12. Dvouvýběrové testy o shodě parametrů dvou populací

<table>
<tr><th>Typ proměnné</th><th>Požadovaný typ analýzy</th><th colspan="2">Předpoklady</th><th>Testy, resp. intervalové odhady</th></tr>
<tr><td rowspan="8">Dvě nezávislé spojité proměnné</td><td rowspan="3">Ověření shody rozptylů (homoskedasticity)</td><td colspan="2" rowspan="2">Normalita</td><td>F-test (test shody rozptylů)</td></tr>
<tr><td>Intervalový odhad poměru rozptylů, resp. směr. odchylek</td></tr>
<tr><td colspan="2">—</td><td>Leveneho test</td></tr>
<tr><td rowspan="5">Ověření shody měr polohy (středních hodnot, resp. mediánů)</td><td rowspan="4">Normalita</td><td rowspan="2">Shoda rozptylů (homoskedasticita)</td><td>Dvouvýběrový Studentův t-test (test shody stř. hodnot)</td></tr>
<tr><td>Intervalový odhad rozdílu stř.hodnot</td></tr>
<tr><td rowspan="2">Různé rozptyly (heteroskedasticita)</td><td>Aspinové-Welchův test (test shody stř. hodnot)</td></tr>
<tr><td>Intervalový odhad rozdílu stř.hodnot</td></tr>
<tr><td colspan="2">—</td><td>Mannův-Whitneho test test shody mediánů</td></tr>
<tr><td rowspan="4">Párová (spojitá) data</td><td rowspan="4">Ověření shody úrovně párových dat</td><td colspan="2" rowspan="2">Normalita</td><td>Párový studentův t-test</td></tr>
<tr><td>Intervalový odhad střední hodnoty rozdílů</td></tr>
<tr><td colspan="2">Výběry většího rozsahu</td><td>Párový znaménkový test</td></tr>
<tr><td colspan="2">Symetrické rozdělení</td><td>Wilcoxonův párový test</td></tr>
<tr><td rowspan="2">Dvě dichotomické proměnné</td><td rowspan="2">Ověření shody pravděpodobností</td><td colspan="2" rowspan="2">$n_i > \frac{9}{p_i(1-p_i)}, i = 1, 2$</td><td>Test homogenity dvou binomických rozdělení</td></tr>
<tr><td>Intervalový odhad rozdílu parametru binomických rozdělení</td></tr>
</table>

Dvouvýběrový F-test (Test o shodě rozptylů):

Porovnáváme dva nezávislé výběry (mezi čísly v jednom a druhém výběru není souvislost). Tzn. naměřili jsme hodnoty:

$x_{11}, x_{12}, \dots, x_{1n_1}$ náh. veličiny $X_1 \to N(\mu_1, \sigma_1^2)$ a

$x_{21}, x_{22}, \dots, x_{2n}$ náh. veličiny $X_2 \to N(\mu_2, \sigma_2^2)$

(tj. oba výběry pocházejí z normálního rozdělení!)

Hypotézy:

$H_0: \sigma_1 = \sigma_2$

$H_A: \sigma_1 \neq \sigma_2 \quad (\sigma_1 > \sigma_2 ; \sigma_1 < \sigma_2)$

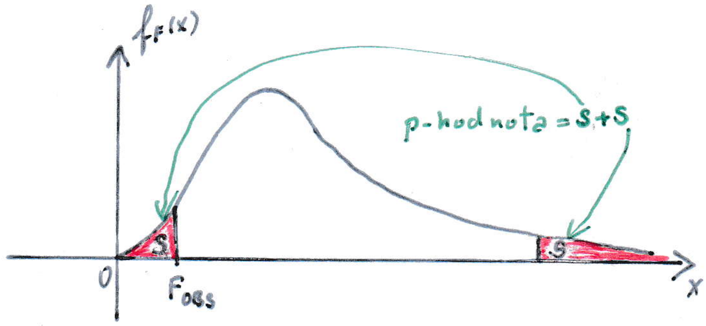
Testová statistika:

$$F = \frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}} \underset{\substack{\uparrow \\ \text{pokud platí } H_0}}{=} \frac{S_1^2}{S_2^2} \longrightarrow F_{n_1-1,\, n_2-1} \quad \dots \text{Fisherovo-Snedecorovo rozdělení!}$$

