Testování

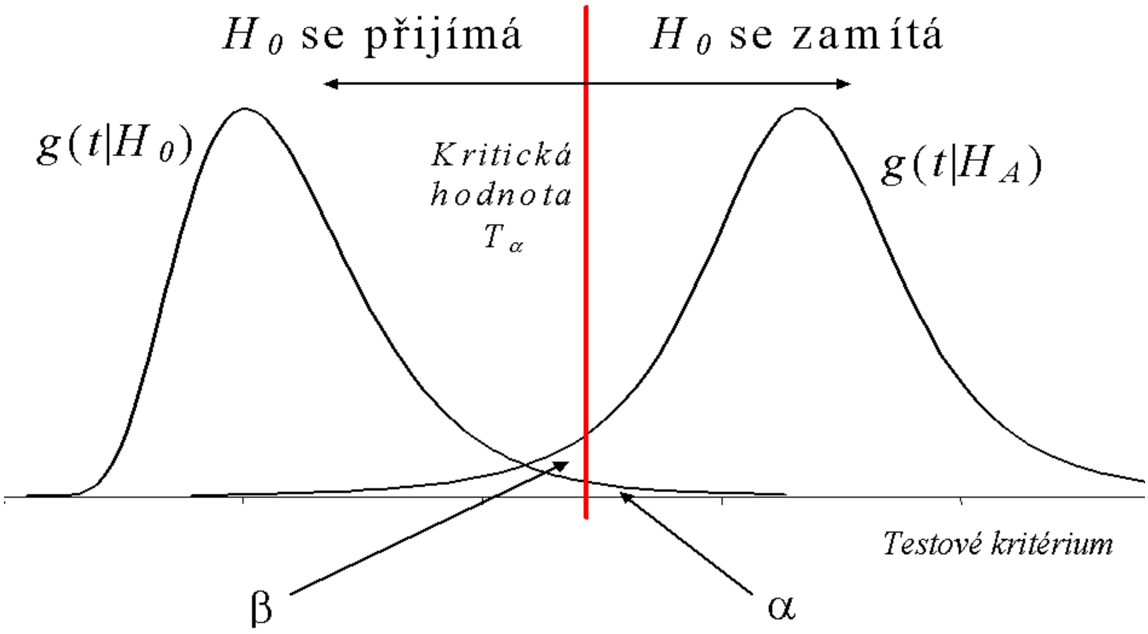
Menu:	QCExpert	Testování

Skupina *Testování* obsahuje následující moduly: *Síla a rozsah výběru*, *Testy* a *Kontingenční tabulka*.

Síla a rozsah výběru

Menu:	QCExpert	Testování	Síla a rozsah výběru

Moduly ve skupině *Síla a rozsah výběru* počítají sílu testu, rozsah výběru a mezní rozdíl parametru a jeho odhadu za daných předpokladů pro normální a binomické rozdělení. Vstupními parametry jsou vždy požadovaná hladina významnosti testu (riziko chyby I. druhu) α, typ testu (jednostranný, nebo oboustranný) a teoretická (očekávaná, požadovaná) hodnota parametru, kterým je u normálního rozdělení střední hodnota a u binomického pravděpodobnost. Dále je nutno definovat vždy dva z těchto tří údajů: Rozsah výběru, předpokládaná výběrová hodnota parametru a síla testu 1 - β. Třetí z těchto údajů se pak dopočítá stiskem příslušného tlačítka. Tento modul nepoužívá žádná data z datového editoru.



Obrázek 1 Pravděpodobnost chyby prvního (α) a druhého (β) druhu

Obrázek 1 ilustruje význam chyby I. a II. druhu. Symbolem H_0 označíme (jednoduchou) hypotézu, kterou chceme testovat, tzv. nulová hypotéza, například shodu průměru s danou hodnotou. Symbolem H_A pak označujeme neplatnost, tzv. alternativní hypotéza. Výsledkem testu je zamítnutí či nezamítnutí hypotézy H_0, obvykle na základě porovnání určité hodnoty vypočítané z dat, tzv. testačního kritéria T, s teoretickou kritickou hodnotou T_α pro toto kritérium. Tak u testování shody průměru x s danou hodnotou μ používáme Studentova *t*-testu, kdy testačním kritériem je $T = |x - \mu|/s$ a kritická hodnota $T_\alpha = t_{1-\alpha/2}(N-1)$ je 100(1–α)%ní kvantil Studentova rozdělení. Křivky na předchozím obrázku ilustrují hustoty testačního kritéria při platnosti H_0 ($g(t|H_0)$) a při neplatnosti H_0, tedy platnosti H_A ($g(t|H_A)$). Lze se dopustit dvou omylů:

Chyby I. druhu, kdy omylem zamítneme H_0, přestože ve skutečnosti H_0 platí. To znamená, že například testační kritérium T pro hypotézu H_0: ($x = \mu$) vyšlo příliš vysoké, neboť jsme k výpočtu průměru a směrodatné odchylky vzácnou (nešťastnou) náhodou vybrali velmi nepříznivá data, viz Obrázek 2. Riziko této chyby, neboli hladina významnosti testu se označuje α a můžeme si jej zadat, obvykle se volí α = 0.05, tedy 5%.

