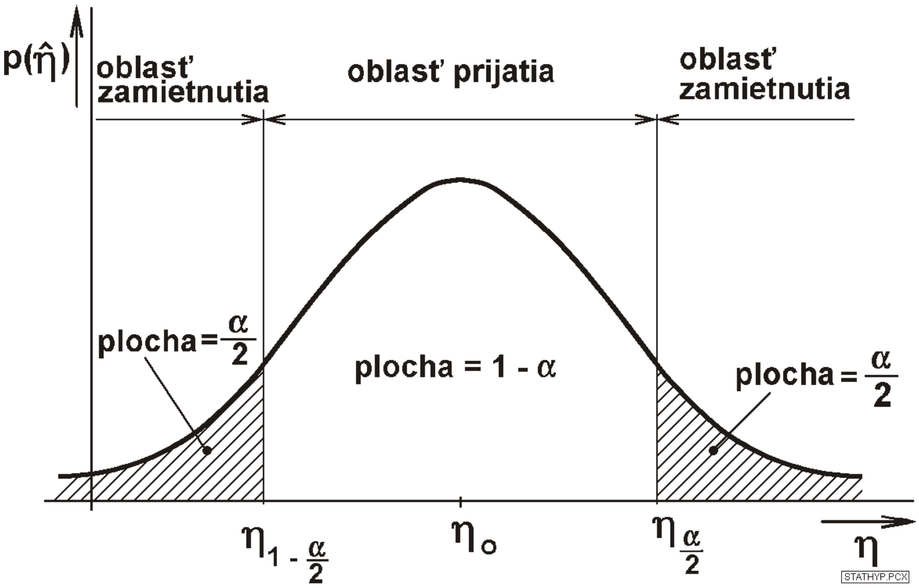
TESTOVANIE ŠTATISTICKÝCH HYPOTÉZ



OBLASŤ PRIJATIA A ZAMIETNUTIA HYPOTÉZY PRI TESTOVANÍ

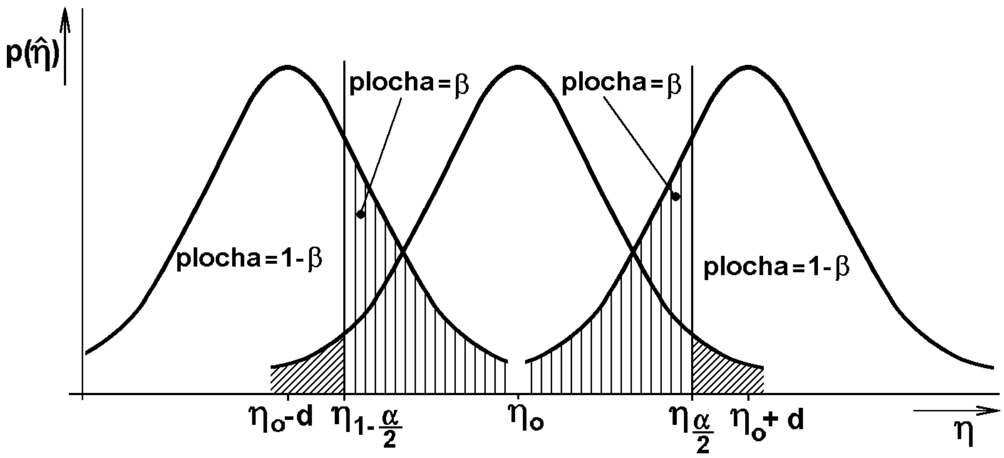
CHYBY I. A II. DRUHU

Chyba I. druhu sa vyskytne vtedy, ak je hypotéza správna, ale napriek tomu je zamietnutá, pretože hodnota príslušnej štatistiky $\hat{\eta}$ je málo pravdpodobná, teda pravdepodobnosť vzniku chyby I. druhu je rovná hladine významnosti α.