$$F_{OBS} = \frac{s_1}{s_2}$$

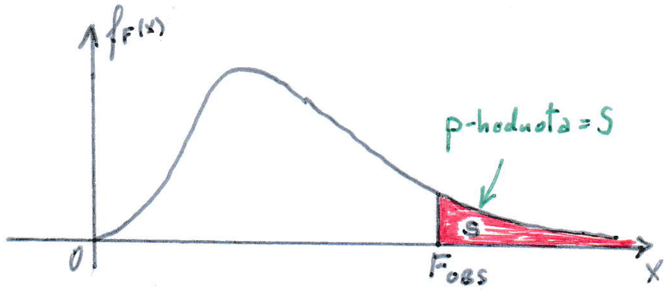
a) $H_A: \sigma_1 \neq \sigma_2$

$$\text{p-hodnota} = 2 \cdot \min\{F(F_{OBS}) ; 1 - F(F_{OBS})\}$$



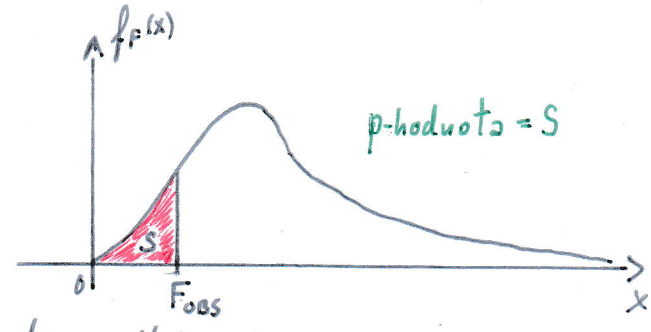
b) $H_A: \sigma_1 > \sigma_2$ (⇒ F_{OBS} je „velké")

$$\text{p-hodnota} = 1 - F(F_{OBS})$$



c) $H_A: \sigma_1 < \sigma_2$ (⇒ F_{OBS} je „malé")

$$\text{p-hodnota} = F(F_{OBS})$$



Závěr: p-hodnota $< \alpha$ ⇒ Na hladině významnosti zamítáme H_0.
$\geq \alpha$ ⇒ nezamítáme H_0.

Př. Změřili jsme hmotnosti 30 brambor odrůdy A a 20 brambor odrůdy B. Výběrová směrodatná odchylka u brambor A je $s_1 = 0{,}3$ kg a u brambor B je $s_2 = 0{,}41$ kg.

Předpokládejme, že hmotnosti brambor A i B jsou náhodné veličiny s normálním rozdělením pravděpodobnosti. Testujte hypotézu, že odrůda B má větší variabilitu hmotnosti.

⇒ Test o shodě rozptylů

Předpoklady: X_1 ... hmotnost brambor A $\rightarrow N(\mu_1, \sigma_1^2)$
X_2 ... hmotnost brambor B $\rightarrow N(\mu_2, \sigma_2^2)$ } Podle zadání splněno!

$H_0: \sigma_1 = \sigma_2$

$H_A: \sigma_1 < \sigma_2$

$$F = \frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}} = \frac{S_1^2}{S_2^2} \longrightarrow \underbrace{F_{30-1,\,20-1}}_{\text{Fisherovo-Snedecorovo rozdělení}}$$

$$F_{OBS} = \frac{0{,}3^2}{0{,}41^2} \doteq 0{,}5354$$

$$\text{p-hodnota} = F(F_{OBS}) = \underline{\underline{0{,}063}}$$

⇒ Na hladině významnosti $\alpha = 0{,}05$ nezamítáme nulovou hypotézu, že rozptyl hmotností je u odrůdy A stejný jako u odrůdy B.

Dvouvýběrový t-test (Test o shodě středních hodnot): Porovnáváme

dva nezávislé výběry. Tzn. naměřili jsme hodnoty:

$x_{11}, x_{12}, \dots, x_{1n_1}$ náh. veličiny $X_1 \to N(\mu_1, \sigma^2)$

$x_{21}, x_{22}, \dots, x_{2n_2}$ náh. veličiny $X_2 \to N(\mu_2, \sigma^2)$

(Tj. X_1 a X_2 mají stejný rozptyl!)

