Teorie chyb a vyrovnávací počet 1

Téma č. 12: **Testování statistických hypotéz 1.**

1. Obecný postup testování statistických hypotéz.
2. Ověření hypotézy o typu rozdělení pravděpodobnosti
   1. Pearsonův $\chi 2$ - test pro jeden výběr
   2. Kolmogorovův - Smirnovův test pro jeden výběr
3. Testování normality náhodných výběrů
   1. Šikmost
   2. Špičatost
   3. D’Agostinův (omnibus) test
4. Testování střední hodnoty
   1. Při známé základní směrodatné odchylce
   2. Při neznámé základní směrodatné odchylce
5. Ověření hypotézy o rovnosti středních hodnot dvou normálně rozdělených základních souborů
6. Testování hypotézy o shodnosti výběrové a základní směrodatné odchylce
7. Testování hypotézy o shodnosti dvou výběrových směrodatných odchylek
8. Závěrem

1. Obecný postup testování statistických hypotéz (opakování).

Vypočítané či změřené hodnoty často potřebujeme porovnat s nějakým parametrem, ověřit statistickou shodnost či naopak. K tomu používáme statistické testy.

Obecný postup testování statistických hypotéz:

1. Formulace nulové hypotézy H_0 a alternativní hypotézy H_1,
2. volba hladiny významnosti α,
3. volba testovacího kritéria,
4. určení rozdělení pravděpodobnosti testovacího kritéria a výpočet kritických hodnot (oboustranných nebo jednostranných) pro hladinu významnosti α,
5. porovnání vypočtené a kritické hodnoty testovacího kritéria a vyslovení závěru o testované hypotéze H_0.

Nulová hypotéza H_0 je předpoklad o existenci základního souboru s jistým parametrem Θ. Závažná je v této souvislosti formulace tzv. alternativní hypotézy H_1, tj. hypotézy, kterou přijmeme, když neplatí nulová hypotéza. Je rozhodující pro určení jednostranné nebo oboustranné kritické hodnoty testovacího kritéria.

1. Obecný postup testování statistických hypotéz (opakování).

Formulace hypotéz o neznámém parametru Θ může být:

$a)$ $H_0: \Theta = \Theta_0, H_1: \Theta \neq \Theta_0$, pak přichází v úvahu oboustranný test,

$b)$ $H_0: \Theta = \Theta_0, H_1: \Theta > \Theta_0$, nebo $H_0: \Theta = \Theta_0, H_1: \Theta < \Theta_0$, pak v obou případech použijeme jednostranný test.

Testovací kritérium je zvolená funkce, obsahující testovaný výběrový parametr.
Hladina významnosti α je pravděpodobnost, že hodnota testovacího kritéria překročí určenou kritickou hodnotu. Prakticky se nejčastěji volí $\alpha = 0{,}05$ nebo $\alpha = 0{,}01$. Hodnoty testovacího kritéria, které se vyskytnou s pravděpodobností menší než α se nazývají statisticky nevýznamné.

Při testování nulové hypotézy se můžeme dopustit dvou druhů chyb:

a) chyby prvního druhu, tj. chyby, že zamítáme nulovou hypotézu, ačkoliv je ve skutečnosti správná - její pravděpodobnost je právě hladina významnosti α,
b) chyby druhého druhu, tj. chyby, že nezamítáme nulovou hypotézu, ačkoliv je nesprávná - její pravděpodobnost označíme β.

O vzájemném vztahu obou druhů chyb platí, že za nezměněných podmínek snižování pravděpodobnosti chyby jednoho druhu vede ke zvyšování pravděpodobnosti chyby druhého druhu.

2. Ověření hypotézy o typu rozdělení pravděpodobnosti.

Pearsonův χ2 - test pro jeden výběr (opakování)

Viz. Předchozí přednášky – rozdělení do tříd, porovnání třídní a teoretické četnosti, testovací kritérium

$$\chi^2 = \sum_{j=1}^{k} \frac{\left(r_j - R_j\right)^{\mathbf{2}}}{R_j}$$

Nulovou hypotézou je předpoklad, že se skutečné a očekávané třídní četnosti liší pouze náhodně, tj. že hodnocený výběr je výběrem ze základního souboru se zvoleným rozdělením pravděpodobnosti.

H_0: Rozdělení odpovídá s P = 100% - α.
H_1: Rozdělení neodpovídá s rizikem α.