Obrázek 2 Vznik chyby I. druhu, H_0 je na základě 4 měření zamítnuta, ačkoliv platí

Chyba II.druhu. Podobně může dojít k chybě II. druhu, kdy hypotéza ve skutečnosti neplatí, ale měření použitá k výpočtu T vedou naopak k přijetí (nezamítnutí) H_0, viz Obrázek 3. Riziko chyby II. druhu je obyčejně v praxi méně důležité. Je tedy zřejmé, že chceme-li mít nízká rizika α a β, musíme mít velký počet měření, nebo musí být závěr testu velmi zřejmý (například značný rozdíl mezi měřenými daty a μ). Na druhé straně, máme-li málo dat a závěr testu není zřejmý a navíc požadujeme nízké riziko α, musíme se spokojit s vysokým rizikem β, tedy nízkou silou testu 1 – β. Všechny testy mají jednostrannou a oboustrannou variantu, kterou je možné zvolit. Při oboustranné variantě testujeme, zda hodnota (průměř dat, resp. podíl četností) vypočítaná z dat je různá od jiné hodnoty bez ohledu na to, zda „různá“ znamená větší nebo menší. Při jednostranné variantě testujeme, zda naměřená hodnota je větší, než vypočítaná, či nikoliv. Jednostranná varianta testuje vždy pouze $x > \mu$ u jednovýběrových, resp. $x_2 > x_1$ u dvouvýběrových testů pro normální rozdělení a $P_A > P_0$ u jednovýběrových, resp. $P_2 > P_1$ u dvouvýběrových testů pro binomické rozdělení.

K výpočtu těchto podmínek za zadaných předpokladů slouží modul *Síla a rozsah výběru*.

Obrázek 3 Vznik chyby II. druhu, H_0 je na základě 4 měření přijata, ačkoliv neplatí

Tento modul odpovídá tedy na tři druhy základních otázek:

- Kolik dat by bylo nejméně zapotřebí, aby se prokázal zadaný rozdíl mezi skutečnou a z dat odhadnutou střední hodnotou při zadané praděpodobnosti chyb I. a II. druhu α, β;
- Jaký je nejmenší rozdíl mezi skutečnou a z dat odhadnutou střední hodnotou, který by bylo možné prokázat na základě daného počtu dat a dané praděpodobnosti chyb I. a II. druhu α, β;
- Jaká by byla síla testu 1 – β, který má prokázat rozdíl mezi zadanou skutečnou a z dat odhadnutou střední hodnotou při daném počtu dat a dané chybě I. druhu α;

Normální rozdělení, 1 výběr

Menu:	QCExpert	Testování	Síla a rozsah výběru	Normální rozdělení, 1 výběr

Tento modul počítá parametry pro testování shody aritmetického průměru s danou konstantou.

Parametry

V dialogovém okně se musí zadat požadovaná hladina významnosti testu α, testovaná střední hodnota μ a předpokládaná směrodatná odchylka dat σ. Je možno zvolit typ testu výběrem varianty *Jednostranný* nebo *Oboustranný*. Dále je třeba vyplnit libovolné dvě hodnoty v polích *Rozsah výběru N*, *Průměr výběru X* a *Síla testu* 1 – β. U pole, které chceme dopočítat stiskneme příslušné tlačítko. Poslední vypočítaná hodnota se označí červeným rámečkem. Po výpočtu se okno nezavře ani se nic nezapisuje do protokolu a je možné provést další výpočet. Tlačítkem *Výstup do protokolu* se poslední výpočet zapíše do protokolu a okno zůstane otevřené. Tlačítkem *Zavřít* se okno zavře bez zápisu do protokolu.

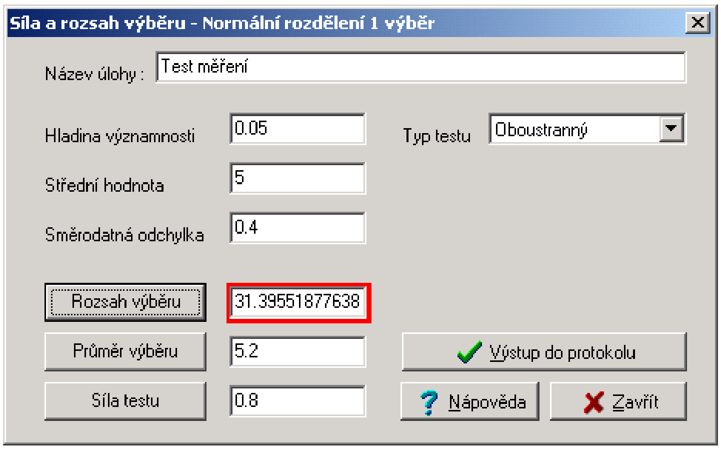
Síla testu se dá vypočítat ze zadných hodnot N, μ, X, σ, α podle vztahu

$$1-\beta=\Phi\left(\frac{\sqrt{N}\left(\mu-X\right)}{\sigma}-Z_{1-\alpha/2}\right)+\Phi\left(\frac{\sqrt{N}\left(X-\mu\right)}{\sigma}-Z_{1-\alpha/2}\right)$$

minimální počet potřebných dat je dán vztahem

$$N = \left\{ \left(\sigma \left(Z_{1-\alpha/2} + Z_{1-\beta} \right) \right) \Big/ \left| \mu - X \right| \right\}^2 ,$$

přičemž Z_α je α-kvantil normálního rozdělení a Φ distribuční funkce normálního rozdělení. Neznámá hodnota *X* (průměr výběru) se dá vypočítat iterativně.



Obrázek 4 Dialogový panel Síla a rozsah výběru, Normální rozdělení 1 výběr

Příklad

Předpokládaná, resp. požadovaná střední hodnota nějaké veličiny má být μ=5. Máme za úkol zjistit na základě budoucí série měření, zda se skutečná neliší od 5 o více, než dvě desetiny na libovolnou stranu. Otázkou je kolik (nákladných) měření k tomu budeme potřebovat. Riziko chyby I. druhu chceme mít α = 0.05 = 5%, což je obvyklá hodnota a riziko chyby II. druhu β chceme mít 20%. Ze zkušenosti předpokládáme směrodatnou odchylku měření kolem σ=0.4. Zadáme tedy *Hladinu významnosti* 0.05, *Střední hodnotu* 5, *Směrodatnou odchylku* 0.4. Protože nám vadí odchylka od 5 na obě strany, zadáme *Typ testu Oboustranný*. Do pole *Průměr výběru* zadáme mezní ***kladnou*** hodnotu 5.2 a do pole *Síla testu* zadáme 0.8 = 1 – β. Nakonec tlačítkem *Rozsah výběru* vypočítáme kolik dat budeme potřebovat k potvrzení, či zamítnutí hypotézy o odchylce od 5. Obrázek 4 ukazuje, že dat je třeba nejméně 32.