Hypotézy: $H_0: \mu_1 - \mu_2 = C$

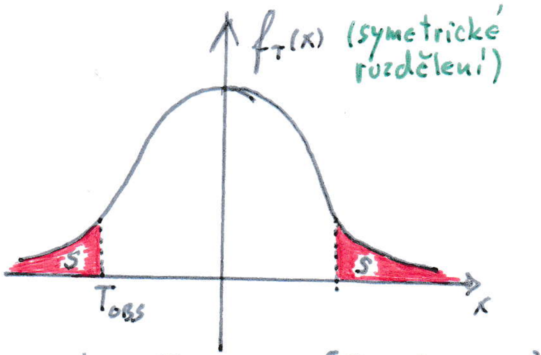
$H_A: \mu_1 - \mu_2 \neq C$

Testová statistika: $$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} \cdot \frac{(n_1+n_2)}{n_1 \cdot n_2}}} \longrightarrow t_{n_1+n_2-2}$$

|| (Pokud platí H_0)

$$T = \frac{\bar{X}_1 - \bar{X}_2 - C}{\sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} \frac{(n_1+n_2)}{n_1 \cdot n_2}}}$$

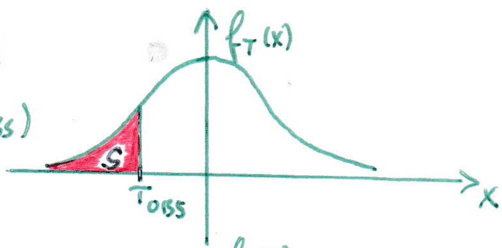
$$T_{OBS} = \frac{\bar{x}_1 - \bar{x}_2 - C}{\sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} \frac{(n_1+n_2)}{n_1 \cdot n_2}}}$$



p-value $= 2S = 2 \cdot \min\{F(T_{OBS}), 1-F(T_{OBS})\}$

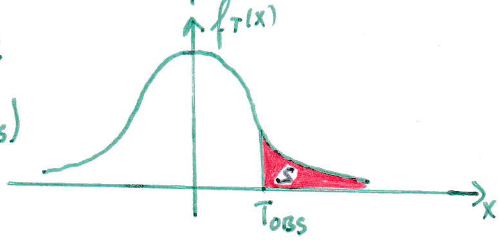
$H_A: \mu_1 - \mu_2 < C$

p-value $= S = F(T_{OBS})$



$H_A: \mu_1 - \mu_2 > C$

p-value $= S = 1 - F(T_{OBS})$



Př. Dva stroje vyrábí podložky o daném poloměru. Byly změřeny poloměry 20-ti podložek, které vyrobil první stroj a 22 podložek, které vyrobil druhý stroj. Ověřte by měření, že poloměry podložek vyrobených prvním strojem mají stejnou střední hodnotu jako ty vyrobené druhým strojem.

1.) Zjistíme, zda data neobsahují odlehlá pozorování:

NE =>

2.) Ověříme, zda jsou splněny předpoklady t-testu:

X_1... poloměry podložek z 1. stroje - musí mít normální rozdělení

boxplot, qqnorm + qqline, moments: skewness, moments: kurtosis, hist
OK, OK, $\in \langle -2,2 \rangle$ => OK, OK

shapiro.test => p-hodnota = 0,9375 => OK => Nezamítáme normalitu X_1.

X_2... poloměry podložek z 2. stroje - musí mít normální rozdělení

analogicky jako u X_1 => Nezamítáme normalitu X_2.

Rovnost rozptylů: $H_0: \sigma_1 = \sigma_2$ var.test => p-hodnota = 0,76 => Nezamítáme $\sigma_1 = \sigma_2$

3.) Provedeme dvouvýběrový t-test: t.test (polomery$X1, polomery$X2, mu=0, alternative = "two.sided", var.equal=TRUE, conf.level=0.95)

$H_0: \mu_1 = \mu_2$ $(\Rightarrow c=0)$

$H_A: \mu_1 \neq \mu_2$

$$T = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2} \cdot \frac{n_1+n_2}{n_1 \cdot n_2}}} \rightarrow t_{n_1+n_2-2}$$

$T_{OBS} = 16{,}511$

p-value $\ll 0{,}05$ =>

Zamítáme H_0, střední hodnoty se liší!

=>

$H_0: \mu_1 = \mu_2$ $(\mu_1 - \mu_2 = 0 = c)$

$H_A: \mu_1 > \mu_2$ $(\mu_1 - \mu_2 > 0 = c)$

T = —— || ——

T_{OBS} = —— || ——

p-value $\ll 0{,}05$ =>

Zamítáme H_0 =>

Střední hodnota poloměrů podložek z 1. stroje je statisticky významně větší než těch ze 2. stroje (p-value $\ll 0{,}05$).

