Matematika IV – 14. přednáška
Testování hypotéz

Michal Bulant

Masarykova univerzita
Fakulta informatiky

19. 5. 2008

Obsah přednášky

Doporučené zdroje

- Martin Panák, Jan Slovák, **Drsná matematika**, *e*-text.
- Karel Zvára, Josef Štěpán, **Pravděpodobnost a matematická statistika**, Matfyzpress, 4. vydání, 2006, 230 stran, ISBN 80-867-3271-1.
- Marie Budíková, **Statistika** , Masarykova univerzita, ESF, distanční studijní opora, Brno 2004, 176 stran, `http://www.math.muni.cz/~budikova/esf/Statistika.zip`

Doporučené zdroje

- Martin Panák, Jan Slovák, **Drsná matematika**, *e*-text.
- Karel Zvára, Josef Štěpán, **Pravděpodobnost a matematická statistika**, Matfyzpress, 4. vydání, 2006, 230 stran, ISBN 80-867-3271-1.
- Marie Budíková, **Statistika** , Masarykova univerzita, ESF, distanční studijní opora, Brno 2004, 176 stran, `http://www.math.muni.cz/~budikova/esf/Statistika.zip`
- Marie Budíková, Štěpán Mikoláš, Pavel Osecký, **Teorie pravděpodobnosti a matematická statistika (sbírka příkladů)**, Masarykova univerzita, 3. vydání, 2004, 117 stran, ISBN 80-210-3313-4.

Plán přednášky

Motivační úvod

Testování hypotéz umožňuje na základě náhodného výběru s danou pravděpodobností ověřovat domněnky o rozdělení, z něhož pochází daný náhodný výběr.

Motivační úvod

Testování hypotéz umožňuje na základě náhodného výběru s danou pravděpodobností ověřovat domněnky o rozdělení, z něhož pochází daný náhodný výběr. **Hypotézou** budeme rozumět nějaké tvrzení o parametrech tohoto rozdělení.

Motivační úvod

Testování hypotéz umožňuje na základě náhodného výběru s danou pravděpodobností ověřovat domněnky o rozdělení, z něhož pochází daný náhodný výběr. **Hypotézou** budeme rozumět nějaké tvrzení o parametrech tohoto rozdělení.

Definice

H_0 ... nulová hypotéza (např. $\theta = c$, kde c vyjadřuje naši domněnku o hodně parametru θ)
H_1 ... (oboustranná) alternativní hypotéza (obvykle negace nulové)
Testováním H_0 oproti alternativní hypotéze rozumíme postup založený na náhodném výběru, s jehož pomocí platnost H_0 *zamítneme* nebo *nezamítneme* (= připouštíme).

Motivační úvod

Testování hypotéz umožňuje na základě náhodného výběru s danou pravděpodobností ověřovat domněnky o rozdělení, z něhož pochází daný náhodný výběr. **Hypotézou** budeme rozumět nějaké tvrzení o parametrech tohoto rozdělení.

Definice

H_0 ... nulová hypotéza (např. $\theta = c$, kde c vyjadřuje naši domněnku o hodně parametru θ)
H_1 ... (oboustranná) alternativní hypotéza (obvykle negace nulové)
Testováním H_0 oproti alternativní hypotéze rozumíme postup založený na náhodném výběru, s jehož pomocí platnost H_0 *zamítneme* nebo *nezamítneme* (= připouštíme).
Chyba 1. druhu ... H_0 platí a my ji zamítneme (závažnější)
Chyba 2. druhu ... H_0 neplatí a my ji nezamítneme

Motivační úvod

Testování hypotéz umožňuje na základě náhodného výběru s danou pravděpodobností ověřovat domněnky o rozdělení, z něhož pochází daný náhodný výběr. **Hypotézou** budeme rozumět nějaké tvrzení o parametrech tohoto rozdělení.