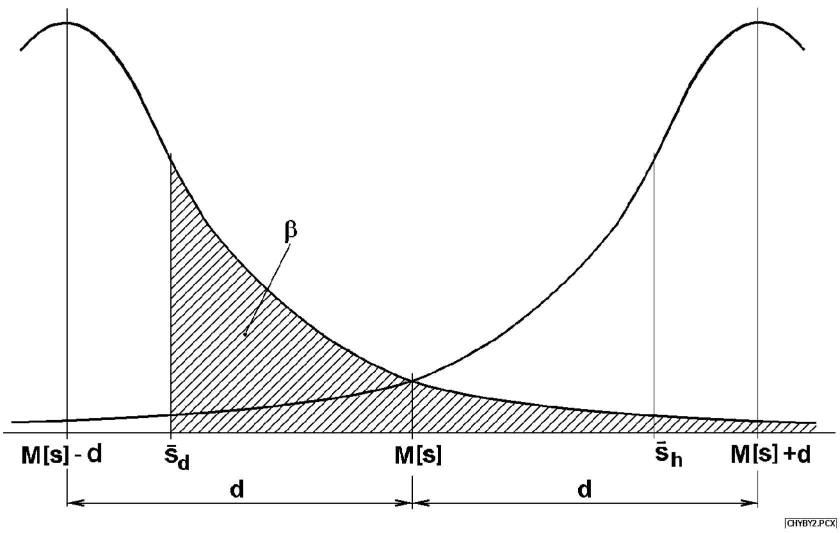
K druhému druhu chýb je možné urobiť nasledovnú úvahu. Predpokladajme, že hodnota istého parametra je buď $\eta = \eta_0 + d$, alebo $\eta = \eta_0 - d$. Ak hypotéza je postavená tak, že $\eta = \eta_0$ (avšak v skutočnosti platí $\eta = \eta_0 \pm d$), potom pravdepodobnosť toho, že $\hat{\eta}$ padne do oblasti prijatia hypotézy (ohraničenej medzami $\eta_{1-\alpha/2}$ a $\eta_{\alpha/2}$) je rovná β. Teda výskyt chýb II. druhu pre detekovateľné diferencie $\pm\, d$ od hypotetickej hodnoty η_0 je možné predvídať s pravdepodobnosťou β.

Pravdepodobnosť *(1-β)* je tzv. sila testu.

Pri nemennom rozsahu náhodného výberu vedie zmenšenie chýb I. druhu (zmenšením hladiny významnosti α) k zväčšeniu pravdepodobnosti vzniku chýb II. druhu (redukuje sa sila testu).



OBLASTI PRIJATIA A ZAMIETNUTIA HYPOTÉZY SPÔSOBUJÚCEJ CHYBY II. DRUHU PRI TESTOVANÍ HYPOTÉZY



URČENIE CHYBY II. DRUHU

Výber z tabuľky pre kritické hodnoty normálneho rozloženia

1 - "	0,8	0,9	0,95	0,99	0,9973
u"/2	1,282	1,643	1,96	2,576	3

Test významnosti rozdielu dvoch stredných hodnôt, ktorých odhady boli získané z dvoch nezávislých výberov

- testovaná (nulová) hypotéza H_0: $\bar{x}_1 = \bar{x}_2$,
- alternatívna hypotéza H_1: $\bar{x}_1 \neq \bar{x}_2$.

a) Ak sú disperzie oboch základných súborov známe

aa) ak sú obe disperzie rovnaké: $D[x_1] = D[x_2] = D[x] = \sigma^2$.

Testová charakteristika je: $u = \dfrac{\bar{x}_1 - \bar{x}_2}{\sigma_d}$; $kde\ \sigma_d = \sigma\sqrt{\dfrac{n_1 + n_2}{n_1 . n_2}}$;

kde n_1 je rozsah 1. náhodného výberu,

n_2 je rozsah 2. náhodného výberu,

ktorá má normované Gaussovo rozdelenie N(0, 1).