2. Ověření hypotézy o typu rozdělení pravděpodobnosti.

Kolmogorovův - Smirnovův test pro jeden výběr

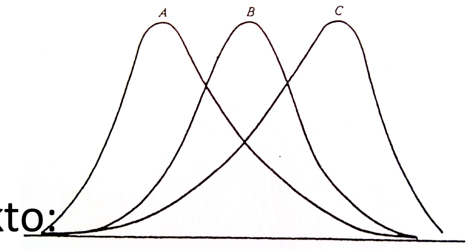
Prvky hodnoceného náhodného výběru roztřídíme do intervalů. Zjistíme skutečné třídní četnosti r_j, tyto hodnoty postupně zleva sčítáme a dostaneme skutečné sčítané četnosti $\tilde{r}_j$. Ze zvoleného rozdělení, které předpokládáme, že platí pro základní soubor, můžeme vypočítat očekávané sčítané třídní četnosti $\tilde{R}_j$, tak, že zleva postupně sčítáme očekávané třídní četnosti R_j. Nulová hypotéza je stejná jako v předchozím případě, test se používá jako pravostranný.

$$D = \frac{1}{n} \max\left|\tilde{r}_j - \tilde{R}_j\right|$$

kde n je celkový počet prvků v souboru a výraz $\max\left|\tilde{r}_j - \tilde{R}_j\right|$ značí hodnotu rozdílu sčítaných skutečných a očekávaných četností v tom intervalu, kde je jeho absolutní velikost maximální. Testovací kritérium má své speciální rozdělení, závislé na rozsahu výběru n.

$\alpha \backslash n$	10	15	20	25	30	35	40	n > 40
0,05	0,40	0,34	0,29	0,26	0,24	0,22	0,21	1,36/$\sqrt{n}$
0,01	0,49	0,40	0,35	0,32	0,29	0,27	0,25	1,63/$\sqrt{n}$

3. Testování normality náhodných výběrů.



Šikmost

šikmost β_3 je třetí standardizovaný (centrální) moment definovaný takto:

$$\beta_3 = \frac{E(x-\mu)^3}{(E(x-\mu)^2)^{3/2}} = \frac{E(x-\mu)^3}{\sigma^3}$$

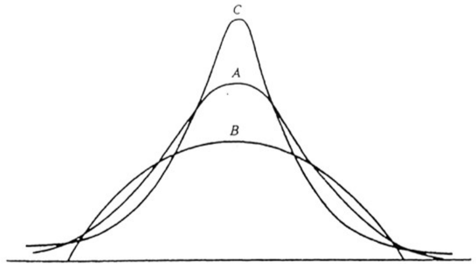
Pro testovaný výběrový soubor lze vypočítat výběrovou hodnotu b_3:

$$b_3 = \frac{m_3}{m_2^{3/2}}\text{, kde } m_k = \frac{\sum(x_i-\bar{x})^k}{n} = \frac{\sum(v_i)^k}{n} \text{ a } \bar{x} = \frac{\sum(x_i)}{n}.$$

Lze testovat přímo (speciální rozdělení) nebo pro více než 8 hodnot po transformaci na veličinu $Z(b_3)$ testovat pomocí normálního rozdělení:

$$Y = b_3 \cdot \left\{\frac{(n+1)\cdot(n+3)}{6\cdot(n-2)}\right\}^{\frac{1}{2}}, \beta_2(b_3) = \frac{3\cdot(n^2+27\cdot n-70)\cdot(n+1)\cdot(n+3)}{(n-2)\cdot(n+5)\cdot(n+7)\cdot(n+9)}, W^2 = -1 + \{2\cdot(\beta_2(b_3) - 1)\}^{\frac{1}{2}} \quad \delta = \frac{1}{\sqrt{\ln(W)}}, \alpha = \left\{\frac{2}{W^2-1}\right\}^{\frac{1}{2}}, \qquad Z(b_3) = \delta \cdot \ln\left(\frac{Y}{\alpha} + \left\{\left(\frac{Y}{\alpha}\right)^2 + 1\right\}^{\frac{1}{2}}\right).$$

Pokud pro oboustranný test platí s vybraným rizikem α výraz: $|Z(b_3)| < t_{\alpha/2}$, nezamítne se nulová hypotéza.



3. Testování normality náhodných výběrů.

Špičatost

špičatost je definována:

$$\beta_4 = \frac{E(x-\mu)^4}{(E(x-\mu)^2)^2} = \frac{E(x-\mu)^4}{\sigma^4}$$

Testovat lze opět přímo nebo pro výběry větší než 20 transformovat na $Z(b_4)$ - přibližně normální rozdělení.