Definice

H_0 ... nulová hypotéza (např. $\theta = c$, kde c vyjadřuje naši domněnku o hodně parametru θ)
H_1 ... (oboustranná) alternativní hypotéza (obvykle negace nulové)
Testováním H_0 oproti alternativní hypotéze rozumíme postup založený na náhodném výběru, s jehož pomocí platnost H_0 *zamítneme* nebo *nezamítneme* (= připouštíme).
Chyba 1. druhu ... H_0 platí a my ji zamítneme (závažnější)
Chyba 2. druhu ... H_0 neplatí a my ji nezamítneme
Pravděpodobnost chyby 1. druhu se nazývá *hladina významnosti* (α, obvykle $\alpha = 0,05$), pravděpodobnost chyby 2. druhu se značí β a číslo $1 - \beta$ se nazývá *síla testu*.

Způsoby testování nulové hypotézy

1. pomocí intervalu spolehlivosti

Způsoby testování nulové hypotézy

1. pomocí intervalu spolehlivosti
2. pomocí kritického oboru

Způsoby testování nulové hypotézy

1. pomocí intervalu spolehlivosti
2. pomocí kritického oboru
3. pomocí tzv. p–hodnoty (p-value)

Způsoby testování nulové hypotézy

1. pomocí intervalu spolehlivosti
2. pomocí kritického oboru
3. pomocí tzv. p−hodnoty (p-value)

Interval spolehlivosti Na základě realizace náhodného výběru sestrojíme $100(1-\alpha)\%$ interval spolehlivosti pro neznámý parametr θ a zjistíme, zda c patří do tohoto intervalu. Pokud ano, hypotézu H_0 nezamítáme (v opačném případě zamítáme) na hladině významnosti α.

Způsoby testování nulové hypotézy

1. pomocí intervalu spolehlivosti
2. pomocí kritického oboru
3. pomocí tzv. p−hodnoty (p-value)

Interval spolehlivosti Na základě realizace náhodného výběru sestrojíme $100(1-\alpha)\%$ interval spolehlivosti pro neznámý parametr θ a zjistíme, zda c patří do tohoto intervalu. Pokud ano, hypotézu H_0 nezamítáme (v opačném případě zamítáme) na hladině významnosti α.

Kritický obor Stanovení kritického oboru je postup do jisté míry obrácený. Nejprve (i bez náhodného výběru) zvolíme vhodnou statistiku T a množinu hodnot, jichž může T nabývat, rozdělíme na dvě disjunktní podmnožiny: obor nezamítnutí H_0 (značíme V) a **kritický obor** W (obor zamítnutí H_0). Pokud realizace T padne do W, pak H_0 zamítneme, jinak nezamítáme.

Stanovení kritického oboru na hladině α

Pro statistiku T (*testové kritérium*) stanovíme obor nezamítnutí V jako interval, jehož hraniční body tvoří kvantil $\alpha/2$ a $1-\alpha/2$, odtud je

$$W = (-\infty, F^{-1}(\alpha/2)) \cup (F^{-1}(1-\alpha/2), \infty).$$

Způsoby testování nulové hypotézy

p-hodnota Testování pomocí *p*-hodnoty se jednoduchý test, umožněný rozšířením statistických balíků. *p*-hodnota udává nejnižší možnou hladinu významnosti, při níž H_0 zamítáme. Je-li *p*-hodnota $> \alpha$, hypotézu H_0 nezamítáme, pro *p*-hodnotu menší než α, hypotézu zamítneme.

Způsoby testování nulové hypotézy

p-hodnota Testování pomocí *p*-hodnoty se jednoduchý test, umožněný rozšířením statistických balíků. *p*-hodnota udává nejnižší možnou hladinu významnosti, při níž H_0 zamítáme. Je-li *p*-hodnota $> \alpha$, hypotézu H_0 nezamítáme, pro *p*-hodnotu menší než α, hypotézu zamítneme.

Způsoby testování nulové hypotézy

p-hodnota Testování pomocí p-hodnoty se jednoduchý test, umožněný rozšířením statistických balíků. p-hodnota udává nejnižší možnou hladinu významnosti, při níž H_0 zamítáme. Je-li p-hodnota $> \alpha$, hypotézu H_0 nezamítáme, pro p-hodnotu menší než α, hypotézu zamítneme.

p-hodnota se stanoví rovněž se znalostí konkrétní realizace t_0 statistiky T náhodného výběru jako

$$p = 2\min\{P(T \leq t_0), P(T \geq t_0)\}.$$

Testování hypotézy proti jednostranné alternativě

Je-li H_0 hypotéza $\theta = c$, pak *levostranná alternativní hypotéza* je tvrzení $\theta < c$, *pravostranná alternativní hypotéza* je tvrzení $\theta > c$.