Ak platí: $|u| < u_{\alpha/2}$, , nie je dôvod zamietnuť testovanú hypotézu.

ab) ak disperzie nie sú rovnaké: $D[x_1] \neq D[x_2]$

Testová charakteristika je: $u = \dfrac{\bar{x}_1 - \bar{x}_2}{\sigma_d}$, $kde\ \sigma_d = \sqrt{\dfrac{D[x_1]}{n_1} + \dfrac{D[x_2]}{n_2}}$.

b) Ak disperzie oboch základných súborov nie sú známe

Na posúdenie toho, či rozdiely medzi odhadmi rozptylov oboch súborov sú štatisticky významné alebo nie, sa použije F-test.

ba) ak rozdiely medzi výberovými rozptylmi nie sú štatisticky významné, testová charakteristika má Studentovo rozdelenie v tvare:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{S_d};\quad kde\ S_d^2 = \frac{n_1 + n_2}{n_1 . n_2}\cdot\frac{S_1^2.(n_1 - 1) + S_2^2.(n_2 - 1)}{n_1 + n_2 - 2},$$

kde S_1 a S_2 sú výberové smerodadatné odchýlky 1. a 2. súboru.

Niekedy sa výraz S_d^2 používa v tvare: $S_d^2 = \dfrac{n_1 + n_2}{n_1 . n_2}\cdot\dfrac{n_1 . S_1^2 + n_2 . S_2^2}{n_1 + n_2 - 2}$.

Ak $|t| < t_{\alpha/2}(\nu)$, kde $\nu = n_1 + n_2 - 2$ je počet stupňov voľnosti, nie je dôvod na zamietnutie hypotézy o rovnosti oboch stredných hodnôt.

bb) ak rozdiely medzi výberovými rozptylmi sú štatisticky významné, testová charakteristika má Studentovo rozdelenie v tvare:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{S_1^2}{n_1 - 1} + \dfrac{S_2^2}{n_2 - 1}}}.$$

kde S_1^2 a S_2^2 sú výberové rozptyly 1. a 2. súboru s rozsahmi n_1 a n_2.

Testová charakteristika je: $$t_k = \frac{t_1 \dfrac{S_1^2}{n_1 - 1} + t_2 \dfrac{S_2^2}{n_2 - 1}}{\sqrt{\dfrac{S_1^2}{n_1 - 1} + \dfrac{S_2^2}{n_2 - 1}}},$$

kde t_1 a t_2 sú kritické hodnoty **t-rozdelenia** s počtom stupňov voľnosti $\nu_1 = n_1 - 1$ a $\nu_2 = n_2 - 1$ na hladine významnosti α.

Ak $|t| < t_k$, nie je dôvod zamietnutia testovanej hypotézy H_0.

TEST ROZDIELU DVOCH ROZPTYLOV (F-test)

Testovacie kritérium: $F = \dfrac{S_1^2}{S_2^2}; \quad S_1^2 > S_2^2,$ má F-rozdelenie s počtom stupňov voľnosti: $\nu_1 = n_1 - 1$ a $\nu_2 = n_2 - 1$ na hladine významnosti α.

Hypotézy:

- testovaná (nulová) hypotéza H_0: $S_1^2 = S_2^2,$
- alternatívna hypotéza H_1: $S_1^2 > S_2^2,$

Ide o **jednostranný test** a o výsledku testu rozhodne nerovnosť: $F < F_{\alpha, \nu_1, \nu_2}$.