Protokol

Rozsah výběru a síla Normální rozd., 1 výběr	
Výpočet rozsahu výběru	Uvádí, který parametr byl počítán
Název úlohy	Název úlohy z dialogového okna
Hladina významnosti	Zadaná hladina významnosti testu
Střední hodnota M	Zadaná střední hodnota μ
Předpokládaný průměr X	Zadaný nebo vypočtený aritmetický průměr výběru.
Typ testu	Zadaný typ testu: jednostranný nebo oboustranný.
Nulová hypotéza H_0	X = M
Alternativní hypotéza H_A	X <> M v případě oboustranného testu, X > M v případě jednostranného testu. Vždy se počítá případ $x > \mu$.
Směrodatná odchylka	Zadaná předpokládaná směrodatná odchylka dat.
Rozsah výběru	Zadaný nebo vypočítaný rozsah výběru, neboli počet dat. Necelé číslo je třeba vždy zaokrouhlit nahoru.
Zaokrouhlený rozsah	Rozsah zaokrouhlený nahoru, aby byla zajištěna požadovaná rizika.
Síla testu	Vypočítaná nebo zadaná síla testu.

Normální rozdělení, 2 výběry

Menu:	QCExpert	Testování	Síla a rozsah výběru	Normální rozdělení, 2 výběry

Tento modul počítá parametry pro testování shody dvou aritmetických průměrů.

Parametry

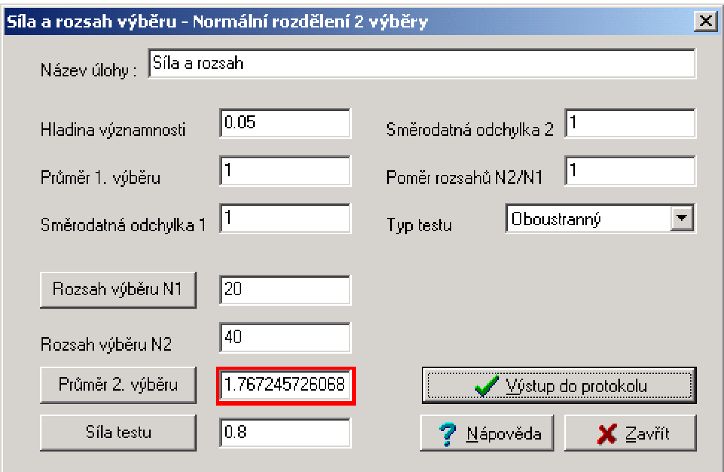
V dialogovém okně se musí zadat požadovaná hladina významnosti testu α, průměr prvního výběru X_1, průměr druhého výběru X_2 a předpokládané směrodatné odchylky obou výběrů. Mají-li se vypočítat rozsahy výběru, je možné zadat poměr rozsahu druhého a prvního výběru. Je možno zvolit typ testu výběrem varianty *Jednostranný* nebo *Oboustranný*. Dále je třeba vyplnit libovolné dvě hodnoty ve dvojici polí *Rozsah výběru N1* a *N2*, a v polích *Průměr výběru X* a *Síla testu* 1 – β. U pole, které chceme dopočítat stiskneme příslušné tlačítko. Hodnota *Poměr rozsahů N2/N1* se bere v úvahu pouze při výpočtu rozsahu výběru. Počítá-li se průměr 2. výběru, nebo síla testu, je možno do polí *Rozsah výběru N1* a *N2* zadat libovolné hodnoty. Poslední vypočítaná hodnota se označí červeným rámečkem. Po výpočtu se okno nezavře ani se nic nezapisuje do protokolu a je možné provést další výpočet. Tlačítkem *Výstup do protokolu* se poslední výpočet zapíše do protokolu a okno zůstane otevřené. Tlačítkem *Zavřít* se okno zavře bez zápisu do protokolu.

Minimální počty potřebných dat (rozsah výběrů N_1 a N_2) jsou dány vztahy

$$N_1 = \left(\sigma_1^2 + \frac{\sigma_2^2}{k}\right)\left\{\frac{\left(Z_{1-\alpha/2} + Z_{1-\beta}\right)}{\left|X_2 - X_1\right|}\right\}^2 ,$$

$$N_2 = kN_1$$

kde k je zadaný poměr N_2/N_1, Z_α je α-kvantil normálního rozdělení.



Obrázek 5 Dialogový panel Síla a rozsah výběru, Normální rozdělení 2 výběry

Poznámka

Ačkoliv je celkový rozsah $N_1 + N_2$ nejmenší při $N_1 = N_2$, tedy poměru $N_2/N_1 = 1$, může být výhodné zadáním poměru výběrů zmenšit jeden rozsah na úkor zvětšení druhého, například tehdy, když je získání dat jednoho výběru nákladnější, než u výběru druhého.

Protokol

Rozsah výběru a síla Normální rozd., 2 výběry	
Výpočet minimálního rozdílu	Uvádí, který parametr byl počítán.
Název úlohy :	Název úlohy z dialogového okna.

