PRAVDĚPODOBNOST A STATISTIKA

Testování hypotéz-dvouvýběrové testy

Katedra kvantitativních metod FVL UO Brno

13. PŘEDNÁŠKA

Obsah

Nechť $X_1, X_2, \dots, X_{n_1}$ je náhodný výběr z rozdělení $N(\mu_1, \sigma_1^2)$ a $Y_1, Y_2, \dots, Y_{n_2}$ je náhodný výběr z rozdělení $N(\mu_2, \sigma_2^2)$.
Označme S_X^2 a S_Y^2 odpovídající výběrové rozptyly.
Statistika

$$F = \frac{S_X^2}{S_Y^2} \cdot \frac{\sigma_2^2}{\sigma_1^2}$$

má Fisherovo-Snedecorovo rozdělení s $\nu_1 = n_1 - 1$ a $\nu_2 = n_2 - 1$ stupni volnosti.
Využijeme tohoto tvrzení pro konstrukci testu o shodě rozptylů.

Nechť $X_1, X_2, \ldots, X_{n_1}$ je náhodný výběr z rozdělení $N(\mu_1, \sigma_1^2)$ a $Y_1, Y_2, \ldots, Y_{n_2}$ je náhodný výběr z rozdělení $N(\mu_2, \sigma_2^2)$.
Označme S_X^2 a S_Y^2 odpovídající výběrové rozptyly.
Statistika

$$F = \frac{S_X^2}{S_Y^2} \cdot \frac{\sigma_2^2}{\sigma_1^2}$$

má Fisherovo-Snedecorovo rozdělení s $\nu_1 = n_1 - 1$ a $\nu_2 = n_2 - 1$ stupni volnosti.
Využijeme tohoto tvrzení pro konstrukci testu o shodě rozptylů.

Test shody rozptylů

Nechť $x_1, x_2, \ldots, x_{n_1}$ značí hodnoty náhodného výběru z $N(\mu_1, \sigma_1^2)$,
$y_1, y_2, \ldots, y_{n_2}$ značí hodnoty náhodného výběru z $N(\mu_2, \sigma_2^2)$,
s_x^2 a s_y^2 jsou odpovídající hodnoty výběrových rozptylů.

Testujeme hypotézu, že parametr σ_1^2 je roven hodnotě σ_2^2:

$$H : \sigma_1^2 = \sigma_2^2,$$

testové kriterium je statistika

$$F = \frac{s_x^2}{s_y^2},$$

která má při platnosti hypotézy H Fisherovo-Snedecorovo rozdělení s $\nu_1 = n_1 - 1$ a $\nu_2 = n_2 - 1$ stupni volnosti.

Test shody rozptylů

Nechť $x_1, x_2, \ldots, x_{n_1}$ značí hodnoty náhodného výběru z $N(\mu_1, \sigma_1^2)$,
$y_1, y_2, \ldots, y_{n_2}$ značí hodnoty náhodného výběru z $N(\mu_2, \sigma_2^2)$,
s_x^2 a s_y^2 jsou odpovídající hodnoty výběrových rozptylů.

Testujeme hypotézu, že parametr σ_1^2 je roven hodnotě σ_2^2:

$$H : \sigma_1^2 = \sigma_2^2,$$

testové kriterium je statistika

$$F = \frac{s_x^2}{s_y^2},$$

která má při platnosti hypotézy H Fisherovo-Snedecorovo rozdělení s $\nu_1 = n_1 - 1$ a $\nu_2 = n_2 - 1$ stupni volnosti.

Podle alternativní hypotézy volíme následující kritické obory:

alternativní hypotéza	**kritický obor**
$A : \sigma_1^2 > \sigma_2^2$	$W_\alpha = \{F, F \geq F_{1-\alpha}(\nu_1, \nu_2)\}$
$A : \sigma_1^2 < \sigma_2^2$	$W_\alpha = \{F, F \leq F_{\alpha}(\nu_1, \nu_2)\}$
$A : \sigma_1^2 \neq \sigma_2^2$	$W_\alpha = \left\{F, F \leq F_{\frac{\alpha}{2}}(\nu_1, \nu_2) \vee F \geq F_{1-\frac{\alpha}{2}}(\nu_1, \nu_2)\right\}$

kde $F_\alpha(\nu_1, \nu_2)$, $F_{1-\alpha}(\nu_1, \nu_2)$, $F_{1-\frac{\alpha}{2}}(\nu_1, \nu_2)$ jsou kvantily Fisher-Snedecorova rozdělení, $\nu_1 = n_1 - 1$, $\nu_2 = n_2 - 1$.

