σ^2 neznámé
$\mu = \mu_0$	$\mu > \mu_0$	$(\bar{X} - \mu_0)\sqrt{n} \geq S t_{1-\alpha}(n-1)$	σ^2 neznámé
$\mu = \mu_0$	$\mu < \mu_0$	$(\bar{X} - \mu_0)\sqrt{n} \leq -S t_{1-\alpha}(n-1)$	σ^2 neznámé
$\sigma^2 = \sigma_0^2$	$\sigma^2 \neq \sigma_0^2$	$\frac{(n-1)S^2}{\sigma_0^2} \notin \left(\chi^2_{\frac{\alpha}{2}}(n-1), \chi^2_{1-\frac{\alpha}{2}}(n-1)\right)$	μ neznámé
$\sigma^2 = \sigma_0^2$	$\sigma^2 > \sigma_0^2$	$\frac{(n-1)S^2}{\sigma_0^2} \geq \chi^2_{1-\alpha}(n-1)$	μ neznámé
$\sigma^2 = \sigma_0^2$	$\sigma^2 < \sigma_0^2$	$\frac{(n-1)S^2}{\sigma_0^2} \leq \chi^2_{\alpha}(n-1)$	μ neznámé

Testy dvou nezávislých výběrů

- první náhodný výběr $\perp\!\!\!\perp\{X_1,\dots,X_{n_1}\} \sim N(\mu_1,\sigma_1^2)$,
- druhý náhodný výběr $\perp\!\!\!\perp\{Y_1,\dots,Y_{n_2}\} \sim N(\mu_2,\sigma_2^2)$,
- označme

$$S_{12}^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2},$$

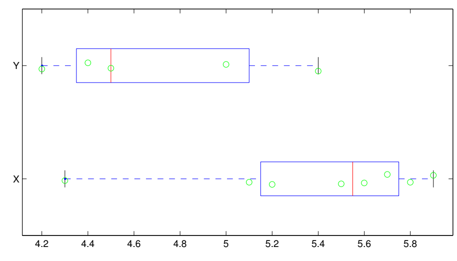
H_0	H_1	H_0 zamítáme, pokud $(\mathbf{X}',\mathbf{Y}')' \in W_\alpha$	Předpoklady
$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$	$\lvert\bar{X}-\bar{Y}\rvert \geq u_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$	σ_1^2, σ_2^2 známé
$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$	$\lvert\bar{X}-\bar{Y}\rvert \geq t_{1-\frac{\alpha}{2}}(n_1+n_2-2)\; S_{12}\sqrt{\frac{n_1+n_2}{n_1 n_2}}$	$\sigma_1^2 = \sigma_2^2$ neznámé
$\sigma_1^2 = \sigma_2^2$	$\sigma_1^2 \neq \sigma_2^2$	$\frac{S_1^2}{S_2^2} \notin \left(F_{\frac{\alpha}{2}}(n_1-1,n_2-1), F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)\right)$	μ_1, μ_2 neznámé

Příklad

Příklad 5 (Dva nezávislé náhodné výběry z normálního rozdělení při neznámých ale stejných rozptylech)

Bylo vybráno 13 polí stejné kvality. Na 8 z nich se zkoušel nový způsob hnojení, zbývajících 5 bylo ošetřeno běžným způsobem. Výnosy pšenice uvedené v tunách na hektar jsou označeny X_i u nového a Y_i u běžného způsobu hnojení. Je třeba zjistit, zda způsob hnojení má vliv na výnos pšenice.

X_i	5,7	5,5	4,3	5,9	5,2	5,6	5,8	5,1
Y_i	5,0	4,5	4,2	5,4	4,4			



Řešení

Řešení Budeme nejprve testovat hypotézu $H_0 : \frac{\sigma_1^2}{\sigma_2^2} = 1$ proti alternativě $H_1 : \frac{\sigma_1^2}{\sigma_2^2} \neq 1$. Za pivotovou statistiku zvolíme statistiku

$$F = \frac{S_1^2}{S_2^2}\frac{\sigma_2^2}{\sigma_1^2} \sim F(n_1 - 1, n_2 - 1).$$

(a) Můžeme například vypočítat statistiku F za platnosti nulové hypotézy a porovnat ji s příslušnými oboustrannými kvantily.