$$E(b_4) = \frac{3\cdot(n-1)}{n+1}, var(b_4) = \frac{24\ (n-2)\cdot(n-3)}{(2+1)^2\cdot(n+3)\cdot(n+5)}, \quad x = \frac{b_4 - E(b_4)}{\sqrt{var(b_4)}},$$

$$\sqrt{\beta_1(b_4)} = \frac{6\cdot\left(n^2-5\cdot n+2\right)}{(n+7)\cdot(n+9)} \cdot \sqrt{\frac{6\cdot(n+3)\cdot(n+5)}{n\cdot(n-2)\cdot(n-3)}}, A = 6 + \frac{8}{\sqrt{\beta_1(b_4)}} \cdot \left[\frac{2}{\beta_1(b_4)} + \sqrt{1+\frac{4}{\beta_1(b_4)}}\right]$$

$$Z(b_4) = \left(\left(1 - \frac{2}{9\cdot A}\right) - \left[\frac{1-\frac{2}{A}}{1+x\cdot\sqrt{\frac{2}{A-4}}}\right]^{\frac{1}{3}}\right) / \sqrt{\frac{2}{9\cdot A}}.$$

Nulovou hypotézou je zde předpoklad normálního rozdělení na základě šikmosti, alternativní hypotézou nenormální rozdělení. Pokud pro oboustranný test platí s vybraným rizikem α výraz:

$$|Z(b_4)| < t_{\alpha/2},$$

3. Testování normality náhodných výběrů.

D'Agostinův (omnibus) test

Jedná se o souhrnný test, kdy se normalita posuzuje současně jak pomocí šikmosti, tak pomocí špičatosti. Testovací kritérium má tvar:

$$K^2 = Z^2(b_3) + Z^2(b_4),$$

kde $Z(b_3)$ a $Z(b_4)$ jsou normální aproximace šikmosti a špičatosti. Jestliže má testovaný soubor normální rozdělení, pak veličina K^2 má přibližně χ^2 rozdělení se dvěma stupni volnosti.

4. Testování střední hodnoty.

Při známé základní směrodatné odchylce

Testujeme hypotézu, že náhodný výběr s výběrovým průměrem x je proveden ze základního souboru se střední hodnotou $E(x) = \bar{x}$. Nulová hypotéza $H_0 : E(x) = \bar{x}$. Nejčastěji se používá alternativní hypotéza $H_1 : E(x) \neq \bar{x}$, tedy provádí se oboustranný test. Testovacím kritériem bude veličina

$$t = \frac{|x-\bar{x}|}{\sigma_x} = \frac{|x-\bar{x}|}{\sigma} \cdot \sqrt{n}\,,\ \sigma_x = \frac{\sigma}{\sqrt{n}}\,,$$

která má normální rozdělení. n je počet opakování měření se základní směrodatnou odchylkou σ. Pro hladinu významnosti α najdeme z tabulek $t_{\alpha/2}$ a v případě:

$t > t_{\alpha/2}$: zamítneme nulovou hypotézu na hladině významnosti α,
$t < t_{\alpha/2}$: nezamítneme nulovou hypotézu na hladině významnosti α.

4. Testování střední hodnoty.

Při neznámé základní směrodatné odchylce

Neznáme základní směrodatnou odchylku σ. V tomto případě ji nahradíme výběrovou:

$$s = \sqrt{\frac{[vv]}{n-1}}, \text{ kde } v_i = x - x_i.$$

Myšlenka i postup testu je obdobný jako v předchozím testu. Testovacím kritériem bude veličina

$$t = \frac{|x-\bar{x}|}{s_x} = \frac{|x-\bar{x}|}{s} \cdot \sqrt{n},$$

která má Studentovo t-rozdělení s $(n-1)$ stupni volnosti. Hodnocení nulové hypotézy bude stejné jako v předchozím případě.

5. Ověření hypotézy o rovnosti středních hodnot dvou normálně rozdělených základních souborů.

Testujeme hypotézu, že dva výběry s výběrovými průměry x_1 a x_2 a výběrovými směrodatnými odchylkami s_1 a s_2 jsou výběry ze dvou základních souborů, pro které platí rovnost jejich středních hodnot $E(x)_1 = E(x)_2$. Test se většinou používá jako oboustranný. Testovací kritérium volíme podle toho, zda rozptyly základních souborů jsou, či nejsou stejné. Musíme tedy nejdříve provést test rozdílu mezi dvěma rozptyly, kterým rozhodneme, zda:

oba rozptyly se významně neliší; pak použijeme testovací kritérium

$$t = \frac{|x_1 - x_2|}{\sqrt{(n_1-1)\cdot s_1^2 + (n_2-1)\cdot s_2^2}} \cdot \sqrt{n_1 + n_2 - 2} \cdot \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}},$$

kde $s_1 = \sqrt{\frac{[vv]_1}{n_1-1}}$, $s_2 = \sqrt{\frac{[vv]_2}{n_2-1}}$

a t má Studentovo rozdělení s $(n_1 + n_2 - 2)$ stupni volnosti. Z tabulek Studentova rozdělení najdeme pro zvolenou hladinu významnosti hodnotu $t_{\alpha/2}$. Nulovou hypotézu budeme zamítat při $t > t_{\alpha/2}$.