Hladina významnosti	Zadaná hladina významnosti testu.
Předpokládaný průměr X1	Zadaný průměr prvního výběru.
Předpokládaný průměr X2	Zadaný, nebo vypočtený průměr druhého výběru.
Typ testu	Zadaný typ testu: jednostranný nebo oboustranný.
Nulová hypotéza H0	X1 = X2
Alternativní hypotéza HA	X1 <> X2 v případě oboustranného testu, X1 < X2 v případě jednostranného testu. Vždy se počítá případ $X_1 < X_2$.
Směrodatná odchylka 1	Zadaná předpokládaná směrodatná odchylka 1. výběru.
Směrodatná odchylka 2	Zadaná předpokládaná směrodatná odchylka 2. výběru.
Poměr výběrů N2/N1	Zadaný požadovaný poměr rozsahu výběrů v případě výpočtu rozsahu.
Rozsah výběrů N1, N2	Zadaný nebo vypočítaný rozsah výběru, neboli počet dat 1. a 2. výběru. Necelé číslo je třeba vždy zaokrouhlit nahoru.
Zaokrouhlené rozsahy	Rozsahy zaokrouhlené nahoru, aby byla zajištěna požadovaná rizika.
Síla testu	Vypočítaná nebo zadaná síla testu.

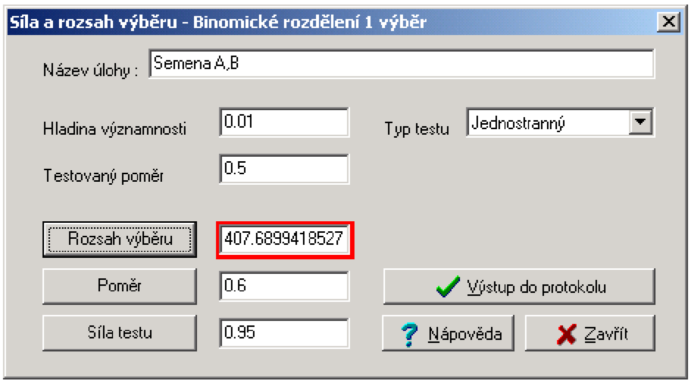
Binomické rozdělení, 1 výběr

Menu:	QCExpert	Testování	Síla a rozsah výběru	Binomické rozdělení, 1 výběr

Tento modul počítá parametry pro testování shody relativní četnosti jevu, tedy poměru výskytů k celkovému počtu případů (položka *Poměr*) se zadanou pravděpodobností tohoto jevu (položka *Testovaný poměr*).

Parametry

V dialogovém okně se musí zadat požadovaná hladina významnosti testu α, testovanou hodnotu poměru P_0 a předpokládaný poměr spočítaný z dat. Je možno zvolit typ testu výběrem varianty *Jednostranný* nebo *Oboustranný*. Dále je třeba vyplnit libovolné dvě hodnoty v polích *Rozsah výběru N*, *Poměr P* a *Síla testu* 1 – β. U pole, které chceme dopočítat stiskneme příslušné tlačítko. Poslední vypočítaná hodnota se označí červeným rámečkem. Po výpočtu se okno nezavře ani se nic nezapisuje do protokolu a je možné provést další výpočet. Tlačítkem *Výstup do protokolu* se poslední výpočet zapíše do protokolu a okno zůstane otevřené. Tlačítkem *Zavřít* se okno zavře bez zápisu do protokolu.



Obrázek 6 Dialogový panel Síla a rozsah výběru, Binomické rozdělení 1 výběr

Potřebný rozsah výběru *N* je dán vztahem

$$N = \left\{ \frac{\sqrt{P_0(1-P_0)}Z_{1-\alpha/2} + \sqrt{P_0(1-P_0)}Z_{1-\beta}}{|P-P_0|} \right\}^2 + \frac{2}{|P-P_0|}$$

kde Z_α je α-kvantil normálního rozdělení. Při výpočtu se využívá aproximace normálním rozdělením, která je dostatečně spolehlivá při $NP(1 - P) > 5$. Druhý člen v rovnici představuje korekci na náhradu diskrétního rozdělení spojitým.

Příklad

Podle našeho názoru by měla rostlinná populace produkovat 50% semen s vlastností A a 50% semen s vlastností B. Podle konkurenční teorie je tento poměr 40 : 60. Chceme pracovat na hladině $\alpha = 0.01$ a s pravděpodobností chyby II. druhu $\beta = 0.05$ a chceme znát počet náhodně vybraných semen, které bude třeba vyšetřit, abychom za těchto podmínek mohli potvrdit nebo vyvrátit naši domněnku. Označíme P podíl semen s vlastností B. Zadáme hladinu významnosti 0.01, testovaný poměr 0.5. Protože nepřipouštíme možnost poměru $P<0.5$, stačí testovat zda není $P>0.5$, zvolíme tedy *Typ testu Jednostranný*. Pak zadáme *Poměr P* 0.6 a požadovanou *Sílu testu* 1-β, 0.95. Tlačítkem *Rozsah výběru* získáme potřebný počet dat $N = 408$. Při méně přísných požadavcích na „neomylnost“ testu, např. při typických hodnotách $\alpha = 0.05$ a $\beta = 0.2$ by nutný rozsah souboru vyšel pouze 173.

Protokol

Rozsah výběru a síla Binomické rozd., 1 výběr	
Výpočet síly testu	Uvádí, který parametr byl počítán.
Název úlohy :	Název úlohy z dialogového okna.
Hladina významnosti	Zadaná hladina významnosti testu.
Testovaný poměr P_0	Zadaný testovaný poměr P_0.
Předpokládaný poměr P_A	Zadaný předpokládaný poměr P_0.
Typ testu	Zadaný typ testu: jednostranný nebo oboustranný.
Nulová hypotéza H_0	P0 = PA
Alternativní hypotéza H_A	P0 <> PA v případě oboustranného testu, P0 < PA v případě jednostranného testu. Vždy se počítá případ $P_0 < P_A$.
Rozsah výběru	Zadaný nebo vypočítaný rozsah výběru, neboli počet dat. Necelé číslo je třeba vždy zaokrouhlit nahoru.
Síla testu	Vypočítaná nebo zadaná síla testu.