$$\begin{aligned}
&f = \tfrac{0{,}27}{0{,}24} = 1{,}1243\\
&F_{\frac{\alpha}{2}}(n_1-1, n_2-1) = 0{,}1811\\
&F_{1-\frac{\alpha}{2}}(n_1-1, n_2-1) = 9{,}0741
\end{aligned}$$

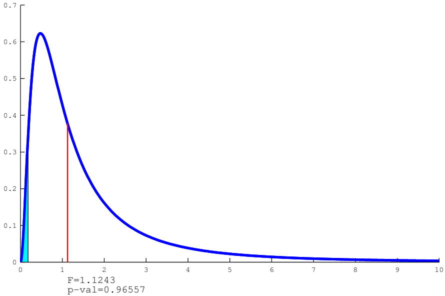
H_0 **nezamítáme**.

Řešení

(b) Další možností je spočítat p-hodnotu a srovnat se zvolenou hladinou testu α:

$$p\text{-value} = 0,9656 \gg 0,05$$

Protože p-hodnota je výrazně větší než zvolená hladina testu, hypotézu o rovnosti rozptylů proti alternativě nerovnosti **nezamítáme**.



Řešení

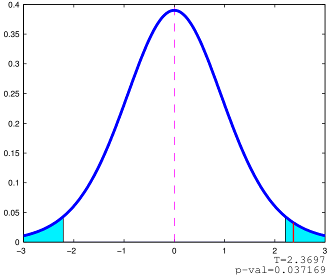
(c) A naposledy můžeme ještě zkonstruovat $100(1-\alpha)\%$ interval spolehlivosti pro podíl rozptylů $\frac{\sigma_1^2}{\sigma_2^2}$

$$\left\langle \frac{S_1^2}{S_2^2}\frac{1}{F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)}, \frac{S_1^2}{S_2^2}\frac{1}{F_{\frac{\alpha}{2}}(n_1-1,n_2-1)} \right\rangle.$$

a zjistit, zda pokrývá hodnotu 1. Protože dostáváme interval $\langle 0,1239; 6,2088\rangle$, který pokrývá jedničku, hypotézu nezamítáme.

Řešení

(I) Testování pomocí statistiky T a kritické hodnoty



Vypočítáme-li hodnotu statistiky

$$T_{\bar{X}-\bar{Y}} = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{S_{12}} \sqrt{\frac{n_1 n_2}{n_1 + n_2}}$$

a porovnáme s kvantilem Studentova rozdělení, tj.

$$t_{\bar{x}-\bar{y}} = 2{,}3697 > t_{1-\alpha/2}(11) = 2{,}201,$$

takže **hypotézu**

$$H_0 : \mu_1 - \mu_2 = 0$$

zamítáme.

Řešení

(II) TESTOVÁNÍ POMOCÍ p-HODNOTY

Vypočítáme-li p-hodnotu a porovnáme se zvolenou hladinou významnosti $\alpha = 0,05$

$$p = 2P(|T_{\bar{X}-\bar{Y}}| > t) = 2(1 - P(|T_{\bar{X}-\bar{Y}}| \leq t)) = 0,037169 < \alpha$$

takže **hypotézu**

$$H_0 : \mu_1 - \mu_2 = 0$$

zamítáme.

Řešení

(III) Testování nulové hypotézy pomocí intervalu spolehlivosti

$$
\begin{aligned}
\bar{x}-\bar{y}\pm t_{1-\frac{\alpha}{2}}(\nu)\, s_{12}\sqrt{\tfrac{n_1+n_2}{n_1 n_2}} &= \langle 0,6875 \pm 2,201 \cdot 0,5089/1,7541\rangle \\
&= \langle 0,048958; 1,326\rangle
\end{aligned}
$$

Protože interval spolehlivosti nepokrývá nulu, na dané hladině významnosti **hypotézu zamítáme** ve prospěch alternativy.