Testování hypotézy proti jednostranné alternativě

Je-li H_0 hypotéza $\theta = c$, pak *levostranná alternativní hypotéza* je tvrzení $\theta < c$, *pravostranná alternativní hypotéza* je tvrzení $\theta > c$. Volba typu alternativní hypotézy vyplývá z konkretní situace.

Příklad

- V předmětu Matematika 3 psali studenti písemku rozdělení na 2 skupiny. Hypotéza H_0 : *obě zadání mají stejnou průměrnou obtížnost* je testována oproti oboustranné alternativní hypotéze *zadání nejsou stejně obtížná*.

Testování hypotézy proti jednostranné alternativě

Je-li H_0 hypotéza $\theta = c$, pak *levostranná alternativní hypotéza* je tvrzení $\theta < c$, *pravostranná alternativní hypotéza* je tvrzení $\theta > c$. Volba typu alternativní hypotézy vyplývá z konkretní situace.

Příklad

- V předmětu Matematika 3 psali studenti písemku rozdělení na 2 skupiny. Hypotéza H_0 : *obě zadání mají stejnou průměrnou obtížnost* je testována oproti oboustranné alternativní hypotéze *zadání nejsou stejně obtížná*.
- V předmětu Matematika 3 se dříve po studentech nevyžadovalo řešení domácích úloh. Toto bylo nyní nově zavedeno s cílem dosažení lepších výsledků studentů u závěrečné zkoušky.

Testování hypotézy proti jednostranné alternativě

Je-li H_0 hypotéza $\theta = c$, pak *levostranná alternativní hypotéza* je tvrzení $\theta < c$, *pravostranná alternativní hypotéza* je tvrzení $\theta > c$. Volba typu alternativní hypotézy vyplývá z konkretní situace.

Příklad

- V předmětu Matematika 3 psali studenti písemku rozdělení na 2 skupiny. Hypotéza H_0 : *obě zadání mají stejnou průměrnou obtížnost* je testována oproti oboustranné alternativní hypotéze *zadání nejsou stejně obtížná*.
- V předmětu Matematika 3 se dříve po studentech nevyžadovalo řešení domácích úloh. Toto bylo nyní nově zavedeno s cílem dosažení lepších výsledků studentů u závěrečné zkoušky.

Testování hypotézy proti jednostranné alternativě

Je-li H_0 hypotéza $\theta = c$, pak *levostranná alternativní hypotéza* je tvrzení $\theta < c$, *pravostranná alternativní hypotéza* je tvrzení $\theta > c$. Volba typu alternativní hypotézy vyplývá z konkretní situace.

Příklad

- V předmětu Matematika 3 psali studenti písemku rozdělení na 2 skupiny. Hypotéza H_0 : *obě zadání mají stejnou průměrnou obtížnost* je testována oproti oboustranné alternativní hypotéze *zadání nejsou stejně obtížná*.
- V předmětu Matematika 3 se dříve po studentech nevyžadovalo řešení domácích úloh. Toto bylo nyní nově zavedeno s cílem dosažení lepších výsledků studentů u závěrečné zkoušky.
  V tomto případě zřejmě použijeme nulovou hypotézu H_0 : *výsledné bodové hodnocení se nezlepšilo* oproti pravostranné alternativní hypotéze H_1 : *bodový výsledek studentů se zlepšil*

Jednoduchý příklad

Příklad

Náš protivník hodil 60x kostkou a padla mu 16x šestka. Testujme na hladině významnosti $\alpha = 0,05$ nulovou hypotézu H_0 : *kostka není upravená* oproti jednostranné alternativní hypotézeH_1 : *kostka je upravená tak, aby padalo více šestek.*