BARTLETTOV TEST

slúži pre porovnanie súboru k výberových rozptylov, získaných z výberov s rozsahmi $n_1, n_2, \dots, n_k$. Tu je testovacia veličina B s približne χ^2 rozdelením s $(k - 1)$ stupňami voľnosti.

$$B = \frac{2{,}30259}{C}\left[K.\log S^2 - \sum_{j=1}^{k}(n_j - 1).\log S_j^2\right],$$

kde

$$K = \sum_{j=1}^{k}(n_j - 1), \qquad S^2 = \frac{1}{K}\sum_{j=2}^{k}(n_j - 1).S_j^2, \qquad C = 1 + \frac{\sum\limits_{j=1}^{k}\dfrac{1}{n_j - 1} - \dfrac{1}{K}}{3.(k - 1)}.$$

Nulová hypotéza sa nezamieta, ak platí: $\chi_{1-\alpha/2,\nu} < B < \chi_{\alpha/2,\nu}\ ; \quad \nu = k - 1.$

TESTY DOBREJ ZHODY

Na objektívne posúdenie toho, či empirická distribučná funkcia alebo hustota rozloženia pravdepodobnosti je dobre aproximovaná zvoleným typom rozloženia, alebo či boli dobre odhadnuté parametre daného rozloženia, slúžia testy známe pod názvom **Testy dobrej zhody**.

TEST χ^2

1. Pre predpokladaný typ rozdelenia sa vypočítajú z náhodného výberu odhady jeho parametrov.

2. Obor výskytu experimentálnych dát sa rozdelí do q tried tak, aby v každej triede bolo aspoň 5 očakávaných výskytov náhodnej veličiny.

3. Vypočítajú sa teoretické počty výskytov v jednotlivých triedach (za predpokladu platnosti zvoleného typu rozdelenie) $\hat{n}_j$, $j = 1, 2, ..., q$.

4. Zistia sa skutočné (experimentálne) početnosti výskytov v jednotlivých triedach n_j.

5. Vypočíta sa hodnota testovej charakteristiky v tvare: $\chi_v^2 = \sum_{j=1}^{q} \frac{(n_j - \bar{n}_j)^2}{\bar{n}_j}$, ktoré má približne rozdelenie $\chi^2(\nu)$, s počtom stupňov voľnosti $\nu = q - p - 1$, kde p je počet parametrov príslušného rozdelenia.

6. Kritický obor testu je $\chi_v^2 > \chi_\alpha^2(\nu)$, kde α sa volí obvykle $0,05$, v ktorom sa príjme alternatívna hypotéza, t. j. že zvolený typ rozloženia nevyhovuje experimentálnym údajom.

Počet tried v závislosti od rozsahu náhodného výberu

n	50	100	200	400	600	800	1000	1500	2000	4000	8000	10000
q	7	10	16	20	24	27	30	35	40	51	68	76

KOLMOGOROVOV - SMIRNOVOV TEST DOBREJ ZHODY

- pre jeden výber,
- pre dva výbery.

K-S TEST PRE JEDEN VÝBER

Testová charakteristika: $D_1 = max\left| F_n(x) - F(x) \right|.$

Kritické hodnoty K-S štatistiky pre $n > 40$ až 50: $D_{n;1-\alpha} \approx \left[-\frac{ln(\alpha/2)}{2.n} \right]^{1/2}$

Modifikovaná kritická hodnota testu: $D_{n;1-\alpha} = \frac{\delta_{1-\alpha}}{\sqrt{n}}$, $pre\ n > 50.$

Súčinitele $*_{1-}$" pre konštrukciu konfidenčných pásov distribučných funkcií K - S testu

Rozdelenie hustoty pravdepodobnosti	" = 0,01	" = 0,05	" = 0,1
Všeobecné so známymi parametrami (pôvodný K-S test)	1,628	1,358	1,224
Exponenciálne s 1 odhadovaným parametrom (merítka)	1,25	1,06	0,96
Normálne s oboma parametrami odhadovanými	1,03 *) 1,10 *)	0,89 *) 0,84 *)	0,81 *) 0,73 *)
Weibullovo s 2 odhadovanými parametrami (merítka a tvaru)	1,06	0,84	0,7

*) Podľa rôznych literárnych prameňov

V prípade *n < 50* je nutné použiť príslušné tabuľky kritických hodnôt alebo ich vypočítať z Kolmogorovho - Smirnovovho rozdelenia, ktorého distribučná funkcia má tvar:

$$F_{KS} = 0, \quad pre\ x \leq 0,$$

$$F_{KS} = 1 - 2\sum_{k=1}^{\infty} (-1)^{k-1}.exp(-2.k^2.x^2), \quad pre\ x > 0.$$

ak platí: $F_{KS}(\sqrt{n}.D_1) > 1 - \alpha,$ **zamieta sa testovaná hypotéza na hladine významnosti** *"*.

