15. Testování hypotéz

Na základě hodnot náhodného výběru činíme rozhodnutí o platnosti hypotézy o hodnotách parametrů rozdělení nebo o jeho vlastnostech. Rozeznáváme dva základní typy testů:

Parametrické testy jsou testy o hodnotách parametrů rozdělení, ze kterého je proveden náhodný výběr.

Neparametrické testy jsou testy o typu rozdělení, shodě rozdělení, symetrii rozdělení.

Testování provádíme na základě funkce náhodného výběru, *statistiky,* jejíž rozdělení je známé a rozhodnutí činíme na základě hodnot této statistiky.

Strategie testování.

1. Na základě hodnot náhodného výběru a charakteru úlohy zvolíme: *nulovou hypotézu* H_0 a *alternativní hypotézu* H_1,
kterou přijímáme v případě odmítnutí nulové hypotézy.

2. Volíme *testovací kritérium.* Vybereme *statistiku,* funkci náhodného výběru, jejíž rozdělení známe a která charakterizuje testovanou vlastnost rozdělení.

3. Stanovíme *hladinu významnosti testu* jako hodnotu α, číslo α je blízké nule. Obvykle z intervalu $(0,01; 0,1)$, nejčastěji $0,05$, která bývá zadavaná ve statistických programech.

4. Na základě hodnoty hladiny, stanovíme *kritický obor* W_α testu, kdy v případě, že zvolená statistika má hodnotu z kritického oboru odmítneme nulovou hypotézu H_0 a příjmeme alternativní hypotézu H_1.

Chyby testu.

Je-li T testovací statistika, α je hladina významnosti testu a W_α je kritický obor testu, pak při rozhodovaní nastanou následující situace.

Skutečnost

	H_0	H_1
H_0	$T \notin W_\alpha$ správně	$T \notin W_\alpha$, chyba 2. druhu $\leq \beta$
H_1	$T \in W_\alpha$, chyba 1. druhu $\leq \alpha$	$T \in W_\alpha$ správně

Stanovení kritického oboru. Požadujeme, aby chyba 1. druhu, kdy odmítneme nulovou hypotézu H_0, ačkoliv platí, byla menší než α. K tomu stačí, aby byl kritický obor W_α doplňkem k $(1-\alpha)100\%$ intervalu spolehlivosti pro testovaný parametr rozdělení. Chybu 2. druhu můžeme pouze odhadnout.

Testy o parametrech rozdělení.

15.1. Test o střední hodnotě, *jednovýběrový t-test.* Předpokládáme, $X_1, X_2, \dots X_n$ je náhodný výběr z normálního rozdělení $N(\mu;\sigma^2)$. Jako odhad střední hodnoty μ použijeme výběrový průměr $\overline{X}$ a jako odhad rozptylu σ^2 použijeme výběrový rozptyl S^2.

a) Testujeme nulovou hypotézu

$H_0: \quad \mu = \mu_0$

proti alternativní hypotéze

$H_1: \quad \mu \neq \mu_0.$

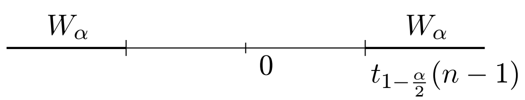
Za testovou statistiku volíme

$$T = \frac{\overline{X} - \mu_0}{S}\sqrt{n},$$

o které je známo, že má Studentovo $t(n-1)$ rozdělení. Kritickým oborem je

$$W_\alpha = \{T;\ |T| > t_{1-\frac{\alpha}{2}}(n-1)\}$$

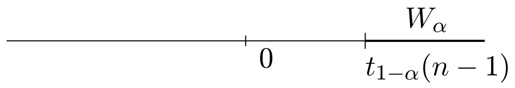
doplněk k $(1-\alpha)100\%$ intervalu spolehlivosti pro parametr μ. Při této volbě je chyba 1. druhu menší než α. To znamená, že ve $100\alpha\%$ případů odmítneme pravdivou skutečnost a příjmeme alternativní hypotézu, ačkoliv neplatí. Situace je znázorněná na obrázku.



Obdobně provádíme test jednostranných hypotéz:

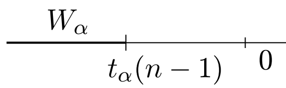
b) $H_0: \quad \mu \leq \mu_0, \qquad H_1: \quad \mu > \mu_0$, pak

$$W_\alpha = \{T;\ T > t_{1-\alpha}(n)\};$$



c) $H_0: \quad \mu \geq \mu_0, \qquad H_1: \quad \mu < \mu_0$, pak

$$W_\alpha = \{T;\ T < t_\alpha(n)\}.$$



Kritické hodnoty testu. Krajní body intervalů, které tvoří kritické obory se nazývají *kritické hodnoty* testu. Označují se symbolem $t(\alpha)$, ačkoliv jsou to $1 - \frac{\alpha}{2}$ kvantily. Při práci s tabulkami je třeba dávat pozor, jak je přesně kritická hodnota definována. V záhlaví tabulky je toto vždy uvedeno.

Poznamenejme, že pro rozsahy výběru $n \geq 30$ můžeme nahradit kvantily, či kritické hodnoty Studentova $t-$rozdělení hodnotami z normovaného normálního rozdělení.

Poznámka: p-hodnota V současné době se využívá možnosti počítačů a rozhodování provádíme pomocí $p-$hodnoty. Ta je stanovena tak, že je to hladina, při které se hraniční hodnotou kritického oboru stává hodnota t_0 testovací statistiky T. Je tedy $p-$hodnota definována jako:

a) $p = 2min\{P(T \leq t_0);\ P(T \geq t_0)\}$ pro oboustranný test;

b) $p = P(T \leq t_0)$ pro levostranný test;

c) $p = P(T \geq t_0)$ pro pravostranný test.

Potom při $p \leq \alpha$ hypotézu H_0 zamítáme a při $p > \alpha$ hypotézu H_0 nezamítáme. Lze říci, že čím je p větší, tím je i chyba 2. druhu menší.