5. Ověření hypotézy o rovnosti středních hodnot dvou normálně rozdělených základních souborů.

oba rozptyly se významně liší; pak použijeme testovací kritérium

$$t = \frac{|x_1 - x_2|}{\sqrt{\frac{m_1^2}{n_1} + \frac{m_2^2}{n_2}}}$$

a jeho vypočítanou hodnotu porovnáme s hodnotou

$$t^*_{\frac{\alpha}{2}} = \frac{t_{n_1', \frac{\alpha}{2}} \cdot \frac{m_1^2}{n_1} + t_{n_2', \frac{\alpha}{2}} \cdot \frac{m_2^2}{n_2}}{\frac{m_1^2}{n_1} + \frac{m_2^2}{n_2}}$$, kde $m_1 = \sqrt{\frac{[vv]_1}{n_1 - 1}}$, $m_2 = \sqrt{\frac{[vv]_2}{n_2 - 1}}$,

a $t_{n_1', \alpha/2}$ a $t_{n_2', \alpha/2}$ jsou kritické hodnoty z tabulek Studentova rozdělení pro hladinu významnosti α a stupně volnosti $n_1' = n_1 - 1$ a $n_2' = n_2 - 1$. Nulovou hypotézu zamítáme při $t > t^*_{\alpha/2}$.

6. Testování hypotézy o shodnosti výběrové a základní směrodatné odchylce.

Testujeme hypotézu, že náhodný výběr s výběrovou směrodatnou odchylkou s je proveden se základního souboru se směrodatnou odchylkou σ. Nulová hypotéza $H_0: s = \sigma$. Podle formulace úlohy se používá buď jednostranný, nebo oboustranný test.
Testovacím kritériem bude veličina

$$\chi^2 = \frac{n-1}{\sigma^2} \cdot s^2, \mathrm{s} = \sqrt{\frac{[vv]}{n-1}},$$

která má χ^2-rozdělení s $(n-1)$ stupni volnosti. Pro hladinu významnosti α najdeme z tabulek χ^2-rozdělení kritické hodnoty:

a) pro oboustranný test rozdělíme α na $\alpha/2$ na levé i pravé straně grafu rozdělení a kritické hodnoty budou $\chi^2_{1-\alpha/2}$ a $\chi^2_{\alpha/2}$,
b) pro levostranný test bude kritická hodnota $\chi^2_{1-\alpha}$ na levé straně grafu rozdělení, pro pravostranný test bude kritická hodnota χ^2_{α} na pravé straně grafu rozdělení.

Nulovou hypotézu budeme zamítat, pokud:
a) při oboustranném testu $\chi^2 < \chi^2_{1-\alpha/2}$, nebo$\chi^2 > \chi^2_{\alpha/2}$,
b) při levostranném testu $\chi^2 < \chi^2_{1-\alpha}$,
c) při pravostranném testu $\chi^2 > \chi^2_{\alpha}$.

7. Testování hypotézy o shodnosti dvou výběrových směrodatných odchylek.

Testujeme hypotézu, že dva výběrové rozptyly s_1^2 a s_2^2 ze dvou výběrů o rozsahu n_1 a n_1 odpovídají výběrům ze dvou základních souborů, pro které platí rovnost základních směrodatných odchylek, tedy $\sigma_1 = \sigma_2$. Test se většinou používá jako oboustranný.
Testovacím kritériem bude veličina

$$F = \frac{s_1^2}{s_2^2}\text{, kde } s_1 = \sqrt{\frac{[vv]_1}{n_1-1}},\ s_2 = \sqrt{\frac{[vv]_2}{n_2-1}},$$

která má F-rozdělení s $n_1' = n_1 - 1$ a $n_2' = n_2 - 1$ stupni volnosti. Ve vzorci volíme vždy $s_1^2 > s_2^2$. Z tabulek F-rozdělení najdeme pro zvolenou hladinu významnosti kritickou hodnotu $F_{\alpha/2}$ na pravé straně grafu rozdělení.
Nulovou hypotézu budeme zamítat při $F > F_{\alpha/2}$.

8. Závěrem.

Ke zkoušce:

1. Test – 3 příklady, početní + odvození, max 1 F, každý se známkuje zvlášť.
2. Test – 3 otázky teoretické nebo odvození, max 1 F, každý se známkuje zvlášť.
3. Vyhodnocení, známka, případně ústní dozkoušení.

Známka ze cvičení?

Termín?

☺ Konec ☺