Jednoduchý příklad

Příklad

Náš protivník hodil 60x kostkou a padla mu 16x šestka. Testujme na hladině významnosti $\alpha = 0,05$ nulovou hypotézu H_0 : *kostka není upravená* oproti jednostranné alternativní hypotézeH_1 : *kostka je upravená tak, aby padalo více šestek.*

Řešení

Statistika T (*počet šestek*) ma rozdělení $T \sim Bi(60, 1/6)$. Kritický obor je dán 95. percentilem tohoto rozdělení. Snadno vypočteme, že $P(T > 14) = 0,065$ a $P(T > 15) = 0,034$, proto $p-$hodnota rovna 0,034 (nebo jinými slovy: kritickým oborem na hladině 0,05 je interval $\langle 16, \infty)$. Hypotézu H_0 tedy zamítáme – na hladině $0,05$ můžeme tvrdit, že kostka je upravená.

Jednoduchý příklad – pokr.

Řešení (pomocí aproximace)

Porovnejme předchozí řešení příkladu s řešením, při kterém využijeme aproximaci pomocí de Moive-Laplaceovy věty. Náhodnou veličinu

$$X = \frac{T - 10}{\sqrt{50/6}}$$

lze považovat za veličinu mající normální rozdělení $N(\mu, \sigma^2)$ s jednotkovým rozptylem $\sigma^2 = 1$, testovat budeme hypotézu $\mu = 0$.

Jednoduchý příklad – pokr.

Řešení (pomocí aproximace)

Porovnejme předchozí řešení příkladu s řešením, při kterém využijeme aproximaci pomocí de Moive-Laplaceovy věty. Náhodnou veličinu

$$X = \frac{T - 10}{\sqrt{50/6}}$$

lze považovat za veličinu mající normální rozdělení $N(\mu, \sigma^2)$ s jednotkovým rozptylem $\sigma^2 = 1$, testovat budeme hypotézu $\mu = 0$. **Kritickým oborem** $N(0, 1)$ je interval $(1,65, \infty)$ (stále uvažujeme *pravostranou alternativu*). Přitom pro realizaci statistiky X platí $x = (16 - 10)/\sqrt{50/6} \approx 2,08$ a hypotézu tedy opět zamítáme.

Jednoduchý příklad – pokr.

Řešení (pomocí aproximace)

Porovnejme předchozí řešení příkladu s řešením, při kterém využijeme aproximaci pomocí de Moive-Laplaceovy věty. Náhodnou veličinu

$$X = \frac{T - 10}{\sqrt{50/6}}$$

lze považovat za veličinu mající normální rozdělení $N(\mu, \sigma^2)$ s jednotkovým rozptylem $\sigma^2 = 1$, testovat budeme hypotézu $\mu = 0$. **Kritickým oborem** $N(0, 1)$ je interval $(1,65, \infty)$ (stále uvažujeme *pravostranou alternativu*). Přitom pro realizaci statistiky X platí $x = (16 - 10)/\sqrt{50/6} \approx 2,08$ a hypotézu tedy opět zamítáme. **Jednostranným intervalem spolehlivosti** pro X je $((2,08 - 1,65)/\sqrt{60}, \infty)$ a protože do něj nepatří hodnota 0 zamítáme nulovou hypotézu (všimněte si, že v obou případech *rozhodlo* porovnání $1,65 < 2,08$).

Jednoduchý příklad – pokr.

Řešení (pomocí aproximace a p-hodnoty)

Určeme nejmenší pravděpodobnost p, při níž stále ještě zamítáme nulovou hypotézu $\mu = 0$ oproti pravostranné hypotéze $\mu > 0$ (tj. p-hodnotu). Má-li X rozdělení $N(0,1)$, pak
$p = P(X \geq 2,08) = 1 - 0,981 = 0,019.$

Jednoduchý příklad – pokr.

Řešení (pomocí aproximace a p-hodnoty)

Určeme nejmenší pravděpodobnost p, při níž stále ještě zamítáme nulovou hypotézu $\mu = 0$ oproti pravostranné hypotéze $\mu > 0$ (tj. p-hodnotu). Má-li X rozdělení $N(0,1)$, pak
$p = P(X \geq 2,08) = 1 - 0,981 = 0,019.$
Protože je $\alpha = 0,05 > 0,019$, opět hypotézu zamítáme.