15.2. Test o rozptylu normálního rozdělení. Pro náhodný výběr $X_1, X_2, \ldots X_n$ je náhodný výběr z normálního rozdělení $N(\mu; \sigma^2)$ hledáme hodnotu rozptylu σ^2. Jako jeho odhad použijeme výběrový rozptyl S^2.

a) Testujeme nulovou hypotézu

$H_0: \quad \sigma^2 = \sigma_0^2$

proti alternativní hypotéze

$H_1: \quad \sigma^2 \neq \sigma_0^2$.

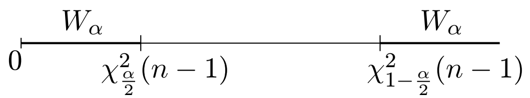
Za testovou statistiku volíme

$$V = \frac{(n-1)S^2}{\sigma_0^2},$$

o které je známo, že má $\chi^2(n-1)$ rozdělení. Kritickým oborem je

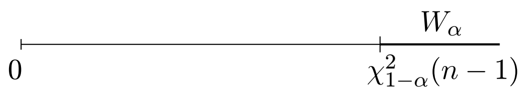
$$W_\alpha = \{V;\ V < \chi^2_{\frac{\alpha}{2}}(n-1) \text{ nebo } V > \chi^2_{1-\frac{\alpha}{2}}(n-1)\}$$

doplněk k $(1-\alpha)100\%$ intervalu spolehlivosti pro parametr σ^2. Při této volbě je chyba 1. druhu menší než α. To znamená, že ve $100\alpha\%$ případů odmítneme pravdivou skutečnost a příjmeme alternativní hypotézu, ačkoliv neplatí. Situace je znázorněná na obrázku.



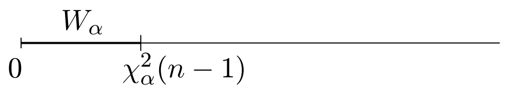
Obdobně provádíme test jednostranných hypotéz:
b) $H_0: \quad \sigma^2 \le \sigma_0^2, \qquad H_1: \quad \sigma^2 > \sigma_0^2$, pak

$$W_\alpha = \{V;\ V > \chi^2_{1-\alpha}(n-1)\};$$



c) $H_0: \quad \sigma^2 \ge \sigma_0^2, \qquad H_1: \quad \sigma^2 < \sigma_0^2$, pak

$$W_\alpha = \{V;\ V < \chi^2_{\alpha}(n-1)\}.$$



15.3. Test pro parametr δ exponenciálního rozdělení $Exp(0;\delta)$.

Pro náhodný výběr $X_1, X_2, \ldots, X_n$ z exponenciálního rozdělení $Exp(0;\delta)$ hledáme hodnotu parametru δ. Testujeme nulovou hypotézu $H_0: \quad \delta = \delta_0$ proti alternativě $H_1: \quad \delta \ne \delta_0$.

Za testovou statistiku volíme

$$T = \frac{2n\overline{X}}{\delta_0},$$

která má rozdělení $\chi^2(2n)$. Kritickým oborem je

$$W_\alpha = \{V;\ V < \chi^2_{\frac{\alpha}{2}}(n-1) \text{ nebo } V > \chi^2_{1-\frac{\alpha}{2}}(n-1)\}$$

doplněk k $(1-\alpha)100\%$ intervalu spolehlivosti pro parametr δ.

15.4. Test o rovnosti středních hodnot, *dvouvýběrový t-test.*

Předpokládáme, že $X_1, X_2, \ldots, X_n$ je náhodný výběr z normálního rozdělení $N(\mu_1; \sigma_1^2)$ a $Y_1, Y_2, \ldots, Y_m$ je náhodný výběr z normálního rozdělení $N(\mu_2; \sigma_2^2)$. Jako odhady středních hodnot μ_1 a μ_2 použijeme výběrové průměry $\overline{X}$ a $\overline{Y}$ a jako odhady rozptylů σ_1^2 a σ_2^2 použijeme výběrové rozptyly S_X^2 a S_Y^2. Předpokládáme, že jsou výběry nezávislé a že se rozptyly rovnají.

Testujeme nulovou hypotézu

$H_0 :\quad \mu_1 - \mu_2 = \Delta$, obvykle $\Delta = 0$,

proti alternativní hypotéze

$H_1 :\quad \mu_1 - \mu_2 \neq \Delta$.

A) **Dvouvýběrový t-test.** Za testovou statistiku volíme

$$T = \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{(n-1)S_X^2 + (m-1)S_Y^2}} \sqrt{\frac{nm(n+m-2)}{n+m}},$$

o které je známo, že má Studentovo $t(n+m-2)$ rozdělení. Kritickým oborem je

$$W_\alpha = \{T;\ |T| > t_{1-\frac{\alpha}{2}}(n+m-2)\}$$

doplněk k $(1-\alpha)100\%$ intervalu spolehlivosti pro parametr Δ. Při této volbě je chyba 1. druhu menší než α. To znamená, že ve $100\alpha\%$ případů odmítneme pravdivou skutečnost a příjmeme alternativní hypotézu, ačkoliv neplatí.

Porušení normality výběru se ve výsledcích testů výrazněji neprojeví. Shodu rozptylů před výpočtem ověříme testem pro jejich rovnost.

Pokud nám test pro rovnost rozptylů dá negativní výsledek, použijeme **Cochranův-Coxův** test nebo neparametrický **dvouvýběrový Wilcoxonův test**.

B) **Cochranův-Coxův test** volíme v případě, že není splněn předpoklad o rovnosti rozptylů. Za testovou statistiku volíme

$$T^* = \frac{\overline{X} - \overline{Y} - \Delta}{S}, \quad S = \sqrt{v_X + v_Y}, \quad v_X = \frac{S_X^2}{n},\ v_Y = \frac{S_Y^2}{m}.$$

Kritickým oborem je

$$W_\alpha = \{T^*;\ |T^*| > t^*\}, \quad t^* = \frac{v_X t_{n-1}(\alpha) + v_Y t_{m-1}(\alpha)}{v_X + v_Y},$$

kde $t_k(\alpha)$ je kritická hodnota jednovýběrového $t-$testu.