Kritické hodnoty K-S testu pre jeden výber

n	10	15	20	25	30	35	40	> 40
$D_{n;\ 0,95}$	0,409	0,338	0,294	0,264	0,242	0,224	0,210	$1.36/\sqrt{n}$

K-S TEST PRE DVA VÝBERY

$$\{x_i\},\ i = 1, ..., n;\ \{y_i\},\ i = 1, ..., m.$$

Testová charakteristika: $D_2 = max|\,F(x) - G(y)\,|,$

kde $F(x)$ **je experimentálna distribučná funkcia výberu** $\{x_i\}$,
$G(y)$ **je experimentálna distribučná funkcia výberu** $\{y_i\}$.
Testovacie kritérium v tvare $\sqrt{n.m/(n+m)}.D_2$ **má limitné K-S rozdelenie:**

$$\lim_{n,\,m \to \infty} P\left(\sqrt{\frac{n.m}{n+m}}\cdot D_2 < x\right) = F_{KS}(x).$$

Ak $F_{KS}\left(\sqrt{\frac{n.m}{n+m}}\cdot D_2\right) > 1 - \alpha,$ **zamietne sa hypotéza o zhode rozdelení oboch skúmaných náhodných výberov.**

Pre malé výbery (do 40 až 50) použiť len vtedy, ak sú oba výbery rovnakého rozsahu (t. j. $m = n$).

Pre väčšie výbery $n > 40$ až 50, $m > 40$ až 50 nemusí byť $n = m$. V tomto prípade sa môže určiť kritická hodnota testu na hladine významnosti ″ = 0,05 zo vzťahu

$$D_{2;\,0,05} = 1,36\sqrt{\frac{n+m}{n.m}}.$$

W-testy (Shapiro - Wilkov test)

a) Vypočíta sa výraz $S^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2.$

b) Súbor experimentálnych dát x_i; $i = 1, 2, 3, ..., n$ **sa usporiada do vzostupnej postupnosti** $x_1 \# x_2 \# \dots \# x_n$.

c) Vypočíta sa výraz $b = a_n(x_n - x_1) + a_{n-1}(x_{n-1} - x_2) + \dots,$ **kde hodnoty koeficientov** a_{n-i+1}; $i = 1, 2, ...$ **pre vybrané rozsahy náhodných výberov sú uvedené v tabuľke.**

d) Vypočíta sa testová charakteristika $W = \frac{b^2}{S^2}.$

e) Hodnota veličiny W pre N rozdelenie je rovná 1. Pri odchýlkach od N rozdelenia je $W < 1$, kritický obor testu je jednostranný, na dolnom okraji rozdelenia, teda ak je splnená nerovnosť $W < W_{1-\alpha}(n),$ **hypotéza o normálnosti rozloženia sa zamietne.**

Kritické hodnoty premenej $W_{1-″}(n)$ pre W - test normálnosti rozdelenia

″ / n	0,99	0,95	0,9	0,10	0,05	0,01
10	0,781	0,842	0,869	0,972	0,978	0,986
12	0,805	0,859	0,883	0,973	0,979	0,986
14	0,825	0,874	0,895	0,975	0,980	0,986
16	0,844	0,887	0,906	0,976	0,981	0,987
18	0,858	0,897	0,914	0,978	0,982	0,988
20	0,868	0,905	0,920	0,979	0,983	0,988
22	0,878	0,911	0,926	0,980	0,984	0,989
25	0,888	0,918	0,931	0,981	0,985	0,989
30	0,900	0,927	0,939	0,983	0,985	0,990
35	0,910	0,934	0,944	0,984	0,986	0,990
40	0,919	0,940	0,949	0,985	0,987	0,991