12.2 Příklady

1. Data v souboru cholesterol2.xls udávají hladinu cholesterolu v krvi mužů dvou různých věkových skupin (20-30 letých a 40-50 letých). Ověřte na hladině významnosti 0,05 hypotézu, zda se hladina cholesterolu v krvi starších mužů neliší od hladiny cholesterolu v krvi mladších mužů.

Testované parametry: μ_M, μ_S $(x_{0,5,M}, x_{0,5,S})$

Možné testy:	Dvouvýběrový Studentův t-test	Aspinové-Welshův test	Mannův-Whitneyův test
	↓	↓	↓
Předpoklady testů:	normalita homoskedasticita nezávislost výběrů	normalita nezávislost výběrů	nezávislost výběrů

Normalita u mladších mužů: H_0 : data pocházejí z normálního rozdělení
OK
H_A : data nepocházejí z normálního rozdělení
explorační grafy ⇒ OK
Shapiro-Wilkův test:
$x_{OBS} = 0,97$
p-hodnota = 0,29 > 0,05
Na hladině významnosti 0,05 nezamítáme nulovou hypotézu o normalitě dat (S-W test, p-hodnota = 0,29).

Normalita u starších mužů: H_0 : data pocházejí z normálního rozdělení
OK
H_A : data nepocházejí z normálního rozdělení
explorační grafy ⇒ OK
Shapiro-Wilkův test:
$x_{OBS} = 0,99$
p-hodnota = 0,92 > 0,05
Na hladině významnosti 0,05 nezamítáme nulovou hypotézu o normalitě dat (S-W test,p-hodnota = 0,92).

Homoskedasticita: $H_0 : \sigma_M^2 = \sigma_S^2$ (M: n = 99, s = 0,115)
NE!
$H_A : \sigma_M^2 \neq \sigma_S^2$ (S: n = 85, s = 0,368)
$x_{OBS} = 0,10$
p-hodnota F-testu = 0,001
Na hladině významnosti 0,05 zamítáme nulovou hypotézu o shodě rozptylů (F test, p-hodnota = 0,001).

Nezávislost výběrů: Z podstaty zadání splněno.

Zvolený test: Aspinové-Welchův test

Nulová hypotéza H_0: $\mu_M = \mu_S$ (M: $\overline{x}_M$ = 4,597; S: $\overline{x}_S$ = 5,101)

Alternativní hypotéza H_A: $\mu_M < \mu_S$

Pozorovaná hodnota x_{OBS}: -12,923

p-hodnota: $<< 0,001$ ⇒ zamítáme nulovou hypotézu

Rozhodnutí: Na hladině významnosti 0,05 zamítáme nulovou hypotézu (Aspinové-Welshův test, p-hodnota $<< 0,001$) ⇒ s pravděpodobností 95 % můžeme tvrdit, že je hladina cholesterolu u starších mužů statisticky významně vyšší než u mladších mužů. [p-hodnota $<< 0,001$]

Upraveno

2. Údaje v souboru deprese.xls představují délku remise ve dnech z prostého náhodného výběru ze dvou různých skupin pacientů (pacienti s endogenní depresi a pacienti s neurotickou depresí). Ověřte, zda je pozorovaný rozdíl mezi průměrnou délkou remise u těchto dvou skupin pacientů statisticky významný.

Testované parametry: μ_M, μ_S $(x_{0,5,M}, x_{0,5,S})$

Možné testy:	Dvouvýběrový Studentův t-test	Aspinové-Welshův test	Mannův-Whitneyův test
	↓	↓	↓
Předpoklady testů:	normalita homoskedasticita nezávislost výběrů	normalita nezávislost výběrů	nezávislost výběrů

Normalita u end. skupiny:

NE!

H_0 : data pocházejí z normálního rozdělení
H_A : data nepocházejí z normálního rozdělení
Shapiro-Wilkův test:
$x_{OBS} = 0,68$
p-hodnota $<< 0,001$
Na hladině významnosti 0,05 zamítáme nulovou hypotézu o normalitě dat (S-W test, p-hodnota $<< 0,001$).
⇒ Není nutné zkoumat další předpoklady.