Tento test má ještě některé jiné varianty, které pro menší rozsahy výběrů dávají poněkud jiné kritické obory. Uvedeme si na ukázku dvě z nich.

C) **Satterthwaite (1946).**

Kritickým oborem je

$$W_\alpha = \{T^*;\ |T^*| > t_f(\alpha)\}, \quad f = \frac{S^4}{\frac{v_X^2}{n-1} + \frac{v_Y^2}{m-1}},$$

kde $t_k(\alpha)$ je kritická hodnota jednovýběrového $t-$testu.

D) **Welch (1947).**

Kritickým oborem je

$$W_\alpha = \{T^*;\ |T^*| > t_h(\alpha)\}, \quad h = \frac{S^4}{\frac{v_X^2}{n} + \frac{v_Y^2}{m}} - 2,$$

kde $t_k(\alpha)$ je kritická hodnota jednovýběrového $t-$testu.

15.5. Test o rovnosti rozptylů, *F-test.*

Předpokládáme, že $X_1, X_2, \ldots, X_n$ je náhodný výběr z normálního rozdělení $N(\mu_1; \sigma_1^2)$ a $Y_1, Y_2, \ldots, Y_m$ je náhodný výběr z normálního rozdělení $N(\mu_2; \sigma_2^2)$. Jako odhady středních hodnot μ_1 a μ_2 použijeme výběrové průměry $\overline{X}$ a $\overline{Y}$ a jako odhady rozptylů σ_1^2 a σ_2^2 použijeme výběrové rozptyly S_X^2 a S_Y^2. Předpokládáme, že jsou náhodné výběry nezávislé.

Testujeme nulovou hypotézu

$H_0: \quad \sigma_1^2 = \sigma_2^2$

proti alternativní hypotéze

$H_1: \quad \sigma_1^2 \neq \sigma_2^2.$

Jako výběr X_i označíme ten, pro který je $S_X^2 > S_Y^2$. Za testovou statistiku volíme

$$F = \frac{S_X^2}{S_Y^2},$$

o které je známo, že má $F_{n-1,m-1}$ rozdělení. Kritickým oborem je

$$W_\alpha = \{F;\ F > F_{n-1,m-1}(\alpha),\}$$

kde $F_{n-1,m-1}(\alpha)$ je kritická hodnota z tabulek.

Neparametrické testy

V neparametrických testech má hypotéza charakter tvzení o vlastnostech rozdělení, které nejsou odvozeny od hodnot parametrů. Uvedeme některé z nich.

15.6. Znaménkový test je testem o mediánu rozdělení. Používáme jej jako velice jednoduchou variantu testu na symetrii rozdělení, kdy by se měl medián rovnat střední hodnotě.

Předpokládáme, že $X_1, X_2, \ldots, X_n$ je náhodný výběr ze spojitého rozdělení jehož medián je $x_{0,5} = \tilde{x}$. Testujeme nulovou hypotézu

$H_0: \quad \tilde{x} = x_0$, proti alternativě $H_1: \quad \tilde{x} \neq x_0$.

Označme si $Y_i = X_i - x_0$. Pokud je nulová hypotéza platná, pak by měl být počet kladných a záporných hodnot souboru Y_i stejný. Označíme-li Y počet kladných hodnot v souboru Y_i, je pak Y realizací náhodné veličiny, která má binomické rozdělení $Bi(n, \frac{1}{2})$. Ta nabývá hodnot z množiny $\{0, 1, 2, \ldots, n\}$ a hodnoty blízké nule a n se vyskytují s velmi malou pravděpodobností.

Kritický obor testu je

$$W_\alpha = \{Y;\ Y \leq k_1 \text{ nebo } Y \geq k_2\},$$

kde hodnoty k_1 a k_2 nalezneme v tabulkách. Pro zvolenou hladinu testu je nalezneme tak, že je k_1 největší z hodnot a k_2 je nejmenší z hodnot, pro které platí

$$P(Y \leq k_1) \leq \frac{\alpha}{2}, \quad P(Y \geq k_2) \leq \frac{\alpha}{2},$$

jestliže má Y zmiňované binomické rozdělení $Bi(n, \frac{1}{2})$.

Pokud má výběr větší rozsah, $n > 36$, můžeme nahradit binomické rozdělení $Bi(n, \frac{1}{2})$ normálním rozdělením $N(\frac{n}{2}, \frac{n}{4})$, která mají shodné střední hodnoty $\frac{n}{2}$ a rozptyly $\frac{n}{4}$. Potom má náhodná veličina

$$U = \frac{Y - \frac{n}{2}}{\frac{\sqrt{n}}{2}} = \frac{2Y - n}{\sqrt{n}}$$

normované normální rozdělení $N(0;1)$. Kritický obor je roven

$$W_\alpha = \{U;\ |U| \geq u(\alpha),\}$$

kde $u(\alpha)$ je kritická hodnota pro normální rozdělení, kterou nalezneme z tabulek. Poznamenejme, že je tato kritická hodnota $u(\alpha) = u_{1-\frac{\alpha}{2}}$ rovna $1-\frac{\alpha}{2}$kvantil normovaného normálního rozdělení. Snadno odvodíme i jednostranné varianty testu. Test má poměrně malou sílu a k věrohodnotnějšímu výsledku je potřeba poměrně velký rozsah náhodného výběru.

15.7. Jednovýběrový Wilcoxonův test je testem symetrie rozdělení. Testujeme symetrii rozdělení vzhledem k hodnotě x_0, tedy skutečnost, že pro hustotu či pravděpodobnostní funkci platí
$f(x-x_0) = f(x+x_0)$. Nulovou hypotézu zapisujeme ve tvaru podmínky pro medián $x_{0,5} = \tilde{x}$:
$H_0: \quad \tilde{x} = x_0$, proti alternativě $H_1: \quad \tilde{x} \neq x_0$.