Základní testy hypotéz o parametrech normálního rozdělení

Podobně jako statistiky při konstrukci intervalů spolehlivosti jsou i základní testy standardizované (není divu, jak jsme viděli, jde o úzce propojené pojmy).

Základní testy hypotéz o parametrech normálního rozdělení

Podobně jako statistiky při konstrukci intervalů spolehlivosti jsou i základní testy standardizované (není divu, jak jsme viděli, jde o úzce propojené pojmy).

z-test Nechť je $X_1, \dots, X_n$ náhodný výběr z rozdělení $N(\mu, \sigma^2)$ se známým σ^2 a $n \geq 2$. Test $H_0 : \mu = c$ proti alternativní hypotéze $\mu \neq c$ se nazývá **z-test**.

Základní testy hypotéz o parametrech normálního rozdělení

Podobně jako statistiky při konstrukci intervalů spolehlivosti jsou i základní testy standardizované (není divu, jak jsme viděli, jde o úzce propojené pojmy).

z-test
: Nechť je $X_1, \ldots, X_n$ náhodný výběr z rozdělení $N(\mu, \sigma^2)$ se známým σ^2 a $n \geq 2$. Test $H_0 : \mu = c$ proti alternativní hypotéze $\mu \neq c$ se nazývá **z-test**.

jednovýběrový t-test
: Nechť je $X_1, \ldots, X_n$ náhodný výběr z rozdělení $N(\mu, \sigma^2)$ s neznámým σ^2 a $n \geq 2$. Test $H_0 : \mu = c$ proti alternativní hypotéze $\mu \neq c$ se nazývá **jednovýběrový t-test**.

Základní testy hypotéz o parametrech normálního rozdělení

Podobně jako statistiky při konstrukci intervalů spolehlivosti jsou i základní testy standardizované (není divu, jak jsme viděli, jde o úzce propojené pojmy).

z-test
: Nechť je $X_1, \dots, X_n$ náhodný výběr z rozdělení $N(\mu, \sigma^2)$ se známým σ^2 a $n \geq 2$. Test $H_0 : \mu = c$ proti alternativní hypotéze $\mu \neq c$ se nazývá **z-test**.

jednovýběrový t-test
: Nechť je $X_1, \dots, X_n$ náhodný výběr z rozdělení $N(\mu, \sigma^2)$ s neznámým σ^2 a $n \geq 2$. Test $H_0 : \mu = c$ proti alternativní hypotéze $\mu \neq c$ se nazývá **jednovýběrový t-test**.

dvouvýběrový t-test
: Nechť je $X_{11}, \dots, X_{m1}$ náhodný výběr z rozdělení $N(\mu_1, \sigma^2)$ a $X_{12}, \dots, X_{n2}$ na něm nezávislý náhodný výběr z rozdělení $N(\mu_2, \sigma^2)$ s $m, n \geq 2$ a neznámým σ^2. Test $H_0 : \mu_1 - \mu_2 = c$ proti $H_1 : \mu_1 - \mu_2 \neq c$ se nazývá **dvouvýběrový t-test**.

Základní testy hypotéz o parametrech normálním rozdělení

párový t-test Nechť je $(X_1, Y_1)^T, \ldots, (X_n, Y_n)$ výběr z rozdělení

$$N\left(\begin{matrix}\mu_1 \\ \mu_2\end{matrix}, \begin{matrix}\sigma_1^2 & \sigma_{12} \\ \sigma_{12}^2 & \sigma_2^2\end{matrix}\right)$$

s $n \geq 2$ a neznámými parametry. Test $H_0 : \mu_1 - \mu_2 = c$ oproti $H_1 : \mu_1 - \mu_2 \neq c$ se nazývá **párový t-test**.