Nechť $X_1, X_2, \ldots, X_{n_1}$ je náhodný výběr z rozdělení $N(\mu_1, \sigma_1^2)$ a $Y_1, Y_2, \ldots, Y_{n_2}$ je náhodný výběr z rozdělení $N(\mu_2, \sigma_2^2)$. Dále předpokládejme, že tyto výběry jsou **nezávislé**.

Označme $\overline{X}$ a $\overline{Y}$ odpovídající výběrové průměry, S_X^2 a S_Y^2 odpovídající výběrové rozptyly. Platí-li $\sigma_1^2 = \sigma_2^2$, pak statistika

$$T = \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{S} \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}},$$

kde

$$S = \left[\frac{(n_1 - 1)S_X^2 + (n_2 - 1)S_Y^2}{n_1 + n_2 - 2} \right]^{1/2}$$

má Studentovo rozdělení s $\nu = n_1 + n_2 - 2$ stupni volnosti.

Test shody středních hodnot ($\sigma_1^2 = \sigma_2^2$)

Nechť $X_1, X_2, \dots, X_{n_1}$ je náhodný výběr z rozdělení $N(\mu_1, \sigma_1^2)$ a $Y_1, Y_2, \dots, Y_{n_2}$ je náhodný výběr z rozdělení $N(\mu_2, \sigma_2^2)$. Dále předpokládejme, že tyto výběry jsou **nezávislé**.

Označme $\overline{X}$ a $\overline{Y}$ odpovídající výběrové průměry, S_X^2 a S_Y^2 odpovídající výběrové rozptyly. Platí-li $\sigma_1^2 = \sigma_2^2$, pak statistika

$$T = \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{S} \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}},$$

kde

$$S = \left[\frac{(n_1 - 1)S_X^2 + (n_2 - 1)S_Y^2}{n_1 + n_2 - 2} \right]^{1/2}$$

má Studentovo rozdělení s $\nu = n_1 + n_2 - 2$ stupni volnosti.

Test shody středních hodnot ($\sigma_1^2 = \sigma_2^2$)

Nechť $x_1, x_2, \ldots, x_{n_1}$ značí hodnoty náhodného výběru z $N(\mu_1, \sigma_1^2)$, $y_1, y_2, \ldots, y_{n_2}$ značí hodnoty náhodného výběru z $N(\mu_2, \sigma_2^2)$,
$\overline{x}$, $\overline{y}$, s_x^2 a s_y^2 jsou odpovídající hodnoty výběrových průměrů a rozptylů.

Testujeme hypotézu, že parametr μ_1 je roven hodnotě μ_2 ($\sigma_1^2 = \sigma_2^2$):

$$H : \mu_1 = \mu_2,$$

testové kriterium je statistika

$$t = \frac{\overline{x} - \overline{y}}{S} \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}},$$

kde

$$S = \left[\frac{(n_1 - 1)s_x^2 + (n_2 - 1)s_y^2}{n_1 + n_2 - 2} \right]^{1/2}$$

která má při platnosti nulové hypotézy H Studentovo rozdělení s $\nu = n_1 + n_2 - 2$ stupni volnosti.

Test shody středních hodnot ($\sigma_1^2 = \sigma_2^2$)

Nechť $x_1, x_2, \ldots, x_{n_1}$ značí hodnoty náhodného výběru z $N(\mu_1, \sigma_1^2)$, $y_1, y_2, \ldots, y_{n_2}$ značí hodnoty náhodného výběru z $N(\mu_2, \sigma_2^2)$,
$\overline{x}$, $\overline{y}$, s_x^2 a s_y^2 jsou odpovídající hodnoty výběrových průměrů a rozptylů.

Testujeme hypotézu, že parametr μ_1 je roven hodnotě μ_2 ($\sigma_1^2 = \sigma_2^2$):

$$H : \mu_1 = \mu_2,$$

testové kriterium je statistika

$$t = \frac{\overline{x} - \overline{y}}{S} \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}},$$

kde

$$S = \left[\frac{(n_1 - 1)s_x^2 + (n_2 - 1)s_y^2}{n_1 + n_2 - 2} \right]^{1/2}$$

která má při platnosti nulové hypotézy H Studentovo rozdělení s $\nu = n_1 + n_2 - 2$ stupni volnosti.