Binomické rozdělení, 2 výběry

Menu:	QCExpert	Testování	Síla a rozsah výběru	Binomické rozdělení, 2 výběry

Tento modul počítá parametry pro testování shody dvou relativních četností jevu, tedy poměru výskytů X k celkovému počtu případů N prvního a druhého souboru (položka *Poměr X1/N1* a *Poměr X2/N2*).

Parametry

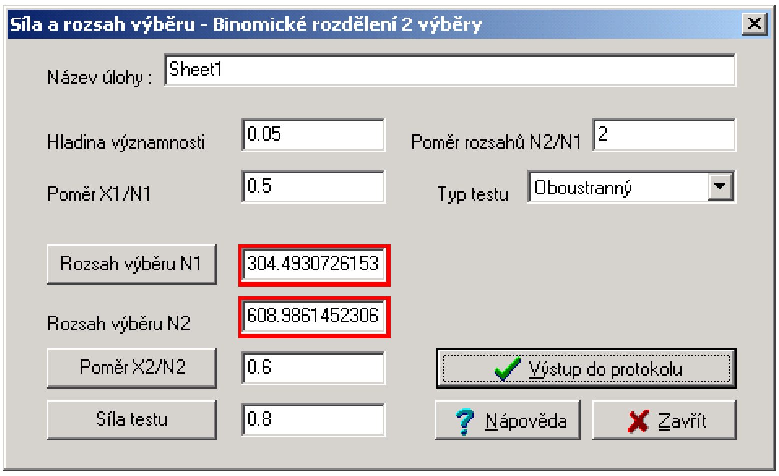
V dialogovém okně se musí zadat požadovaná hladina významnosti testu α, předpokládanou hodnotu poměru $P_1 = X_1/N_1$ prvního výběru a předpokládanou hodnotu poměru $P_2 = X_2/N_2$ druhého výběru. Je možno zvolit typ testu výběrem varianty *Jednostranný* nebo *Oboustranný*. Dále je třeba vyplnit libovolné dvě hodnoty ve dvojici polí *Rozsah výběru N*, a v polích *Poměr X2/N2* a *Síla testu* 1 - β. U pole, které chceme dopočítat pak stiskneme příslušné tlačítko. Poslední vypočítaná hodnota se označí červeným rámečkem. Po výpočtu se okno nezavře ani se nic nezapisuje do protokolu a je možné provést další výpočet. Tlačítkem *Výstup do protokolu* se poslední výpočet zapíše do protokolu a okno zůstane otevřené. Tlačítkem *Zavřít* se okno zavře bez zápisu do protokolu.

Rozsahy výběru N_1 a N_2 jsou dány vztahem

$$N_1 = \left\{ \frac{\sqrt{P_1(1-P_1)+\dfrac{P_2(1-P_2)}{k}}\,Z_{1-\beta} + \sqrt{\bar{P}(1-\bar{P})+1\dfrac{1}{k}}\,Z_{1-\alpha/2}}{|P_2-P_1|} \right\}^2 + \frac{k+1}{k|P_2-P_1|},$$

Korekce N_1 na spojitou aproximaci je obsažena ve druhém členu výrazu.

zde $\bar{P} = \dfrac{P_1 + kP_2}{1+k}$; k je zadaný poměr rozsahů výběru, $k = N_2/N_1$, takže $N_2 = k\,N_1$.



Obrázek 7 Dialogový panel Síla a rozsah výběru, Binomické rozdělení 2 výběry

Protokol

Rozsah výběru a síla Binomické rozd., 2 výběry	
Výpočet rozsahu výběrů	Uvádí, který parametr byl počítán.
Název úlohy :	Název úlohy z dialogového okna.
Hladina významnosti	Zadaná hladina významnosti testu.
Předpokládaný poměr P1	Zadaný poměr prvního výběru P_1.
Předpokládaný poměr P2	Zadaný poměr druhého výběru P_2.
Typ testu	Zadaný typ testu: jednostranný nebo oboustranný.
Nulová hypotéza H0	P1 = P2
Alternativní hypotéza HA	P1 <> P2 v případě oboustranného testu, P1 < P2 v případě jednostranného testu. Vždy se počítá případ $P_1 < P_2$.
Poměr rozsahů N2/N1	Zadaný požadovaný poměr rozsahu výběrů v případě výpočtu rozsahu.
Rozsah výběrů N1, N2	Zadaný nebo vypočítaný rozsah výběru, neboli počet dat 1. a 2. výběru. Necelé číslo je třeba vždy zaokrouhlit nahoru.
Zaokrouhlené rozsahy	Rozsahy zaokrouhlené nahoru, aby byla zajištěna požadovaná rizika.
Síla testu	Vypočítaná nebo zadaná síla testu.

Testy

Skupina *Testy* provádí statistické testování shody pro jednovýběrové a dvouvýběrové binomické a normální rozdělení, pro multinomické rozdělení a pro kontingenční tabulky. Při testování se zde vychází z experimentálních dat, případně ze známých statistik jako aritmetický průměr a směrodatná odchylka.