Pro náhodný výběr $X_1, X_2, \ldots, X_n$ utvoříme soubor $Y_i = X_i - x_0$, ve kterém vypustíme případné nulové hodnoty. Hodnoty $|Y_i|$ uspořádáme podle velikosti a označíme R_i^+ jejich pořadí. Nyní je

$$S^+ = \sum_{Y_i > 0} R_i^+, \quad S^- = \sum_{Y_i < 0} R_i^+.$$

Poznamenejme, že $S^+ + S^- = \frac{1}{2}n(n+1)$.

Pokud je rozdělení symetrické, budou se vyskytovat kladné a záporné hodnoty souměrně kolem hodnoty x_0, tedy součty pořadí kladných a záporných hodnot se od sebe budou málo lišit. Kritický obor testu je stanoven jako

$$W_\alpha: \quad \min(S^+, S^-) < w(\alpha),$$

kde $w(\alpha)$ je kritická hodnota testu, kterou nalezneme v tabulkách. Je-li splněna podmínka pro kritický obor zamítneme nulovou hypotézu, že rozdělení je symetrické.

Poznamenejme, že pro náhodné veličiny S^+ a S^- je

$$E(S^+) = E(S^-) = \frac{1}{4}n(n+1), \text{ a } D(S^+) = D(S^-) = \frac{1}{24}n(n+1)(2n+1).$$

Pro větší hodnoty rozsahu výběru nahradíme rozdělení rozdělením normálním, tedy skutečností, že má náhodná veličina

$$U = \frac{S^+ - \frac{1}{4}n(n+1)}{\sqrt{\frac{1}{24}n(n+1)(2n+1)}}$$

normované normální rozdělení $N(0;1)$. Kritický obor testu je pak

$$W_\alpha = \{U;\ |U| > u(\alpha),\}$$

kde u_α je kritická hodnota testu pro normální rozdělení, která je rovna $u(\alpha) = u_{1-\frac{\alpha}{2}}$kvantilu normovaného normálního rozdělení.

15.8. Dvouvýběrový Wilcoxonův test slouží k porovnání výběrů, kdy testujeme hypotézu, že jsou oba výběry ze stejného rozdělení.

Předpokládáme, že náhodný výběr $\{X_1, X_2, \ldots, X_n\}$ je výběrem z rozdělení s distribuční funkcí F a náhodný výběr $\{Y_1, Y_2, \ldots, Y_m\}$ je výběrem z rozdělení s distribuční funkcí G. Testujeme hypotézu
$H_0: \quad F = G$ proti alternativě $H_1: \quad F \neq G$.

Test je založen na skutečnosti, že pokud jsou obě rozdělení stejná, pak se v obou výběrech budou vyskytovat hodnoty shodné velikosti ve stejném počtu.

Algoritmus testu:

1. Vytvoříme sdružený soubor
$\{Z_1, Z_2, \ldots, Z_{n+m}\} = \{X_1, X_2, \ldots, X_n\} \cup \{Y_1, Y_2, \ldots, Y_m\}$.

2. Stanovíme pořadí prvků souboru, který uspořádáme podle velikosti, přičemž prvkům, které mají stejnou velikost přiřadíme průměr jejich pořadí. Označme
T_1 − je součet pořadí prvků z prvního souboru; T_2 − je součet pořadí prvků z druhého souboru.

Poznamenejme, že $T_1 + T_2 = \frac{1}{2}(n+m)(n+m+1)$.

3. Položme

$$U_1 = nm + \frac{1}{2}n(n+1) - T_1 \text{ a } U_2 = nm + \frac{1}{2}m(m+1) - T_2. \quad (U_1 + U_2 = nm.)$$

Testovací kritérium: Kritický obor

$$W_\alpha : \quad \min\{U_1, U_2\} \le w(\alpha),$$

kde kritickou hodnotu $w(\alpha)$ testu nalezneme v tabulkách.

Poznámka: Pořadí souborů volíme tak, aby $n \ge m$, tabulky bývají pro rozsahy $2 \le m \le 20,\ 5 \le n \le 30$.

Pro větší rozsahy výběrů využíváme skutečnosti, že za platnosti hypotézy H_0 je

$$E(U_1) = E(U_2) = \frac{1}{2}nm \quad \text{a} \quad D(U_1) = D(U_2) = \frac{1}{12}nm(n+m+1).$$

Rozdělení obou veličin můžeme pak považovat za normální a tedy náhodná veličina

$$U = \frac{U_{1,2} - \frac{1}{2}nm}{\sqrt{\frac{1}{12}nm(n+m)(n+m+1)}}$$

má normované normální rozdělení $N(0;1)$.

Kritický obor testu je

$$W_\alpha = \{U;\ |U| > u(\alpha)\},$$

kde $u(\alpha)$ je kritická hodnota pro normální rozdělení, tedy $u_{1-\frac{\alpha}{2}}$ kvantil normálního rozdělení.

Poznámka. Test je citlivý na posun, tedy na situaci, kdy je
$F(x) = G(x - \Delta)$. Pro situace, kdy se soubory liší spíše rozptylem či tvarem je doporučen *Kolmogorovův-Smirnovův test.*

15.9. Kolmogorovův-Smirnovův test.

Nejprve popíšeme *empirickou distribuční funkci*, která se v testu používá.

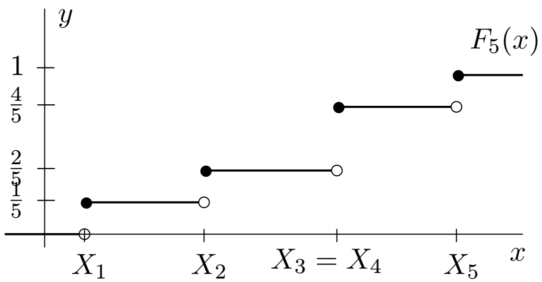
Je-li $\{X_1, X_2, \ldots, X_n\}$ náhodný výběr z rozdělení, které má distribuční funkci F, pak *empirickou distribuční funkcí* nazýváme funkci F_n, která je definována předpisem:

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \xi_i(x), \quad \text{kde} \quad \xi_i(x) = \left\langle \begin{array}{ll} 0, & x < X_i, \\ 1, & x \geq X_i. \end{array} \right.$$

Potom je

$$\lim_{n\to\infty} F_n(x) = F(x), \quad x \in \mathrm{R}.$$

Poznámka. Empirická distribuční funkce je po úsecích konstantní a má skoky velikosti 1 v bodech $x = X_i$, $1 \leq i \leq n$. Znázorníme si průběh empirické distribuční funkce pro náhodný výběr, pro který platí: $X_1 < X_2 < X_3 = X_4 < X_5$.