Základní testy hypotéz o parametrech normálním rozdělení

párový t-test Nechť je $(X_1, Y_1)^T, \dots, (X_n, Y_n)$ výběr z rozdělení

$$N\left(\begin{matrix}\mu_1\\ \mu_2\end{matrix}, \begin{matrix}\sigma_1^2 & \sigma_{12}\\ \sigma_{12}^2 & \sigma_2^2\end{matrix}\right)$$

s $n \geq 2$ a neznámými parametry. Test $H_0 : \mu_1 - \mu_2 = c$ oproti $H_1 : \mu_1 - \mu_2 \neq c$ se nazývá **párový t-test**.

F-test Nechť je $X_{11}, \dots, X_{m1}$ náhodný výběr z rozdělení $N(\mu_1, \sigma_1^2)$ a $X_{12}, \dots, X_{n2}$ na něm nezávislý náhodný výběr z rozdělení $N(\mu_2, \sigma_2^2)$ s $m, n \geq 2$. Test $H_0 : \sigma_1^2/\sigma_2^2 = 1$ proti $H_1 : \sigma_1^2/\sigma_2^2 \neq 1$ se nazývá **F-test**.

Základní testy hypotéz o parametrech normálním rozdělení

párový t-test Nechť je $(X_1, Y_1)^T, \dots, (X_n, Y_n)$ výběr z rozdělení

$$N\left(\begin{matrix}\mu_1 \\ \mu_2\end{matrix}, \begin{matrix}\sigma_1^2 & \sigma_{12} \\ \sigma_{12}^2 & \sigma_2^2\end{matrix}\right)$$

s $n \geq 2$ a neznámými parametry. Test $H_0 : \mu_1 - \mu_2 = c$ oproti $H_1 : \mu_1 - \mu_2 \neq c$ se nazývá **párový t-test**.

F-test Nechť je $X_{11}, \dots, X_{m1}$ náhodný výběr z rozdělení $N(\mu_1, \sigma_1^2)$ a $X_{12}, \dots, X_{n2}$ na něm nezávislý náhodný výběr z rozdělení $N(\mu_2, \sigma_2^2)$ s $m, n \geq 2$. Test $H_0 : \sigma_1^2/\sigma_2^2 = 1$ proti $H_1 : \sigma_1^2/\sigma_2^2 \neq 1$ se nazývá **F-test**.

test rozptylu Nechť je $X_1, \dots, X_n$ náhodný výběr z $N(\mu, \sigma^2)$ s neznámým μ a $n \geq 2$. Test $H_0 : \sigma^2 = c$ proti $H_1 : \sigma^2 \neq c$ se nazývá **test o rozptylu**.

Kritický obor testů normálního rozdělení

z-test $|(M-c)/(\sigma/\sqrt{n})| \geq u_{1-\alpha/2}$

Kritický obor testů normálního rozdělení

z-test $|(M - c)/(\sigma/\sqrt{n})| \geq u_{1-\alpha/2}$

jednovýběrový t-test $|(M - c)/(S/\sqrt{n})| \geq t_{1-\alpha/2}(n - 1)$

Kritický obor testů normálního rozdělení

z-test $|(M-c)/(\sigma/\sqrt{n})| \geq u_{1-\alpha/2}$

jednovýběrový t-test $|(M-c)/(S/\sqrt{n})| \geq t_{1-\alpha/2}(n-1)$

dvouvýběrový t-test

$$\left|\frac{M_1 - M_2 - c}{S_* \sqrt{\frac{1}{m} + \frac{1}{n}}}\right| \geq t_{1-\alpha/2}(m+n-2)$$

Kritický obor testů normálního rozdělení

z-test $|(M-c)/(\sigma/\sqrt{n})| \geq u_{1-\alpha/2}$

jednovýběrový t-test $|(M-c)/(S/\sqrt{n})| \geq t_{1-\alpha/2}(n-1)$

dvouvýběrový t-test

$$\left|\frac{M_1 - M_2 - c}{S_* \sqrt{\frac{1}{m} + \frac{1}{n}}}\right| \geq t_{1-\alpha/2}(m+n-2)$$

párový t-test sestrojením rozdílu $Z_i = X_i - Y_i$ a $\mu = \mu_1 - \mu_2$ úlohu předvedeme na jednovýběrový $t-$ test