Příklad

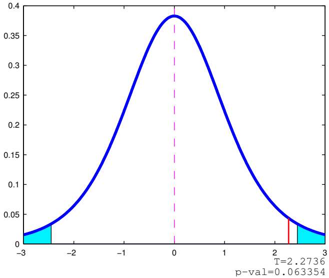
Příklad 6 (Párový test)

Na sedmi rostlinách byl posuzován vliv fungicidního přípravku podle počtu skvrn na listech před a týden po použití přípravku. Otestujte, zdali má přípravek vliv na počet skvrn na listech. Data udávající počet skvrn na listech před a po použití přípravku:

před použitím přípravku	X_1	9	17	31	7	8	20	10
po použití přípravku	X_2	10	11	18	6	7	17	5

Řešení

(I) Testování pomocí statistiky T a kritické hodnoty



Položme

$$Z = X_1 - X_2.$$

Vypočítáme-li hodnotu statistiky

$$T = \frac{\bar{Z}}{S/\sqrt{n}}$$

a porovnáme s kvantilem Studentova rozdělení, tj.

$$t = \frac{\bar{z}}{s/\sqrt{n}} = 2,2736 \not> t_{1-\alpha/2}(n-1) = 2,4469,$$

takže **hypotézu**

$$H_0 : \mu_1 - \mu_2 = 0$$

nezamítáme.

Řešení

(II) Testování pomocí p-hodnoty

Vypočítáme-li p-hodnotu a porovnáme se zvolenou hladinou významnosti $\alpha = 0,05$

$$p = 2P(|T| > t) = 2(1 - P(|T| \leq t)) = 0,06335 > \alpha$$

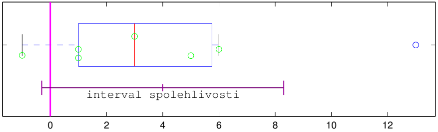
takže **hypotézu**

$$H_0 : \mu_1 - \mu_2 = 0$$

nezamítáme.

Řešení

(III) Testování nulové hypotézy pomocí intervalu spolehlivosti



$$\bar{z} \pm t_{1-\alpha/2}(n-1) \cdot s/\sqrt{n}$$
$$= 4 \pm 2,4469 \cdot 4,6547/2,6458 = [-0,30492; 8,3049]$$

Protože interval spolehlivosti pokrývá hodnotu $Z = 0$, na dané hladině významnosti **hypotézu nezamítáme**.
Shrneme-li předchozí výsledky slovně, pak nulovou hypotézu o tom, že

přípravek nemá vliv na počet skvrn

na hladině významnosti $\alpha = 0,05$ **nemůžeme zamítnout** oproti alternativě o jeho vlivu.

Asymptotické testy

Nechť $\perp\!\!\!\perp\{X_1,\ldots,X_n\} \simeq \mathcal{L}(\mu,\sigma^2)$ s konečnými druhými momenty (s výběrovým průměrem $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ a se $S_*^2 = S_*^2(\mathbf{X})$, což je (**slabě**) **konzistentní odhad** rozptylu σ^2):

H_0	H_1	H_0 zamítáme, pokud $\mathbf{X} \in W_\alpha$	Předpoklady
$\mu = \mu_0$	$\mu \neq \mu_0$	$\frac{\lvert\bar{X}-\mu_0\rvert}{S_*}\sqrt{n} \geq u_{1-\frac{\alpha}{2}}$	$0 < \sigma^2 < \infty$
$\mu = \mu_0$	$\mu \neq \mu_0$	$\frac{\lvert\bar{X}-\mu_0\rvert}{\sqrt{\mu_0}}\sqrt{n} \geq u_{1-\frac{\alpha}{2}}$	$\perp\!\!\!\perp\{X_1,\ldots,X_n\} \simeq Po(\mu)$
$p = p_0$	$p \neq p_0$	$\frac{\lvert\bar{X}-p_0\rvert}{\sqrt{p_0(1-p_0)}}\sqrt{n} \geq u_{1-\frac{\alpha}{2}}$	$\perp\!\!\!\perp\{X_1,\ldots,X_n\} \simeq A(p)$

Příklad

Příklad 7 (Volby)

Starosta obdržel při posledních volbách 60% hlasů. Bude stejně úspěšný i při příštích, když ze 100 náhodně vybraných občanů je pro něj 48?

Řešení Označme $X_i, \ i = 1, \ldots, 100$ náhodnou veličinu nabývající hodnoty 1, pokud i-tý volič dá hlas starostovi a hodnoty 0, pokud nedá. Zřejmě $X_i \sim A(p)$. Testujeme hypotézu $H_0 : \ p = 0,6$ proti alternativní hypotéze $H_1 : \ p \neq 0,6$. Vypočteme průměr $\bar{x} = \frac{48}{100} = 0,48$ a směrodatné odchylky $s_0 = \sqrt{p_0(1 - p_0)} = 0,4899$, $s = \sqrt{\bar{x}(1 - \bar{x})} = 0,4996$.