Podle alternativní hypotézy volíme následující kritické obory:

alternativní hypotéza	**kritický obor**
$A: \mu_1 > \mu_2$	$W_\alpha = \{t, t \geq t_{1-\alpha}(\nu)\}$
$A: \mu_1 < \mu_2$	$W_\alpha = \{t, t \leq -t_{1-\alpha}(\nu)\}$
$A: \mu_1 \neq \mu_2$	$W_\alpha = \left\{t, \lvert t\rvert \geq t_{1-\frac{\alpha}{2}}(\nu)\right\}$

kde $t_{1-\alpha}(\nu)$, $t_{1-\frac{\alpha}{2}}(\nu)$ jsou kvantily Studentova rozdělení, $\nu = n_1 + n_2 - 2$.

Nechť $X_1, X_2, \ldots, X_{n_1}$ je náhodný výběr z rozdělení $N(\mu_1, \sigma_1^2)$ a $Y_1, Y_2, \ldots, Y_{n_2}$ je náhodný výběr z rozdělení $N(\mu_2, \sigma_2^2)$. Dále předpokládejme, že tyto výběry jsou **nezávislé**.

Označme $\overline{X}$ a $\overline{Y}$ odpovídající výběrové průměry, S_X^2 a S_Y^2 odpovídající výběrové rozptyly. Pro $\sigma_1^2 \neq \sigma_2^2$, má statistika

$$T = \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{S_X^2}{n_1} + \frac{S_Y^2}{n_2}}},$$

přibližně Studentovo rozdělení s ν stupni volnosti.

Test shody středních hodnot ($\sigma_1^2 \neq \sigma_2^2$)

Nechť $X_1, X_2, \ldots, X_{n_1}$ je náhodný výběr z rozdělení $N(\mu_1, \sigma_1^2)$ a $Y_1, Y_2, \ldots, Y_{n_2}$ je náhodný výběr z rozdělení $N(\mu_2, \sigma_2^2)$. Dále předpokládejme, že tyto výběry jsou **nezávislé**.

Označme $\overline{X}$ a $\overline{Y}$ odpovídající výběrové průměry, S_X^2 a S_Y^2 odpovídající výběrové rozptyly. Pro $\sigma_1^2 \neq \sigma_2^2$, má statistika

$$T = \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{S_X^2}{n_1} + \frac{S_Y^2}{n_2}}},$$

přibližně Studentovo rozdělení s ν stupni volnosti.

Test shody středních hodnot $(\sigma_1^2 \neq \sigma_2^2)$

Nechť $x_1, x_2, \ldots, x_{n_1}$ značí hodnoty náhodného výběru z $N(\mu_1, \sigma_1^2)$,
$y_1, y_2, \ldots, y_{n_2}$ značí hodnoty náhodného výběru z $N(\mu_2, \sigma_2^2)$,
$\overline{x}$, $\overline{y}$, s_x^2 a s_y^2 odpovídající hodnoty výběrových průměrů a rozptylů.

Testujeme hypotézu, že parametr μ_1 je roven hodnotě μ_2 $(\sigma_1^2 \neq \sigma_2^2)$:

$$H : \mu_1 = \mu_2,$$

testové kriterium je statistika

$$t = \frac{\overline{x} - \overline{y}}{\sqrt{\frac{s_x^2}{n_1} + \frac{s_y^2}{n_2}}},$$

která má při platnosti nulové hypotézy H přibližně Studentovo rozdělení s ν stupni volnosti.

Stupně volnosti daného rozdělení určíme ze vztahu

$$\nu \approx \frac{\left(\frac{s_x^2}{n_1} + \frac{s_y^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_x^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_y^2}{n_2}\right)^2}$$

zaokrouhleno dolů na nejbližší celé číslo.

Podle alternativní hypotézy volíme následující kritické obory:

alternativní hypotéza	**kritický obor**
$A: \mu_1 > \mu_2$	$W_\alpha = \{t, t \geq t_{1-\alpha}(\nu)\}$
$A: \mu_1 < \mu_2$	$W_\alpha = \{t, t \leq -t_{1-\alpha}(\nu)\}$
$A: \mu_1 \neq \mu_2$	$W_\alpha = \left\{t, \lvert t\rvert \geq t_{1-\frac{\alpha}{2}}(\nu)\right\}$

kde $t_{1-\alpha}(\nu)$, $t_{1-\frac{\alpha}{2}}(\nu)$ jsou kvantily Studentova rozdělení, stupně volnosti ν se určí pomocí zmíněného vztahu.