Zvolený test: Mannův-Whitneyův test

Nulová hypotéza H_0: $x_{0,5,E} = x_{0,5,N}$ $(E$: $x_{0,5,E} = 159{,}5$; N: $x_{0,5,N} = 512{,}5)$

Alternativní hypotéza H_A: $x_{0,5,E} < x_{0,5,N}$

Pozorovaná hodnota x_{OBS}: 81

p-hodnota: $0,62 > 0,05 \Rightarrow$ nezamítáme nulovou hypotézu

Rozhodnutí:

Na hladině významnosti 0,05 nezamítáme nulovou hypotézu (Mannův-Whitneyův test, p-hodnota = 0,62) ⇒ s pravděpodobností 95 % není pozorovaný rozdíl mezi průměrnou délkou remise statisticky významný.

[p-hodnota $= 0,62$]

Párový t-test : U n statistických jednotek zjišťujeme páry závislých pozorování. Máme hodnoty:

$x_1, x_2, \dots, x_n$ náh. veličiny X ... např. hodnoty před...

$y_1, y_2, \dots, y_n$ náh. veličiny Y ... např. hodnoty po...

snažíme se zjistit, zda $E(X) = \mu_1 = \mu_2 = E(Y)$.

Definujeme proto n.v. $R = X - Y \Rightarrow E(R) = E(X) - E(Y) \Rightarrow$

$\Rightarrow E(R) = \mu = \mu_1 - \mu_2 \Rightarrow$

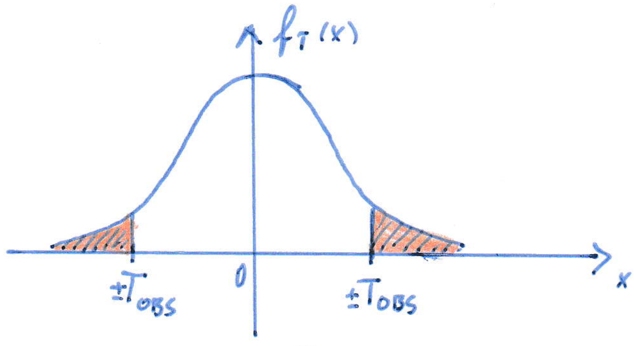
Pokud $R \rightarrow N(\mu, \sigma^2)$, můžeme jednovýběrovým t-testem ověřit hypotézu

$H_0: \mu = \mu_1 - \mu_2 = C$ (tzn $\mu_1 = \mu_2 + C$)

$H_A: \mu \neq C$ (případně jednostranné alternativy)

$$T = \frac{\bar{R} - \mu}{S}\sqrt{n} \longrightarrow t_{n-1} \text{ ... studentovo rozdělení}$$

pokud platí $H_0 \Rightarrow$ $$T_{OBS} = \frac{\bar{r} - C}{s}\sqrt{n}$$



p-value $= 2 \cdot \min\{F(T_{OBS}), 1 - F(T_{OBS})\}$

p-value $< \alpha \Rightarrow$ na hladině významnosti zamítáme H_0

3. Sledujeme osmolalitu moči na lůžkové stanici v 08:00 hodin a v 11:00 hodin u 16 mužů. Na základě výsledků uvedených v souboru osmolalita.xls ověřte, zda se osmolalita statisticky významně zvýšila.

Testované parametry: μ_{11-8}

Možné testy:	Párový Studentův t-test	Párový znaménkový test	Wilcoxnův párový test
	↓		↓
Předpoklady testů:	normalita počítat přes diference		spojité rozdělení symetrické kolem mediánu

Normalita diferencí (11-8): H_0 : data pocházejí z normálního rozdělení
H_A : data nepocházejí z normálního rozdělení
Shapiro-Wilkův test:
$x_{OBS} = 0,87$
p-hodnota $= 0,06$
Na hladině významnosti 0,05 nezamítáme nulovou hypotézu o normalitě dat (S-W test, p-hodnota = 0,06).

Zvolený test: Párový Studentův t-test

Nulová hypotéza H_0: $\mu_{11-8} = 0$ $\quad \overline{x}_{11-8} = 8,5; s = 4,8; n = 16$

Alternativní hypotéza H_A: $\mu_{11-8} > 0$

Pozorovaná hodnota x_{OBS}: 7,14

p-hodnota: $<< 0,001 < 0,05 \Rightarrow$ zamítáme nulovou hypotézu

Rozhodnutí:

Na hladině významnosti 0,05 zamítáme nulovou hypotézu (párový t-test, p-hodnota $<< 0,001$) $\Rightarrow$ s pravděpodobností 95 % můžeme tvrdit, že se hodnota osmolality statisticky významně zvýšila.

[p-hodnota $<< 0,001$]

Test o shodě parametrů dvou binomických rozdělení:

Mějme $X_1 \to Bi(n_1, \pi_1)$ a $X_2 \to Bi(n_2, \pi_2)$.