Binomický test, 1 výběr

Menu:	QCExpert	Testování	Testy	Binomický test, 1 výběr

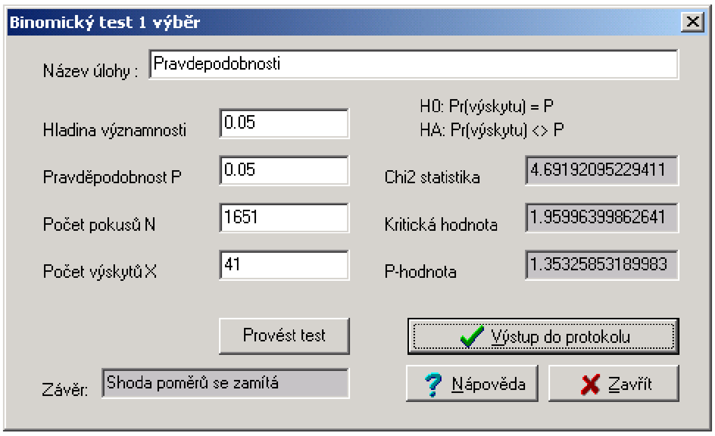
Tento modul testuje hypotézu H_0, zda pozorovaný počet X výskytů jevu A v celkem N případech odpovídá předpokládané pravděpodobnosti výskytu tohoto jevu P. Používá se standardního chi-kvadrát testu. Za předpokladu, že skutečná (nám principiálně neznámá) pravděpodobnost P_A výskytu jevu A je rovna předpokládané pravděpodobnosti P, byla by limitně pro $N \to \infty$ pravděpodobnost P rovna zjištěnému poměru X/N. Testuje se tedy, zda experiment vyvrací tento předpoklad, či nikoliv. Je dobré mít na paměti, že nezamítnutí shody neznamená ještě automaticky, že $P = P_A$. Ve skutečnosti to většinou znamená, že na zamítnutí H_0 prostě není dost dat. Prakticky lze však často považovat nezamítnutí za „potvrzení“ H_0, či alespoň za prokázání, že rozdíl mezi P_A a P je zanedbatelný.

Parametry

Tento modul nepoužívá žádná data z datového editoru. Do dialogového okna se zadá název úlohy, požadovaná hladina významnosti, předpokládaná *Pravděpodobnost P*, počet N uskutečněných pokusů v nichž mohl nastat jev A a *Počet výskytů X*, při nichž tento jev skutečně nastal. Po stisku tlačítka *Provést test* se zobrazí v poli *Závěr* výsledek testu ve slovní formulaci: Shoda poměrů se zamítá, resp. nezamítá. V pravé části dialogového okna se zobrazí rovněž vypočítaná statistika χ^2, kritická hodnota, nad kterou se H_0 zamítne a *P-hodnota*. Je-li p-hodnota menší, než zadaná hladina významnosti α, shoda poměrů se zamítne. Pro nízké počty výskytů a nízké pravděpodobnosti $XP < 5$ je tento test nespolehlivý. Tlačítkem *Výstup do protokolu* se výsledek testu zapíše do protokolu, stiskem tlačítka *Zavřít* se okno uzavře. Pro testování se používá χ^2 test založený na testovací statistice C, která má asymptoticky rozdělení $\chi^2_{(1)}$.

$$C = \frac{\left(X - NP\right)^2}{NP\left(1 - P\right)}$$

Tato statistika se porovná s kritickým kvantilem $\chi^2_{(1)}(1-\alpha)$. Je-li C větší, než kritický kvantil, hypotéza o shodě pravděpodobností se zamítne.



Obrázek 8 Dialogový panel pro Binomický test - 1 výběr

Příklad

Viz též příklad v odst. 0. Při náhodném dopadu na stůl zůstane prázdná krabička od sirek ležet na své největší stěně (A), střední stěně (B), nebo nejmenší stěně (C). Na základě experimentu chceme testovat hypotézu H_0, že pravděpodobnost, že krabička zůstane ležet na stěně C je 0.05, tedy 5%. Přitom ze 1651 zkušebních hodů zůstala krabička na stěně C ležet v 41 případech. Test chceme provádět na hladině významnosti 5%.

V poli *Hladina významnosti* zadáme 0.05, v poli *Pravděpodobnost* předpokládaný podíl výskytu 0.05 (shoda těchto dvou hodnot je čistě náhodná). V poli *Počet pokusů* zadáme 1651 a *Počet výskytů* 41. Po stisku tlačítka *Provést test* zjistíme, že experimentální výsledky vyvracejí na hladině významnosti $\alpha = 0.05$ náš předpoklad 5% pravděpodobnosti jevu „krabička stojí na stěně C".

Protokol

Binomický test shody, 1 výběr	Název modulu.
Název úlohy :	Název úlohy z dialogového okna.
Celkový rozsah	Počet pokusů N.
Počet výskytů	Počet X pozorovaných výskytů jevu A.
Výběrová pravděpodobnost X/N	Vypočítaný podíl X/N.
Testovaná pravděpodobnost	Předpokládaná pravděpodobnost P_A výskytu jevu A.
Hladina významnosti	Zadaná hladina významnosti α.
Statistika Z	Vypočítaná testová statistika s rozdělením χ^2.
Kritická hodnota U	Maximální přípustná hodnota χ^2 kvantilu pro danou hladinu α.
p-hodnota	Vypočítaná hodnota významnosti, na níž by byla H_0 právě zamítnuta.
Závěr	Slovní vyjádření závěru testu.

Binomický test, N výběrů

Menu:	QCExpert	Testování	Testy	Binomický test, N výběrů

Tento modul představuje rozšíření předchozího testu. Testuje simultánně K hypotéz pomocí výsledků z K skupin testů, zda pozorované počty X_i výskytů jevu A_i v N_i případech odpovídá předpokládaným pravděpodobnostem výskytu tohoto jevu P_i. Hypotéza H_0 je zde tedy H_0: $P_i = P_{Ai}$ pro $i = 1, .. K$, tedy všechny výskyty odpovídají předpokládaným pravděpodobnostem P_i, kde K je počet skupin. Používá se standardního chi-kvadrát testu. Za předpokladu, že všechny skutečné (nám principiálně neznámé) pravděpodobnosti P_{Ai} výskytu jevu A jsou rovny předpokládaným pravděpodobnostem P_i, byly by limitně pro $N_i \rightarrow \infty$ pravděpodobnosti P_i rovny zjištěným poměrům X_i/N_i. Testuje se tedy, zda experiment vyvrací tento předpoklad, či nikoliv. Je dobré mít na paměti, že nezamítnutí shody neznamená ještě automaticky, že $P = P_A$. Ve skutečnosti to často znamená, že na zamítnutí H_0 prostě není dost dat. Prakticky lze však často považovat nezamítnutí za „potvrzení" H_0, či alespoň za prokázání, že rozdíl mezi P_A a P je zanedbatelný.