Kritický obor testů normálního rozdělení

z-test $|(M-c)/(\sigma/\sqrt{n})| \geq u_{1-\alpha/2}$

jednovýběrový t-test $|(M-c)/(S/\sqrt{n})| \geq t_{1-\alpha/2}(n-1)$

dvouvýběrový t-test

$$\left|\frac{M_1 - M_2 - c}{S_* \sqrt{\frac{1}{m} + \frac{1}{n}}}\right| \geq t_{1-\alpha/2}(m+n-2)$$

párový t-test sestrojením rozdílu $Z_i = X_i - Y_i$ a $\mu = \mu_1 - \mu_2$ úlohu předvedeme na jednovýběrový $t-$ test

F-test $S_1^2/S_2^2 \leq F_{\alpha/2}(m-1, n-1)$ nebo
$S_1^2/S_2^2 \geq F_{1-\alpha/2}(m-1, n-1)$

Kritický obor testů normálního rozdělení

z-test $|(M-c)/(\sigma/\sqrt{n})| \geq u_{1-\alpha/2}$

jednovýběrový t-test $|(M-c)/(S/\sqrt{n})| \geq t_{1-\alpha/2}(n-1)$

dvouvýběrový t-test

$$\left|\frac{M_1 - M_2 - c}{S_* \sqrt{\frac{1}{m} + \frac{1}{n}}}\right| \geq t_{1-\alpha/2}(m+n-2)$$

párový t-test sestrojením rozdílu $Z_i = X_i - Y_i$ a $\mu = \mu_1 - \mu_2$ úlohu předvedeme na jednovýběrový $t-$ test

F-test $S_1^2/S_2^2 \leq F_{\alpha/2}(m-1, n-1)$ nebo $S_1^2/S_2^2 \geq F_{1-\alpha/2}(m-1, n-1)$

test rozptylu $(n-1)S^2/c \leq \chi^2_{\alpha/2}(n-1)$ nebo $(n-1)S^2/c \geq \chi^2_{1-\alpha/2}(n-1)$

Komplexní příklad na dvouvýběrový t-test

Příklad

Uvažme bodové výsledky studentů z 2. termínu zkoušky předmětu `MB103`, přičemž výsledky testů skupiny A a skupiny B považujme za dva nezávislé výběry z normálního rozdělení. Úkolem je zjistit, jestli výsledky některé ze skupin byly statisticky významně horší. Testujme nulovou hypotézu $H_0 : \mu_1 - \mu_2 = 0$ oproti alternativní hypotéze $H_1 : \mu_1 \neq \mu_2$.

Komplexní příklad na dvouvýběrový t-test

Příklad

Uvažme bodové výsledky studentů z 2. termínu zkoušky předmětu MB103, přičemž výsledky testů skupiny A a skupiny B považujme za dva nezávislé výběry z normálního rozdělení. Úkolem je zjistit, jestli výsledky některé ze skupin byly statisticky významně horší. Testujme nulovou hypotézu $H_0 : \mu_1 - \mu_2 = 0$ oproti alternativní hypotéze $H_1 : \mu_1 \neq \mu_2$.

Řešení

Nejprve pomocí F-testu otestujeme hypotézu o stejných rozptylech, v případě úspěchu poté použijeme dvouvýběrový t-test. Vypočteme základní statistiky:

Komplexní příklad na dvouvýběrový t-test

Příklad

Uvažme bodové výsledky studentů z 2. termínu zkoušky předmětu `MB103`, přičemž výsledky testů skupiny A a skupiny B považujme za dva nezávislé výběry z normálního rozdělení. Úkolem je zjistit, jestli výsledky některé ze skupin byly statisticky významně horší. Testujme nulovou hypotézu $H_0 : \mu_1 - \mu_2 = 0$ oproti alternativní hypotéze $H_1 : \mu_1 \neq \mu_2$.