Řešení

(I) TESTOVÁNÍ NULOVÉ HYPOTÉZY POMOCÍ PIVOTOVÉ STATISTIKY $U_{\overline{X}}$ A KRITICKÉ HODNOTY.

Protože kritický obor W_0 lze ekvivalentně vyjádřit i takto

$$W_0 = \left\{ \mathbf{x} \in \mathbb{R}^n : u_{\bar{x}} = \left| \frac{\bar{x}-p}{\sqrt{p_0(1-p_0)}} \sqrt{n} \right| > u_{1-\frac{\alpha}{2}} \right\},$$

počítejme $u_{\bar{x}} = \frac{|0{,}48-0{,}6|}{0{,}4899}\sqrt{100} = 2,4495$. Protože $u_{\bar{x}} = 2,4495$ překračuje kritickou hodnotu $u_{1-\frac{\alpha}{2}} = u_{0{,}975} = 1,96$, nulovou hypotézu na 5% hladině **zamítáme**.

Řešení

(II) TESTOVÁNÍ POMOCÍ p-HODNOTY

Vypočítáme p-hodnotu a porovnáme se zvolenou hladinou významnosti $\alpha = 0,05$

$$p = 2P(|U_{\bar{x}}| > u_{\bar{x}}) = 2(1 - P(|U_{\bar{x}}| \leq u_{\bar{x}})) = 0,0143 < \alpha$$

takže **hypotézu**

$$H_0 : p = 0,6$$

zamítáme.

Řešení

(III) Testování nulové hypotézy pomocí intervalu spolehlivosti

Interval spolehlivosti pro p:

$$\bar{x} \pm u_{0,975} \frac{s}{\sqrt{n}} = 0,48 \pm 1,96 \cdot \frac{0,4996}{\sqrt{100}} = \langle 0,382; 0,5779 \rangle.$$

Protože interval spolehlivosti $\langle 0,382; 0,5779 \rangle$ nepokrývá hodnotu $0,6$, nulovou hypotézu na hladině významnosti $\alpha = 0,05$ **zamítáme**.

Úlohy k procvičení

Příklad 8.1

Na hladině významnosti $\alpha = 0,05$ *testujte hypotézu* $H_0: \ \sigma_0 = 300$ *o směrodatné odchylce normálně rozdělené náhodné veličiny, jestliže je zaznamenáno* $n = 25, \ \overline{X} = 3118, \ s = 357.$

[nezamítáme]

Příklad 8.2

Denní přírůstky váhy selat (v dkg) byly při krmení směsí A : 62, 54, 55, 60, 53, 58, u směsi B : 52, 56, 50, 49, 51. Je mezi nimi statisticky významný rozdíl?

[ano]

Úlohy k procvičení

Příklad 8.3

Pro bavlněnou přízi je předepsána horní mez variability pevnosti vlákna: rozptyl pevnosti (která má normální rozdělení) nemá překročit $\sigma_0^2 = 0{,}36$*. Při zkoušce 16 vzorků byly zjištěny výsledky 2,22, 3,54, 2,37, 1,66, 4,74, 4,82, 3,21, 5,44, 3,23, 4,79, 4,85, 4,05, 3,48, 3,89, 4,90, 5,37. Je důvod k podezření na vyšší nestejnoměrnost než je stanoveno?*

[ano]

Příklad 8.4

Bylo provedeno měření obsahu SiO_2 *ve strusce dvěma metodami*

analyticky	20,1	19,6	20,0	19,9	20,1	
fotokolorometricky	20,9	20,1	20,6	20,5	20,7	20,5

Je mezi rozptyly výsledků jednotlivých metod podstatný rozdíl?

[není]

Úlohy k procvičení

Příklad 8.5

Při 40 hodech mincí byl rub zaznamenán 22krát. Je důvod se domnívat, že rub nepadá stejně často jako líc? [ne]

Příklad 8.6

Na základě testu máme na 5% hladině významnosti rozhodnout, zda produkce vajec plemene kornyšek černých je nižší než plemene leghornek bílých. Náhodně jsme vybrali 50 kornyšek a 40 leghornek, u nichž byla zjištěna roční produkce vajec. Byl vypočten roční průměr produkce na slepici – kornyška 275, leghornka 280. Z dřívějška jsou známy rozptyly $\sigma^2_{kor} = 48$, $\sigma^2_{leg} = 41$.

[H_0 zamítáme, kornyšky mají horší produkci vajec než leghornky]