Data a parametry

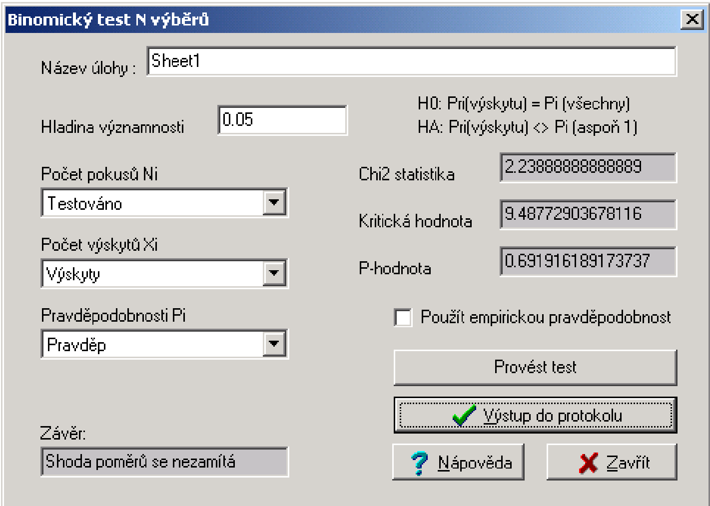
Data pro tento modul musí být ve sloupcích datového editoru. Jeden sloupec obsahuje počty provedených testů, druhý počty pozorovaných výskytů jevu A_i, třetí předpokládané pravděpodobnosti, které mají být testovány. Příklad dat je v následující tabulce, kde $K = 4$. Fyzické pořadí sloupců je libovolné.

Testováno	*Výskyty*	*Pravděp*
200	22	0.1
200	46	0.25
100	56	0.5
250	103	0.4

V poli *Hladina významnosti* se uvede požadovaná hladina významnosti α (typická hodnota je 0.05). V poli *Počet pokusů Ni* se vybere sloupec, v němž jsou uvedeny jednotlivé počty pokusů, v poli *Počet výskytů Xi* se vybere sloupec, v němž jsou příslušné počty výskytů a v poli *Pravděpodobnosti Pi* se vybere sloupec, v němž jsou očekávané, testované pravděpodobnosti výskytů. Je-li zaškrtnuto políčko *Použít empirickou pravděpodobnost*, ignorují se hodnoty ve sloupci *Pravděpodobnosti Pi* a jako P_i se použije jediná hodnota stejná pro všechna i, $P = \Sigma X_i / \Sigma N_i$ pro všechna i. Této volby se používá pro porovnání, zda několik binomických výběrů má stejné rozdělení, tedy stejnou pravděpodobnost výskytu P_i. Stiskem tlačítka *Provést test* se test vypočítá. Výsledky se zapíší do polí v dialogovém okně a toto okno zůstane otevřené. V poli *Chi2 statistika* se zobrazí hodnota vypočítané testovací statistiky s χ^2 rozdělením, v poli *Kritická hodnota* se zobrazí $(1 - \alpha/2)$ - kvantil χ^2 rozdělení, který se porovnává s testovací statistikou. Je-li kritická hodnota větší, H_0 o shodě pravděpodobností se zamítne na hladině α. V poli *P-hodnota* se zobrazí *p*-hodnota, tj. hodnota hladiny významnosti, na které by byla hypotéza H_0 právě zamítnuta. V poli *Závěr* se zobrazí výsledek testu ve slovní formulaci: Shoda poměrů se zamítá, resp. nezamítá. Tlačítkem *Výstup do protokolu* se výsledek testu zapíše do protokolu, stiskem tlačítka *Zavřít* se okno uzavře. Pro testování se používá χ^2 test založený na testovací statistice C, která má asymptoticky rozdělení $\chi^2_{(K-1)}$.

$$C = \sum_{i=1}^{K} \frac{1}{P_i\left(1-P_i\right)}\left(X_i - N_i P_i\right)$$

Tato statistika se porovná s kritickým kvantilem $\chi^2_{(K-1)}\,(1-\alpha)$. Je-li C větší, než kritický kvantil, hypotéza o shodě pravděpodobností se zamítne. Nejsou-li zadány teoretické pravděpodobnosti P_i (je-li zaškrtnuto políčko *Použít empirickou pravděpodobnost*), dosadí se za všechna P_i konstantní hodnota $\Sigma X_i / \Sigma N_i$.



Obrázek 9 Dialogový panel pro Binomický test - N výběrů

Příklad

Čtyři výrobní linky A, B, C, D produkovaly během jednoho měsíce 4200, 4800, 6100 a 2800 výrobků. Přitom byly zaznamenány při 100% automatické kontrole závady na těchto výrobcích v 165, 179, 201, resp. 109 případech. Úkolem je rozhodnout, zda střední frekvence závad (skutečná pravděpodobnost závady na vyrobené jednotce) je na všech linkách stejná. Data uspořádáme do datové tabulky.

Vyrobeno	Závady	Předpokl. zmetkovitost
4200	165	0.035
4800	179	0.035
6100	201	0.035
2800	109	0.035

Zadáme hladinu významnosti 0.05, V poli *Počet pokusů* vybereme sloupec *Vyrobeno*, v poli *Výskyty* vybereme sloupec *Závady*. Teoretická pravděpodobnost je ve sloupci *Předpokl. zmetkovitost*. Můžeme rovněž využít políčko *Použít empirickou pravděpodobnost*, tím se ignorují hodnoty ve třetím sloupci a místo nich se použije hodnota Σ(závady) / Σ(vyrobeno). Stiskem *Provést test* zjistíme, ze shoda poměrů se nezamítá, tedy pravděpodobnosti závady jsou na všech linkách stejné.