Naměřili jsme hodnoty $X_1 = x_1$ a $X_2 = x_2$.

Tzn. úspěch nastal x_1-krát při n_1 pokusech a x_2-krát při n_2 pokusech.

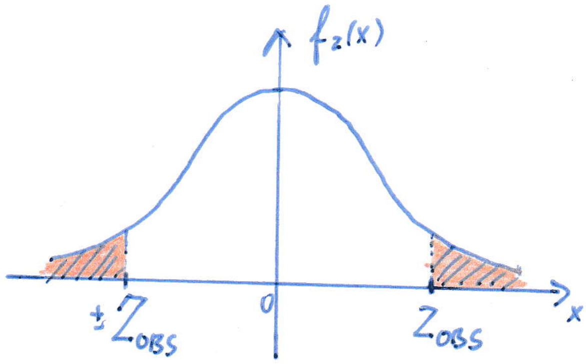
$\Rightarrow p_1 = \frac{x_1}{n_1}$ = odhad π_1 a $p_2 = \frac{x_2}{n_2}$ = odhad π_2

Ověřujeme hypotézu:

$H_0: \pi_1 - \pi_2 = c$

Pokud platí H_0:
$$Z = \frac{(p_1 - p_2) - \overbrace{(\pi_1 - \pi_2)}^{\text{„}c\text{"}}}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}} \xrightarrow{\overset{(CLV)}{\bullet}} N(0,1)$$

$$Z_{OBS} = \frac{p_1 - p_2 - c}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}}$$



p-value $= 2 \cdot \min\{F_Z(Z_{OBS});\ 1 - F_Z(Z_{OBS})\}$

p-value $< 0{,}05$ $\Rightarrow$ Zamítáme H_0 ve prospěch alternativy.

Předpoklady testu: $n_1 > \frac{9}{p_1(1-p_1)}$ a $n_2 > \frac{9}{p_2(1-p_2)}$

Př. Ve skupině 150-ti pacientů operovaných laparoskopicky se v 18-ti případech vyskytly pooperační komplikace. Ve skupině 146-ti pacientů operovaných otevřeně se vyskytlo 21 komplikací. Testujte hypotézu, že pravděpodobnost pooperačních komplikací je u laparoskopických operací stejná jako u otevřených.

$\Rightarrow$ X_1 … počet komplikací po lap. operacích $\rightarrow Bi(n_1, \pi_1) = Bi(150, \pi_1)$

X_2 … počet komplikací po otevř. operacích $\rightarrow Bi(n_2, \pi_2) = Bi(146, \pi_2)$

$\Rightarrow$ Test o shodě parametrů binom. rozdělení

Předpoklady: $n = 150 > \dfrac{9}{p_1(1-p_1)} = \dfrac{9}{\frac{18}{150}\left(1-\frac{18}{150}\right)} = \dfrac{9}{\frac{18 \cdot 132}{150 \cdot 150}} \doteq 85{,}2 \Rightarrow OK$

$n = 146 > \dfrac{9}{p_2(1-p_2)} = \dfrac{9}{\frac{21}{146}\left(1-\frac{21}{146}\right)} = \dfrac{9}{\frac{21 \cdot 125}{146 \cdot 146}} \doteq 73 \Rightarrow OK$

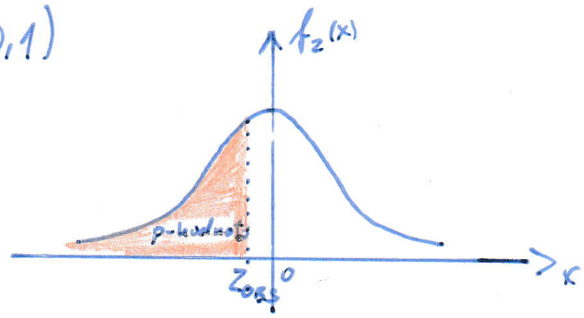
$H_0: \pi_1 - \pi_2 = 0$

$H_A: \pi_1 \neq \pi_2$

$$Z = \frac{(P_1 - P_2) - (\pi_1 - \pi_2)}{\sqrt{\frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}}} \longrightarrow N(0,1)$$

$$Z_{OBS} = \frac{\frac{18}{150} - \frac{21}{146} - 0}{\sqrt{\frac{\frac{18}{150} \cdot \frac{132}{150}}{150} + \frac{\frac{21}{146} \cdot \frac{125}{146}}{146}}} \doteq -0{,}6059$$

← pokud platí H_0



p-hodnota $\doteq 0{,}545$

$\Rightarrow$ Na hladině významnosti nezamítáme H_0. Tzn. rozdíl mezi pravděpodobnostmi pooperačních komplikací po otevřených operacích a po laparoskopických není statisticky významný.