Řešení

Nejprve pomocí F-testu otestujeme hypotézu o stejných rozptylech, v případě úspěchu poté použijeme dvouvýběrový t-test. Vypočteme základní statistiky:

	rozsah	výb. průměr	výb. rozptyl
A	65	10,48	22,49
B	64	7,21	29,75

Řešení (Komplexní příklad na dvouvýběrový t-test (pokr.))

Dostáváme $S_1^2/S_2^2 = 0,76$ a protože $F(0,025;64;63) = 0,61$, **nezamítáme** hypotézu o rovnosti rozptylů.

Řešení (Komplexní příklad na dvouvýběrový t-test (pokr.))

Dostáváme $S_1^2/S_2^2 = 0,76$ a protože $F(0,025; 64; 63) = 0,61$, **nezamítáme** hypotézu o rovnosti rozptylů. O tomtéž se přesvědčíme i vypočtením intervalu spolehlivosti

$$\left(\frac{S_1^2/S_2^2}{F_{1-\alpha/2}(m-1,n-1)}, \frac{S_1^2/S_2^2}{F_{\alpha/2}(m-1,n-1)}\right) \approx (0,46; 1,24),$$

v němž leží testovaný podíl rozptylů 1.

Řešení (Komplexní příklad na dvouvýběrový t-test (pokr.))

Dostáváme $S_1^2/S_2^2 = 0,76$ a protože $F(0,025; 64; 63) = 0,61$, **nezamítáme** hypotézu o rovnosti rozptylů. O tomtéž se přesvědčíme i vypočtením intervalu spolehlivosti

$$\left(\frac{S_1^2/S_2^2}{F_{1-\alpha/2}(m-1, n-1)}, \frac{S_1^2/S_2^2}{F_{\alpha/2}(m-1, n-1)}\right) \approx (0,46; 1,24),$$

v němž leží testovaný podíl rozptylů 1.
Budeme tedy dále s výběry pracovat s předpokladem, že mají stejný rozptyl a použijeme dvouvýběrový t-test.

Řešení (Komplexní příklad na dvouvýběrový t-test (pokr.))

Vypočteme vážený průměr výběrových rozptylů

$$S_*^2 = \frac{(m-1)S_1^2 + (n-1)S_2^2}{m+n-2} \approx 5,11^2,$$

dále $M_1 - M_2 = 3,27$.

Řešení (Komplexní příklad na dvouvýběrový t-test (pokr.))

Vypočteme vážený průměr výběrových rozptylů

$$S_*^2 = \frac{(m-1)S_1^2 + (n-1)S_2^2}{m+n-2} \approx 5,11^2,$$

dále $M_1 - M_2 = 3,27$. V tabulkách najdeme hodnotu $t_{0,975}(65+64-2) = 1,98$, a protože

$$T = \frac{M_1 - M_2}{S_* \sqrt{\frac{1}{65} + \frac{1}{64}}} \approx 3,64,$$

docházíme k závěru, že můžeme hypotézu o stejné střední hodnotě obou rozdělení (tj. hypotézu $\mu_1 = \mu_2$) **zamítnout** (neboť $3,64 > 1,98$).

Řešení (Komplexní příklad na dvouvýběrový t-test (pokr.))

Vypočteme vážený průměr výběrových rozptylů

$$S_*^2 = \frac{(m-1)S_1^2 + (n-1)S_2^2}{m+n-2} \approx 5,11^2,$$

dále $M_1 - M_2 = 3,27$. V tabulkách najdeme hodnotu $t_{0,975}(65+64-2) = 1,98$, a protože

$$T = \frac{M_1 - M_2}{S_* \sqrt{\frac{1}{65} + \frac{1}{64}}} \approx 3,64,$$

docházíme k závěru, že můžeme hypotézu o stejné střední hodnotě obou rozdělení (tj. hypotézu $\mu_1 = \mu_2$) **zamítnout** (neboť $3,64 > 1,98$).Toto opět ověříme výpočtem intervalu spolehlivosti, který má střed v $M_1 - M_2$ a velikost rovnou dvojnásobku $S_* \sqrt{\frac{1}{m} + \frac{1}{n}} t_{1-\alpha/2}(m+n-2) \approx 1,78$, proto je interval spolehlivosti $(1,49; 5,05)$.