Koeficienty a_{n-i+1} pre W-test normálneho rozdelenia

n / i	10	12	14	16	18	20	22	25	30	35	40
1	0,5739	0,5475	0,5251	0,5056	0,4886	0,4734	0,459	0,445	0,4254	0,4096	0,3964
2	0,3291	0,3325	0,3318	0,329	0,3253	0,3211	0,3156	0,3069	0,2944	0,2834	0,2737
3	0,2141	0,2347	0,246	0,2521	0,2553	0,2565	0,2571	0,2543	0,2487	0,2427	0,2368
4	0,1224	0,1586	0,1802	0,1939	0,2027	0,2085	0,2131	0,2148	0,2148	0,2127	0,2098
5	0,0399	0,0922	0,124	0,1447	0,1587	0,1686	0,1764	0,1822	0,187	0,1883	0,1878
6		0,0303	0,0727	0,1005	0,1197	0,1334	0,1443	0,1539	0,163	0,1673	0,1691
7			0,024	0,0593	0,0837	0,1013	0,115	0,1283	0,1415	0,1487	0,1526
8				0,0196	0,0496	0,0711	0,0878	0,1046	0,1219	0,1317	0,1376
9					0,0163	0,0422	0,0618	0,0823	0,1036	0,116	0,1237
10						0,104	0,0368	0,061	0,0862	0,1013	0,1108
11							0,0122	0,0403	0,0697	0,0873	0,0986
12								0,02	0,0537	0,0739	0,087
13								0	0,0381	0,061	0,0759
14									0,0227	0,0484	0,0651
15									0,0076	0,0361	0,0546
16										0,0239	0,0444
17										0,0119	0,0343
18										0	0,0244
19											0,0146
20											0,0049

TESTY EXTRÉMNYCH ODCHÝLOK
GRUBBSOV TEST

Ako testovacie kritérium sa používajú vzťahy:

$$T_N = \frac{x_n - \bar{x}}{S}, \quad resp. \quad T_1 = \frac{\bar{x} - x_1}{S}.$$

Táto veličina má Studentovo t-rozdelenie. Ak $T_n > T_{n,\alpha}$, resp. $T_1 > T_{1,\alpha}$, nulová hypotéza sa zamieta a príslušné hodnoty x_1 alebo x_n, alebo obe sa vylučujú z ďalšieho spracovania. Pre kritické hodnoty testu platí: $T_{1,\alpha} = T_{n,\alpha}$.

Kritické hodnoty Grubbsovho testu extrémnych odchýlok

n	3	4	6	8	10	12	14	16	18	20	22	25
$T_{n,\,0,05}$	1,412	1,689	1,996	2,172	2,294	2,387	2,461	2,523	2,577	2,623	2,664	2,717
$T_{n,\,0,01}$	1,414	1,723	2,13	2,374	2,54	2,663	2,769	2,837	2,903	2,959	3,008	3,071

DIXONOV TEST

Testovacie kritériá sú: $Q_n = \dfrac{x_n - x_{n-1}}{x_n - x_1}$, *resp.* $Q_1 = \dfrac{x_2 - x_1}{x_n - x_1}$.

Kritické hodnoty Dixonovho testu extrémnych odchýlok

n	3	4	5	7	10	12	15	17	20	22	25	30
$Q_{n;\,0,05}$	0,941	0,765	0,642	0,507	0,412	0,376	0,338	0,32	0,3	0,29	0,277	0,26
$Q_{n;\,0,01}$	0,988	0,889	0,78	0,637	0,527	0,482	0,438	0,416	391	0,378	0,362	0,341