4. Byly testovány polovodičové součástky dvou výrobců - MM a PP. MM prohlašuje, že její výrobky mají nižší procento vadných kusů. Pro ověření tohoto tvrzení bylo z produkce MM náhodně vybráno 200 součástek, z nichž 14 bylo vadných. Podobný experiment byl proveden u firmy PP s výsledkem 10 vadných ze 100 náhodně vybraných součástek.

   a) Otestujte tvrzení firmy MM čistým testem významnosti.

   Testované parametry: π_{MM}, π_{PP}

   Možné testy: Test homogenity dvou binomických rozdělení

   Předpoklady testů: $n_{MM} \cdot p_{MM} \cdot (1 - p_{MM}) > 9 \Rightarrow 200 \cdot \frac{14}{200} \cdot (1 - \frac{14}{200}) = 13,02$
   $n_{PP} \cdot p_{PP} \cdot (1 - p_{PP}) > 9 \Rightarrow 100 \cdot \frac{10}{100} \cdot (1 - \frac{10}{100}) = 9$

   Zvolený test: Test homogenity dvou binomických rozdělení

   Nulová hypotéza H_0: $\pi_{MM} = \pi_{PP}$

   Alternativní hypotéza H_A: $\pi_{MM} < \pi_{PP}$

   Pozorovaná hodnota x_{OBS}: $-0,90$

   p-hodnota: $0,18 > 0,05 \Rightarrow$ nezamítáme nulovou hypotézu

   Rozhodnutí:

   Na hladině významnosti 0,05 nezamítáme nulovou hypotézu (test homogenity dvou binomických rozdělení, p-hodnota = 0,18)$\Rightarrow$ s pravděpodobností 95 % nemůžeme tvrdit, že výrobky firmy MM mají nižší procento vadných kusů. [p-hodnota $= 0,18$]

   b) Otestujte tvrzení firmy MM prostřednictvím intervalového odhadu na hladině významnosti 0,05.

   Testované parametry: π_{MM}, π_{PP}

   Předpoklady pro použití intervalového odhadu:

   $n_{MM} \cdot p_{MM} \cdot (1 - p_{MM}) > 9 \Rightarrow 200 \cdot \frac{14}{200} \cdot (1 - \frac{14}{200}) = 13,02$
   $n_{PP} \cdot p_{PP} \cdot (1 - p_{PP}) > 9 \Rightarrow 100 \cdot \frac{10}{100} \cdot (1 - \frac{10}{100}) = 9$

   Nulová hypotéza H_0: $\pi_{MM} = \pi_{PP}$

   Alternativní hypotéza H_A: $\pi_{MM} < \pi_{PP}$

   Rozhodnutí:

   $P(-0,095 < \pi_{MM} - \pi_{PP} < 0,035) = 0,95$, protože intervalový odhad obsahuje nulu, nelze s 95% pravděpodobnostní tvrdit, že výrobky firmy MM mají nižší procento vadných kusů.

   [$P(-0,095 < \pi_{MM} - \pi_{PP} < 0,035) = 0,95$]