Obr. 12.1.

Jednovýběrový test

Předpokládáme, že náhodný výběr X_i, $1 \leq i \leq n$ je výběrem z rozdělení s distribuční funkcí F. Testujeme hypotézu:

H_0 : výběr je z rozdělení s distribuční funkcí F

proti alternativní hypotéze

H_1 : výběr není z rozdělení s distribuční funkcí F.

Algoritmus testu:

1. Vypočteme empirickou distribuční funkce F_n a teoretickou dostribuční funkce F.

2. Určíme maximální rozdíl těchto funkcí,

$$D_n = sup\{|F_n(x) - F(x)|;\ x \in \mathrm{R}\}.$$

Platí-li hypotéza H_0 je $\lim\limits_{n,m\to\infty} D_n = 0$.

3. Určíme testovací statistiku

$$\sqrt{n}D_n,$$

která má rozdělení určené distribuční funkcí $K(\lambda)$, kde

$$K(\lambda) = 1 - 2\sum_{k=1}^{\infty}(-1)^{k+1}\mathrm{e}^{-2k^2\lambda^2},$$

tj.

$$\lim_{n,m\to\infty} P\left(\sqrt{n}D_n < \lambda\right) = K(\lambda), \quad \lambda > 0.$$

4. Kritický obor testu je

$$W_\alpha : \quad \sqrt{n}D_n \geq \lambda_\alpha \Leftrightarrow D_n \geq \frac{\lambda_\alpha}{\sqrt{n}},$$

kde kritickou hodnotu testu $D_n^* = \frac{\lambda_\alpha}{\sqrt{n}}$ nalezneme v tabulkách pro hodnoty $2 \leq n \leq 20$. Pro větší rozsahy výběrů použijeme aproximace

$$K(\lambda) \doteq 1 - 2\mathrm{e}^{-2\lambda^2}$$

a kritickou hodnotu λ_α určíme z podmínky:

$$P\left(D_n < \frac{\lambda_\alpha}{\sqrt{n}}\right) = K(\lambda) = 1-\alpha \Rightarrow 1-\alpha = 1-2\mathrm{e}^{-2\lambda_\alpha^2} \Rightarrow \lambda_\alpha = \sqrt{-\frac{1}{2}\ln\frac{2}{\alpha}}.$$

Pro kritický obor dostaneme

$$W_\alpha : \quad D_n \geq D_n^* = \sqrt{\frac{1}{2n}\ln\frac{2}{\alpha}}.$$

Dvouvýběrový test

Předpokládáme, že náhodný výběr $\{X_1, X_2, \ldots, X_n\}$ je výběrem z rozdělení s distribuční funkcí F a náhodný výběr $\{Y_1, Y_2, \ldots, Y_m\}$ je výběrem z rozdělení s distribuční funkcí G. Testujeme hypotézu
$H_0: \quad F = G$ proti alternativě $H_1: \quad F \neq G$.

Test je založen na skutečnosti, že pokud jsou obě rozdělení stejná, pak se v obou výběrech budou vyskytovat hodnoty shodné velikosti ve stejném počtu.

Algoritmus testu:

1. Vypočteme empirické distribuční funkce F_n a G_m.

2. Určíme maximální rozdíl těchto funkcí,

$$D_{n,m} = sup\{|F_n(x) - G_m(x)|;\ x \in \mathrm{R}\}.$$

Platí-li hypotéza H_0 je $\lim\limits_{n,m\to\infty} D_{n,m} = 0$.

3. Určíme testovací statistiku

$$\sqrt{M} D_{n,m}, \quad M = \frac{nm}{n+m},$$

která má rozdělení určené distribuční funkcí $K(\lambda)$, kde

$$K(\lambda) = 1 - 2\sum_{k=1}^{\infty}(-1)^{k+1}\mathrm{e}^{-2k^2\lambda^2},$$

tj.

$$\lim_{n,m\to\infty} P\left(\sqrt{M} D_{n,m} < \lambda\right) = K(\lambda), \quad \lambda > 0.$$

4. Kritický obor testu je

$$W_\alpha: \quad \sqrt{M} D_{n,m} \geq \lambda_\alpha \Leftrightarrow D_{n,m} \geq \frac{\lambda_\alpha}{\sqrt{M}},$$

kde kritickou hodnotu testu $D^*_{n,m} = \frac{\lambda_\alpha}{\sqrt{M}}$ nalezneme v tabulkách pro hodnoty $2 \leq n \leq 20,\ 4 \leq m \leq 20,\ n + m \geq 8$. Pro větší rozsahy výběrů použijeme aproximace

$$K(\lambda) \doteq 1 - 2\mathrm{e}^{-2\lambda^2}$$

a kritickou hodnotu λ_α určíme z podmínky:

$$P\left(D_{n,m} < \frac{\lambda_\alpha}{\sqrt{M}}\right) = K(\lambda) = 1 - \alpha \Rightarrow \lambda_\alpha = \sqrt{-\frac{1}{2}\ln\frac{2}{\alpha}}.$$

Pro kritický obor dostaneme

$$W_\alpha : \quad D_{n,m} \geq D^*_{n,m} = \sqrt{\frac{1}{2M}\ln\frac{2}{\alpha}}.$$

Poznámka: Test je odvozen za podmínky, že rozdělení je normální. V případě, že rozdělení se významně liší od normálního, je nutné použít jiný test. Pro některá rozdělení jsou ve statistických tabulkách uvedeny kritické hodnoty tesu, které jsou odlišné od dříve uvedených. Distribuční funkce K má pro jiná rozdělení odlišný charakter.