Nechť $X_1, X_2, \ldots, X_{n_1}$ je náhodný výběr z rozdělení se střední hodnotou μ_1 a $Y_1, Y_2, \ldots, Y_{n_2}$ je náhodný výběr z rozdělení se střední hodnotou μ_2. Dále předpokládejme, že tyto výběry jsou *nezávislé* a rozsahy n_1 a n_2 dostatečně velké.

Označme $\overline{X}$ a $\overline{Y}$ odpovídající výběrové průměry, S_X^2 a S_Y^2 odpovídající výběrové rozptyly. Statistika

$$U = \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{S_X^2}{n_1} + \frac{S_Y^2}{n_2}}},$$

má přibližně normální rozdělení $N(0, 1)$.

Nechť $X_1, X_2, \ldots, X_{n_1}$ je náhodný výběr z rozdělení se střední hodnotou μ_1 a $Y_1, Y_2, \ldots, Y_{n_2}$ je náhodný výběr z rozdělení se střední hodnotou μ_2. Dále předpokládejme, že tyto výběry jsou *nezávislé* a rozsahy n_1 a n_2 dostatečně velké.

Označme $\overline{X}$ a $\overline{Y}$ odpovídající výběrové průměry, S_X^2 a S_Y^2 odpovídající výběrové rozptyly. Statistika

$$U = \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{S_X^2}{n_1} + \frac{S_Y^2}{n_2}}},$$

má přibližně normální rozdělení $N(0, 1)$.

Nechť $x_1, x_2, \ldots, x_{n_1}$ značí hodnoty náhodného výběru z prvního rozdělení, $y_1, y_2, \ldots, y_{n_2}$ značí hodnoty náhodného výběru z druhého rozdělení, $\overline{x}$, $\overline{y}$, s_x^2 a s_y^2 jsou odpovídající hodnoty výběrových průměrů a rozptylů.

Testujeme hypotézu, že parametr μ_1 je roven hodnotě μ_2:

$$H : \mu_1 = \mu_2,$$

testové kriterium je statistika

$$u = \frac{\overline{x} - \overline{y}}{\sqrt{\frac{s_x^2}{n_1} + \frac{s_y^2}{n_2}}},$$

která má při platnosti nulové hypotézy H přibližně normální rozdělení $N(0, 1)$.

Test shody středních hodnot dvou výběrů velkého rozsahu

Nechť $x_1, x_2, \ldots, x_{n_1}$ značí hodnoty náhodného výběru z prvního rozdělení, $y_1, y_2, \ldots, y_{n_2}$ značí hodnoty náhodného výběru z druhého rozdělení, $\overline{x}$, $\overline{y}$, s_x^2 a s_y^2 jsou odpovídající hodnoty výběrových průměrů a rozptylů.

Testujeme hypotézu, že parametr μ_1 je roven hodnotě μ_2:

$$H : \mu_1 = \mu_2,$$

testové kriterium je statistika

$$u = \frac{\overline{x} - \overline{y}}{\sqrt{\frac{s_x^2}{n_1} + \frac{s_y^2}{n_2}}},$$

která má při platnosti nulové hypotézy H přibližně normální rozdělení $N(0, 1)$.

Podle alternativní hypotézy volíme následující kritické obory:

alternativní hypotéza	kritický obor
$A: \mu_1 > \mu_2$	$W_\alpha = \{u, u \geq u_{1-\alpha}\}$
$A: \mu_1 < \mu_2$	$W_\alpha = \{u, u \leq -u_{1-\alpha}\}$
$A: \mu_1 \neq \mu_2$	$W_\alpha = \left\{u, \lvert u \rvert \geq u_{1-\frac{\alpha}{2}}\right\}$

kde $u_{1-\alpha}$, $u_{1-\frac{\alpha}{2}}$ jsou kvantily rozdělení $N(0,1)$.

Párový test

Uvažujme situaci, kdy ve výběru o rozsahu n spolu vždy 2 měření určitým způsobem souvisí. (např. měření na jednom prvku je provedeno dvakrát za různých podmínek). Uvažujeme tedy dvě **závislé** náhodné veličiny X a Y se středními hodnotami μ_1 a μ_2, u kterých nás budou zajímat jejich diference $D = X - Y$. Předpokládejme, že máme náhodný výběr $D_1, D_2, \ldots, D_n$, kde diference $D_i = X_i - Y_i$ mají normální rozdělení $N(\mu, \sigma^2)$, kde $\mu = \mu_1 - \mu_2$ (σ^2 není třeba znát).