12.1 Test z teorie

1. Označte všechny neparametrické (robustní) testy.
   a) dvouvýběrový t-test,
   b) párový t-test,
   c) Aspinové-Welchův test,
   d) Mannův-Whitneho test,
   e) znaménkový test nebo párový Wilcoxonův test,
   f) dvouvýběrový F-test (test o shodě rozptylů),
   g) test homogenity dvou binomických rozdělení.
2. Předpokladem pro použití Mannova-Whitneyho testu je
   a) normalita obou výběrů,
   b) normalita a homoskedasticita obou výběrů ,
   c) normalita a heteroskedasticita obou výběrů ,
   d) spojitost a stejný tvar rozdělení obou výběrů,
   e) dostatečný rozsah obou výběrů ($n_i > 9/(p_i(1-p_i)), i = 1, 2$), kde p_i je relativní četnost sledovaného jevu v i-tém výběru.
3. Předpokladem pro použití párového t-testu je
   a) normalita obou výběrů,
   b) normalita a homoskedasticita obou výběrů ,
   c) normalita a heteroskedasticita obou výběrů ,
   d) spojitost a stejný tvar rozdělení obou výběrů,
   e) dostatečný rozsah obou výběrů ($n_i > 9/(p_i(1-p_i)), i = 1, 2$), kde p_i je relativní četnost sledovaného jevu v i-tém výběru.
4. Předpokladem pro použití Aspinové-Welchova testu je
   a) normalita obou výběrů,
   b) normalita a homoskedasticita obou výběrů,
   c) normalita a heteroskedasticita obou výběrů,
   d) spojitost a stejný tvar rozdělení obou výběrů,
   e) dostatečný rozsah obou výběrů ($n_i > 9/(p_i(1-p_i)), i = 1, 2$), kde p_i je relativní četnost sledovaného jevu v i-tém výběru.
5. Neparametrickým protějškem Aspinové-Welchova testu je
   a) dvouvýběrový t-test,
   b) párový t-test,
   c) Mannův-Whitneyho test,
   d) znaménkový test nebo párový Wilcoxonův test,
   e) dvouvýběrový F-test (test o shodě rozptylů),
   f) test homogenity dvou binomických rozdělení.

6. Neparametrickým protějškem párového t-testu je

   a) dvouvýběrový t-test,
   b) Aspinové-Welchův test,
   c) Mannův-Whitneyho test,
   d) znaménkový test nebo párový Wilcoxonův test,
   e) dvouvýběrový F-test (test o shodě rozptylů),
   f) test homogenity dvou binomických rozdělení.

7. Neparametrickým protějškem dvouvýběrového t-testu je

   a) párový t-test,
   b) Aspinové-Welchův test,
   c) Mannův-Whitneyho test,
   d) znaménkový test nebo párový Wilcoxonův test,
   e) dvouvýběrový F-test (test o shodě rozptylů),
   f) test homogenity dvou binomických rozdělení.

8. Tabáková firma TAB prohlašuje, že jejich cigarety mají nižší obsah nikotinu než cigarety NIK. Obsah nikotinu byl změřen ve 100 cigaretách TAB a 100 cigaretách NIK. Na základě obou výběru byla ověřena homoskedasticita obsahů nikotinu v cigaretách TAB a NIK. Bylo ověřeno, že obsah nikotinu v cigaretách má normální rozdělení. Chceme-li ověřit, zda lze tvrzení firmy TAB prohlásit za nepravdivé, použijeme

   a) dvouvýběrový t-test,
   b) párový t-test,
   c) Aspinové-Welchův test,
   d) Mannův-Whitneyho test,
   e) znaménkový test nebo párový Wilcoxonův test,
   f) dvouvýběrový F-test (test o shodě rozptylů),
   g) test homogenity dvou binomických rozdělení.

9. Při testování ojetí (mm) pneumatik 11 automobilů určité značky byla zamítnuta normalita ojetí pneumatik (mm). Chceme-li ověřit, zda se pravé a levé přední pneumatiky automobilů této značky ojíždějí srovnatelně, použijeme

   a) dvouvýběrový t-test,
   b) párový t-test,
   c) Aspinové-Welchův test,
   d) Mannův-Whitneyho test,
   e) znaménkový test nebo párový Wilcoxonův test,
   f) dvouvýběrový F-test (test o shodě rozptylů),
   g) test homogenity dvou binomických rozdělení.

10. Bylo ověřeno, že hmotnost balení cukru má normální rozdělení. Testujeme-li, zda seřízením výrobní linky došlo ke snížení kolísavosti hmotnosti balení cukru, použijeme

    a) dvouvýběrový t-test,
    b) párový t-test,
    c) Aspinové-Welchův test,
    d) Mannův-Whitneyho test,
    e) znaménkový test nebo párový Wilcoxonův test,
    f) dvouvýběrový F-test (test o shodě rozptylů),
    g) test homogenity dvou binomických rozdělení.

11. Určete, zda jsou následující tvrzení pravdivá.

    N a) Při neparametrickém testu homogenity dvou binomických rozdělení nemusíme ověřovat žádné předpoklady o výběrech. [požadavek na rozsah výběrů]
    N b) Mannův-Whitneyho test se používá pro ověření shody úrovně ve dvou závislých výběrech.
    N c) Každý test hypotézy $H_0 : \mu_1 = \mu_2$, tj. hypotézy o shodě dvou středních hodnot, je testem párovým.