15.10. Test shody pro binomické rozdělení. Máme dány hodnoty nezávislých náhodných veličin $X \sim Bi(n, p_1)$ a $Y \sim Bi(m, p_2)$. Testujeme nulovou hypotézu

$$H_0 : \quad p_1 = p_2$$

proti alternativě

$$H_1 : \quad p_1 \neq p_2.$$

Algoritmus testu.

1. Vypočteme hodnoty $x = \frac{X}{n}$ a $y = \frac{Y}{m}$, které jsou odhady parametrů $p_1 \approx x$ a $p_2 \approx y$.

2. Má-li výběr dostatečně velký rozsah, pak mají náhodné veličiny x a y po řadě normální rozdělení

$$x \sim N\left(p_1; \frac{p_1(1-p_1)}{n}\right) \quad \text{a} \quad y \sim N\left(p_2; \frac{p_2(1-p_2)}{m}\right).$$

3. Protože jsou náhodné veličiny x a y nezávislé má náhodná veličina

$$U = \frac{(x-y) - (p_1 - p_2)}{\sqrt{\frac{p_1(1-p_1)}{n} + \frac{p_2(1-p_2)}{m}}}$$

normované normální rozdělení $N(0; 1)$.

4. Pokud platí nulová hypotéza H_0, je $p_1 - p_2 = 0$ a jestliže použijeme aproximací $p_1 = x,\ p_2 = y$, má náhodná veličina

$$U_a = \frac{x-y}{\sqrt{\frac{x(1-x)}{n} + \frac{y(1-y)}{m}}}$$

normované normální rozdělení $N(0;1)$.

5. Kritický obor testu je pak

$$W_\alpha = \{U_a;\ |U_a| \geq u(\alpha)\},$$

kde kritická hodnota $u(\alpha)$ je rovna $1-\frac{\alpha}{2}$–kvantilu normálního rozdělení $N(0;1)$.

Alternativní varianta testu je založena na skutečnosti, že společnou hodnotu $p_1 = p_2$ odhadujeme pomocí hodnoty $z = \frac{X+Y}{n+m} = \frac{nx+my}{n+m}$. Potom má náhodná veličina

$$U_b = \frac{x-y}{\sqrt{z(1-z)\left(\frac{1}{n} + \frac{1}{m}\right)}}$$

normované normální rozdělení $N(0;1)$.

Kritický obor testu je pak

$$W_\alpha = \{U_b;\ |U_b| \geq u(\alpha)\}.$$

Protože je pro $n = m$ hodnota $|U_b| \leq |U_a|$ dává tato varianta častěji jako výsledek testu přijetí nulové hypotézy H_0.

15.11. Multinomické rozdělení. Uvažujme náhodné jevy A_i, $1 \leq i \leq k$, které jsou po dvou disjunktní, $P(A_i) = p_i,\ A_1 \cup A_2 \cup \ldots \cup A_k = U$, tedy $p_1 + p_2 + \ldots + p_k = 1$. Jestliže opakujeme $n-$krát pokus, který jako výsledek dává posloupnost jevů A_i nebo $\overline{A_i}$ a uvažujeme kolikrát se ma $i-$ tém místě objeví jev A_i, pak mluvíme o *multinomickém rozdělení* s parametry n a $p_1, p_2, \ldots, p_k$. Jestliže označíme jako náhodný vektor $(X_1, X_2, \ldots, X_k)$ výsledek pokusu pak pro sdruženou pravděpodobnostní funkci p dostaneme

$$p(i_1, i_2, \ldots, i_k) = P(X_1 = i_1, X_2 = i_2, \ldots, X_k = i_k) =$$

$$= \frac{n!}{i_1!.i_2!\dots i_k!} p_1^{i_1} p_2^{i_2} \dots p_k^{i_k},$$

kde $0 \leq i_j,\ 1 \leq j \leq k,\ i_1 + i_2 + \dots i_k = n$.

Marginální rozdělení každé z veličin X_j je binomické rozdělení $Bi(n, p_j)$ a $E(X_j) = np_j,\ D(X_j) = np_j(1 - p_j),\ 1 \leq j \leq k$. Dále je koeficient korelace $cov(X_i, X_j) = -np_ip_j,\ i \neq j,\ 1 \leq i, j \leq k$.

Takové rozdělení dostaneme, jestliže pro náhodný výběr provedeme diskretizaci jeho hodnot pomocí zvolené škály.

Nechť je $X_1, X_2, \dots X_n$ náhodný výběr z rozdělení s danou distribuční funkcí. Rozdělíme interval, ve kterém se může daná náhodná veličina vyskytovat na systém k disjunktních intervalů tvaru

$$(a_0, a_1\rangle,\ (a_1, a_2\rangle, \dots (a_{k-1}, a_k).$$

Dále označme $p_i = P(a_{i-1} < X \leq a_i), 1 \leq i \leq k$ pravděpodobnost výskytu náhodné veličiny X v $i-$tém intervalu škály. Potom je np_i *teoretická četnost* výskytu hodnot náhodného výběru v $i-$tém intervalu škály. Jestliže si označíme $n_i,\ 1 \leq i \leq k$ *empirickou četnost* výskytu, t.j. počet hodnot X_j z náhodného výběru, které leží v $i-$tém intervalu škály, pak platí tvrzení.

Věta: Náhodná veličina

$$(\spadesuit) \qquad \chi^2 = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i}$$

má přibližně rozdělení $\chi^2(k-1)$.

Poznámka: Hodnota χ^2 je vlastně vážený součet čtverců odchylek empirické a teoretické četnosti, kdy ja každá odchylka vážena proti své teoretické hodnotě. Tato hodnota má být co nejmenší.