Statistika

$$T = \frac{\overline{D} - \mu}{S_D}\sqrt{n},$$

kde $\overline{D}$ je výběrový průměr diferencí a S_D je výběrová směrodatná odchylka diferencí, má potom Studentovo rozdělení s $\nu = n - 1$ stupni volnosti.

Párový test

Uvažujme situaci, kdy ve výběru o rozsahu n spolu vždy 2 měření určitým způsobem souvisí. (např. měření na jednom prvku je provedeno dvakrát za různých podmínek). Uvažujeme tedy dvě **závislé** náhodné veličiny X a Y se středními hodnotami μ_1 a μ_2, u kterých nás budou zajímat jejich diference $D = X - Y$. Předpokládejme, že máme náhodný výběr $D_1, D_2, \ldots, D_n$, kde diference $D_i = X_i - Y_i$ mají normální rozdělení $N(\mu, \sigma^2)$, kde $\mu = \mu_1 - \mu_2$ (σ^2 není třeba znát).
Statistika

$$T = \frac{\overline{D} - \mu}{S_D}\sqrt{n},$$

kde $\overline{D}$ je výběrový průměr diferencí a S_D je výběrová směrodatná odchylka diferencí, má potom Studentovo rozdělení s $\nu = n - 1$ stupni volnosti.

Párový test

Nechť $d_1 = x_1 - y_1, d_2 = x_2 - y_2, \ldots, d_n = x_n - y_n$ jsou naměřené hodnoty diferencí, $\overline{d}$ je jejich průměr a s_d jejich výběrová směrodatná odchylka.

Testujeme hypotézu, že parametr μ_1 je roven hodnotě μ_2:

$$H : \mu_1 = \mu_2,$$

testové kriterium je statistika

$$t = \frac{\overline{d}}{s_d}\sqrt{n},$$

která má při platnosti nulové hypotézy H Studentovo rozdělení s $\nu = n - 1$ stupni volnosti.

Párový test

Nechť $d_1 = x_1 - y_1, d_2 = x_2 - y_2, \dots, d_n = x_n - y_n$ jsou naměřené hodnoty diferencí, $\overline{d}$ je jejich průměr a s_d jejich výběrová směrodatná odchylka.

Testujeme hypotézu, že parametr μ_1 je roven hodnotě μ_2:

$$H : \mu_1 = \mu_2,$$

testové kriterium je statistika

$$t = \frac{\overline{d}}{s_d}\sqrt{n},$$

která má při platnosti nulové hypotézy H Studentovo rozdělení s $\nu = n - 1$ stupni volnosti.

Podle alternativní hypotézy volíme následující kritické obory:

alternativní hypotéza	**kritický obor**
$A: \mu_1 > \mu_2$	$W_\alpha = \{t, t \geq t_{1-\alpha}(\nu)\}$
$A: \mu_1 < \mu_2$	$W_\alpha = \{t, t \leq -t_{1-\alpha}(\nu)\}$
$A: \mu_1 \neq \mu_2$	$W_\alpha = \left\{t, \lvert t\rvert \geq t_{1-\frac{\alpha}{2}}(\nu)\right\}$

kde $t_{1-\alpha}(\nu)$, $t_{1-\frac{\alpha}{2}}(\nu)$ jsou kvantily Studentova rozdělení, $\nu = n - 1$.

Nechť $X_1, X_2, \ldots, X_{n_1}$ je náhodný výběr z alternativního rozdělení s parametrem π_1 a $Y_1, Y_2, \ldots, Y_{n_2}$ je náhodný výběr z alternativního rozdělení s parametrem π_2. Dále předpokládejme, že tyto výběry jsou *nezávislé* a rozsahy n_1 a n_2 dostatečně velké.

Označme $P_1 = \sum_{i=1}^{n_1} X_i/n_1, P_2 = \sum_{i=1}^{n_2} Y_i/n_2$ odhady pravděpodobností (podílů) π_1 a π_2.

Statistika

$$U = \frac{P_1 - P_2 - (\pi_1 - \pi_2)}{\sqrt{\frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}}},$$

má přibližně normální rozdělení $N(0, 1)$.