Uvedeme vzorec, který se někdy lépe hodí k výpočtu hodnoty χ^2. Je totiž

$$\chi^2 = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i} = \sum_{i=1}^{k} \frac{n_i^2 - 2n_inp_i + (np_i)^2}{np_i} =$$

$$= \sum_{i=1}^{k} \frac{n_i^2}{np_i} - 2\sum_{i=1}^{k} n_i + \sum_{i=1}^{k} np_i = \sum_{i=1}^{k} \frac{n_i^2}{np_i} - n.$$

15.12. Test dobré shody, *test* χ^2 (chí kvadrát). Testujeme, že daný náhodný výběr je výběrem ze známého rozdělení. Pokud jsou parametry rozdělení (hustoty či pravděpodobnostní funkce) známy, počítáme uvedené veličiny z rozdělení, které je určeno jejich hodnotami. Pokud tyto parametry neznáme, použijeme pro ně odhady získané některou s metod hledání bodových odhadů (metoda maximální věrohodnosti či metoda momentů).

Máme dán náhodný výběr $X_1, X_2, \ldots, X_n$ z rozdělení se známým typem distribuční funkce (hustoty). Testujeme nulovou hypotézu

H_0 : náhodný výběr je výběrem s daným rozdělením

proti alternativě

H_1 : náhodný výběr je výběrem z jiného rozdělení.

Algoritmus testu.

1. Definiční obor náhodné veličiny X rozdělíme pomocí dělících bodů na škálu k intervalů tvaru

$(-\infty, a_1\rangle,\ (a_1, a_2\rangle, \ldots (a_{k-2}, a_{k-1}\rangle,\ (a_{k-1}, a_k = \infty)$.

2. Vypočteme teoretické četnosti

$$p_i = P(a_{i-1} < X \leq a_i),\ 1 \leq i \leq k$$

a ověříme podmínku použitelnosti testu:

$$np_i \geq 5,\ 1 \leq i \leq k, \qquad \text{nebo} \qquad np_i \geq 5q,$$

kde q je podíl tříd, pro které je $np_i < 5$, v případech kdy $k \geq 3$.

3. Určíme empirické četnosti n_i jako počty hodnot X_j z náhodného výběru, které leží v intervalu $(a_{i-1}, a_i\rangle$, $1 \leq i \leq k$ a vypočteme hodnotu statistiky

$$\chi^2 = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i}.$$

4. Pro zvolenou hladinu významnosti testu stanovíme kritický obor testu

$$W_\alpha = \{\chi^2;\ \chi^2 \leq \chi^2_{k-1}(\alpha)\},$$

kde $\chi^2_{k-1}(\alpha)$ je kritická hodnota testu, která je rovna $1-\alpha-$kvantilu rozdělení $\chi^2(k-1)$.

5. Je-li hodnota $\chi^2 \in W_\alpha$ zamítneme nulovou hypotézu H_0 ve prospěch alternativní hypotézy H_1. V opačném případě, kdy je $\chi^2 < \chi^2_{k-1}(\alpha)$ nulovou hypotézu H_0 přijmeme.

Poznámka: Pokud použijeme místo skutečných hodnot parametrů rozdělení jejich odhadů, pak místo $k-1$ stupňů volnosti rozdělení χ^2 volíme rozdělení s $k-m-1$ stupni volnosti, kde m je počet parametrů rozdělení.

5.13. Test závislosti a nezávislosti. Pro náhodné veličiny X a Y se nejčastěji k popisu závislosti používá *koeficient korelace* $\rho(X,Y)$, který je definován vztahem

$$\rho(X,Y) = \frac{E((X-E(X))(Y-E(Y)))}{\sqrt{D(X)D(Y)}} = \frac{E(XY)-E(X)E(Y)}{\sqrt{D(X)D(Y)}},$$

který je nulový pro nezávislé náhodné veličiny a je roven ± 1 v případě lineární závislosti $Y = aX + b$. Pro normální rozdělení je úplnou charakteristikou závislosti náhodných veličin. Platí totiž: Jestliže má náhodný vektor (X,Y) normální rozdělení, pak je jeho sdružená hustota dána vzorcem

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\,\mathrm{e}^{-\frac{\frac{(x-\mu_1)^2}{\sigma_1^2}+\frac{(y-\mu_2)^2}{\sigma_2^2}-\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}}{2(1-\rho^2)}},$$

kde náhodná veličina X má marginální rozdělení $N(\mu_1;\sigma_1^2)$ a Y má marginální rozdělení $N(\mu_1;\sigma_2^2)$ a ρ je koeficient korelace mezi X a Y. Náhodné veličiny X a Y jsou nezávislé právě když je $\rho = 0$.

Podmíněné náhodné veličiny $X|y$, resp.$Y|x$ mají také normální rozdělení se středními hodnotami

$$E(X|y) = \mu_1 + \beta_{1,2}(y-\mu_2),\ \beta_{12} = \rho\frac{\sigma_1}{\sigma_2},$$

$$E(Y|x) = \mu_2 + \beta_{21}(x-\mu_1),\ \beta_{21} = \rho\frac{\sigma_2}{\sigma_1}$$

a rozptyly

$$D(X|y) = \sigma_1^2(1-\rho^2), \quad \text{resp.} \quad D(Y|x) = \sigma_2^2(1-\rho^2).$$

Podmíněná střední hodnota je lineární funkcí y, resp. x, a její směrnice β_{12}, resp. β_{21}, je regresní koeficinet. Podmíněný rozptyl je konstantní.

Odhad závislosti či nezávislosti pro náhodné výběry provádíme pomocí výběrového koeficientu korelace, který je obdobou výběrových momentů.

Výběrový koeficient korelace je definován pro dvourozměrný náhodný výběr (X_i, Y_i),
$1 \leq i \leq n$ jako

$$r(X,Y) = \frac{S_{XY}}{S_X S_Y},$$

kde

$$S_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2, \quad S_Y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \overline{Y})^2,$$

$$S_{XY} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y}).$$

Vztah lze úpravami, kterými jsme odvodili vyjádření pro výběrový rozptyl upravit na tvar

$$r(X,Y) = \frac{\sum\limits_{i=1}^{n}(X_i Y_i) - n\overline{XY}}{\sqrt{\left(\sum\limits_{i=1}^{n} X_i^2 - n(\overline{X})^2\right)\left(\sum\limits_{i=1}^{n} Y_i^2 - n(\overline{Y})^2\right)}}$$

Test závislosti či nezávislosti je založen na tomto tvrzení: Je-li $(X_i, Y_i),\ 1 \leq i \leq n$ náhodný výběr z dvourozměrného normálního rozdělení, pak má náhodná veličina (statistika)

$$T = \frac{r}{\sqrt{1-r^2}}\sqrt{n-2} \sim t(n-2)$$

$t-$rozdělení s $n-2$ stupni volnosti.

Algoritmus testu

Testovaná hypotéza:

$H_0: \ \rho = 0$ nezávislost; $\quad H_1: \ \rho \neq 0$ závislost.

Kritický obor $W_\alpha = \{T;\ |T| > t_{n-2}(\alpha)\}$, kde $t_{n-2}(\alpha)$ je kritická hodnota $t-$testu, tedy $1-\frac{\alpha}{2}-$kvantil Studentova $t-$rozdělení pro $n-2$ stupňů volnosti.

Existují tabulky, které uvadějí kritické hodnoty $r_n(\alpha)$ přímo pro hodnoty statistiky r. Kritický obor je pak

$W_\alpha = \{r;\ |r| > r_n(\alpha)\}$.

15.14. Testy normality Uvedeme zde test normality rozdělení pro soubor dat, který je náhodným výběrem $\{X_i;\ 1 \le i \le n\}$. Použijeme testy, které jsou založeny na výběrové šikmosti a špičatosti, nebo na jejich kombinaci.

Připomeneme:

$M_k = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(X_i - \overline{X})^k,\ 1 \le k.$

$A_3 = \dfrac{M_3}{(M_2)^{3/2}}$ – výběrová šikmost;

$A_4 = \dfrac{M_4}{M_2^2}$, resp. $A_4 = \dfrac{M_4}{M_2^2} - 3$ – výběrová špičatost.

Je pak

$$E(A_3) = 0, \quad D(A_3) = \frac{6(n-2)}{(n+1)(n+3)}$$

a

$$E(A_4) = 3 - \frac{6}{n+1}, \quad \text{resp.} \quad E(A_4) = -\frac{6}{n+1},$$

$$D(A_4) = \frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}.$$

Algoritmus testu:

1. $H_0:\ A_3 = 0$, resp. $A_4 = 0$ proti $H_1:\ A_3 \ne 0$, resp. $A_4 \ne 0$.

2. Kritický obor testu:

$W_\alpha = \{A_{3,4}:\ |A_{3,4}| \ge a(\alpha)\}$,

kde pro menší rozsahy výběru jsou kritické hodnoty pro statistiky A_3 a A_4 uvedeny v tabulkách.

Pro větší rozsahy výběrů, $n > 200$ pro A_3 a $n > 500$ pro A_4, lze použít aproximace normálním rozdělením, které vychází z centrální limitní věty. Počítáme s tím, že náhodné veličiny

$$U_3 = \frac{A_3}{\sqrt{D(A_3}} \quad \text{a} \quad U_4 = \frac{A_4 - E(A_4)}{D(A_4)}$$

mají normované normální rozdělení. Kritické hodnoty testu nalezneme pomocí kvantilů normálního rozdělení. Kritickým oborem testu je

$$W_\alpha = \{U_3;\ |U_3| > u_{\alpha/2}\},$$

nebo

$$W_\alpha = \{U_4;\ |U_4| > u_{\alpha/2},$$

kde u_α je $\alpha-$kvantil nornálního rozdělení $N(0;1)$.

Existuje podstatné vylepšení postupu, které se dá použít v případě výběrů menšího rozsahu.

Test založený na šikmosti:

Postupně vypočteme

$$b = \frac{3(n^2+27n-70)(n+1)(n+3)}{(n-2)(n+5)(n+7)(n+9)},\ W^2 = \sqrt{2(b-1)} - 1,\ \delta = \frac{1}{\sqrt{\ln W}}$$

$$a = \sqrt{\frac{2}{W^2-1}}, \qquad Z_3 = \delta \ln \left[\frac{U_3}{a} + \sqrt{\left(\frac{U_3}{a}\right)^2 + 1} \right].$$

Potom má náhodná veličina Z_3 přibližně normální rozdělení $N(0;1)$ a hypotézu o normalitě rozdělení zamítáme v případě, že $|Z_3| \geq u_{\alpha/2}$. Test se dá použít pro $n > 8$.

Test založený na špičatosti:

Postupně vypočteme

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

$$Z_4 = \frac{1 - \frac{2}{9A} - \sqrt[3]{\frac{1-\frac{2}{A}}{1+U_4\sqrt{\frac{2}{A-4}}}}}{\sqrt{\frac{2}{9A}}}.$$

Náhodná veličina Z_4 má přibližně normální rozdělení $N(0;1)$ a hypotézu o normalitě zamítáme, pokud je $|Z_4| \geq u_{\alpha/2}$. Aproximace je použitelná pro $n \geq 20$.

Testy založené současně na šikmosti a špičatosti:

Pro výběry kde je rozsah $n > 200$ můžeme použít skutečnosti, že náhodná veličina

$$U_3^2 + U_4^2 \sim \chi_2^2$$

má rozdělení χ^2 o dvou stupních volnosti. Hypotézu o normalitě zamítáme, pokud je

$$U_3^2 + U_4^2 \geq \chi_2^2(\alpha).$$

Pro menší rozsahy, kde $n \geq 20$ lze použít skutečnosti, že má náhodná veličina

$$Z_3^2 + Z_4^2 \sim \chi_2^2$$

přibližně rozdělení χ_2^2 o dvou stupních volnosti. Hypotézu o normalitě zamítáme, pokud je

$$Z_3^2 + Z_4^2 \geq \chi_2^2(\alpha).$$