Test shody dvou parametrů alternativního rozdělení pro velké nezávislé výběry

Nechť $X_1, X_2, \ldots, X_{n_1}$ je náhodný výběr z alternativního rozdělení s parametrem π_1 a $Y_1, Y_2, \ldots, Y_{n_2}$ je náhodný výběr z alternativního rozdělení s parametrem π_2. Dále předpokládejme, že tyto výběry jsou *nezávislé* a rozsahy n_1 a n_2 dostatečně velké.

Označme $P_1 = \sum_{i=1}^{n_1} X_i / n_1, P_2 = \sum_{i=1}^{n_2} Y_i / n_2$ odhady pravděpodobností (podílů) π_1 a π_2.

Statistika

$$U = \frac{P_1 - P_2 - (\pi_1 - \pi_2)}{\sqrt{\frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}}},$$

má přibližně normální rozdělení $N(0,1)$.

Test shody dvou parametrů alternativního rozdělení pro velké nezávislé výběry

Nechť $x_1, x_2, \dots, x_{n_1}$ značí hodnoty náhodného výběru z prvního alternativního rozdělení, $y_1, y_2, \dots, y_{n_2}$ značí hodnoty náhodného výběru z druhého alternativního rozdělení, $p_1 = \sum_{i=1}^{n_1} x_i / n_1$, $p_2 = \sum_{i=1}^{n_2} x_i / n_2$ jsou odpovídající odhady.

Testujeme hypotézu, že parametr π_1 je roven hodnotě π_2:

$$H : \pi_1 = \pi_2,$$

testové kriterium je statistika

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

která má při platnosti nulové hypotézy H přibližně normální rozdělení $N(0, 1)$.

Test shody dvou parametrů alternativního rozdělení pro velké nezávislé výběry

Nechť $x_1, x_2, \ldots, x_{n_1}$ značí hodnoty náhodného výběru z prvního alternativního rozdělení, $y_1, y_2, \ldots, y_{n_2}$ značí hodnoty náhodného výběru z druhého alternativního rozdělení, $p_1 = \sum_{i=1}^{n_1} x_i / n_1$, $p_2 = \sum_{i=1}^{n_2} x_i / n_2$ jsou odpovídající odhady.

Testujeme hypotézu, že parametr π_1 je roven hodnotě π_2:

$$H : \pi_1 = \pi_2,$$

testové kriterium je statistika

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

která má při platnosti nulové hypotézy H přibližně normální rozdělení $N(0, 1)$.

Podle alternativní hypotézy volíme následující kritické obory:

alternativní hypotéza	**kritický obor**
$A: \pi_1 > \pi_2$	$W_\alpha = \{u, u \geq u_{1-\alpha}\}$
$A: \pi_1 < \pi_2$	$W_\alpha = \{u, u \leq -u_{1-\alpha}\}$
$A: \pi_1 \neq \pi_2$	$W_\alpha = \left\{u, \lvert u \rvert \geq u_{1-\frac{\alpha}{2}}\right\}$

kde $u_{1-\alpha}$, $u_{1-\frac{\alpha}{2}}$ jsou kvantily rozdělení $N(0,1)$.

Základní literatura

- MANN, P.S. Introductory Statistics. 6th edition. Hoboken: Wiley, 2007. ISBN 978-0-471-75530-2.
- MOUČKA, J., RÁDL, P. Matematika pro studenty ekonomie. 1. vyd. Grada 2010. ISBN 978-80-247-3260-2.
- NEUBAUER, J., SEDLAČÍK, M., KŘÍŽ, O. Základy statistiky – Aplikace v technických a ekonomických oborech. Grada 2012.ISBN: 978-80-247-4273-1.
- ŘEZANKOVÁ, H. Analýza dat z dotazníkových šetření. 2. vydání, Professional Publishing, 2010. ISBN: 9788074310195.

Doporučená literatura

- AGRESTI, A. Categorical Data Analysis. Second Edition. Wiley 2002. ISBN: 0-471-36093-7.
- ANDĚL, J. Statistické metody. 3. vydání. Praha: Matfyzpress, 2003. ISBN 80-86732-08-8.
- ANDĚL, J. Základy matematické statistiky. 2. vyd. Praha: Matfyzpress, 2007, 358 s. ISBN 978-80-7378-001-2.
- VÁGNER, M. Integrální počet funkcí jedné proměnné. 1. vydání. Brno: UO, 2005,126 s. ISBN 80-7231-025-9.
- VÁGNER, M., KAŠTÁNKOVÁ, V. Posloupnosti a řady. 1. vydání. Brno: UO, 2006. ISBN 80-7231-131-X.