Protokol

Binomický test shody, N výběrů	Název modulu.
Název úlohy :	Název úlohy z dialogového okna.
Počet výběrů K	Počet skupin *K*.
Rozsahy výběrů Ni	Rozsahy jednotlivých skupin.
Počty výskytů Xi	Počty výskytů jevu v jednotlivých skupinách.
Teoretické výskyty Ni*Pi	Ideální hodnoty počtů výskytů při platnosti H_0.
Skutečné podíly Xi/Ni	Skutečně pozorované počty výskytů v jednotlivých skupinách.
Předpokládané podíly Pi	Zadané hodnoty předpokládaných pravděpodobností.
Hypotéza H0	PRi = Pi
Hypotéza HA	PRi <> Pi, počítá se pouze oboustranný test.
Hladina významnosti	Zadaná hladina významnosti, obyčejně se pracuje na hladině 0.05.
Stupně volnosti	Počet stupňů volnosti.
Statistika Chi2	Testační kritérium vypočítané z dat.
Kritická hodnota	Teoretická maximální přijatelná hodnota kritéria při platnosti H_0.
p-hodnota	Vypočítaná *p*-hodnota.
Závěr	Slovně vyjádřený závěr: Shoda poměrů se nezamítá, resp. zamítá.

Multinomický test

Menu:	QCExpert	Testování	Testy	Multinomický test

Tento modul je zobecněním binomického testu na případ, kdy může nastat několik vzájemně se vylučujících jevů A_i, $i = 1, ..., K$, K se nazývá počet tříd. Skutečné pravděpodobnosti těchto jevů označíme P_{Ai}, předpokládané pravděpodobnosti označíme P_i. Provede-li se celkem N pozorování, získá se K četností X_1, X_2, až X_K výskytů jevu A_1, A_2, až A_K. Přitom platí, že $\Sigma X_i = N$. Testuje se shoda pravděpodobností $P_{Ai} = P_i$ pro všechna $i = 1, .., K$ na základě empirických odhadů P_{Ai} poměrem X_i/N. Musí přitom (na rozdíl od předchozího binomického testu pro N výběrů) platit, že $\Sigma P_i = 1$ a $\Sigma P_{Ai} = 1$. Testovaná hypotéza H_0 je zde tedy H_0: $P_i = P_{Ai}$ pro $i = 1, .. K$, tedy všechny počty výskytů odpovídají předpokládaným pravděpodobnostem P_i, kde K je počet tříd. Používá se standardního chi-kvadrát testu. Za předpokladu, že všechny skutečné (nám principiálně neznámé) pravděpodobnosti P_{Ai} výskytu jevu A_i jsou rovny předpokládaným pravděpodobnostem P_i, byly by limitně pro $N \to \infty$ pravděpodobnosti P_i rovny zjištěným poměrům X_i/N. Testuje se tedy, zda experiment vyvrací tento předpoklad, či nikoliv. Je dobré mít na paměti, že nezamítnutí shody neznamená ještě automaticky, že $P = P_A$. Ve skutečnosti to většinou znamená, že na zamítnutí H_0 prostě není dost dat. Prakticky lze však často považovat nezamítnutí za „potvrzení“ H_0, či alespoň za prokázání, že rozdíl mezi P_A a P je zanedbatelný. Při malém počtu výskytu kterékoliv četnosti $X_i < 5$ je test neprůkazný. Pro testování se používá χ^2 test založený na testovací statistice C, která má asymptoticky rozdělení $\chi^2_{(K-1)}$.

$$C = \sum_{i=1}^{K} \frac{\left(X_i - NP_i\right)^2}{NP_i}$$

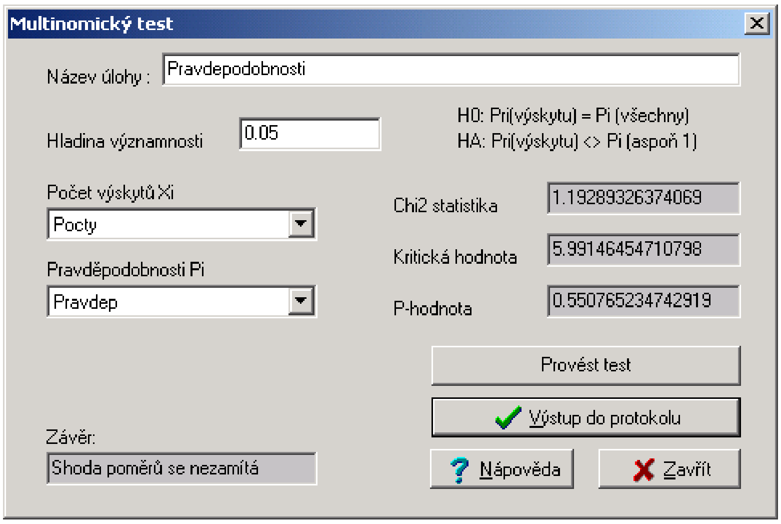
Tato statistika se porovná s kritickým kvantilem $\chi^2_{(K-1)}$ $(1-\alpha)$. Je-li C větší, než kritický kvantil, hypotéza o shodě pravděpodobností se zamítne.

Data a parametry

Data pro tento modul musí být ve sloupcích datového editoru. Jeden sloupec obsahuje počty pozorovaných výskytů jevu A_i, další sloupec předpokládané pravděpodobnosti, které mají být testovány. Příklad dat je v následující tabulce, kde $K = 4$. Fyzické pořadí sloupců je libovolné. Povšimněte si, že součet pravděpodobností musí být vždy roven jedné.

Výskyty	*Pravděp*
120	0.125
140	0.125
260	0.25
480	0.5

V poli *Hladina významnosti* se uvede požadovaná hladina významnosti α (typická hodnota je 0.05). V poli *Počet výskytů Xi* se vybere sloupec, v němž jsou příslušné počty výskytů a v poli *Pravděpodobnosti Pi* se vybere sloupec, v němž jsou očekávané, testované pravděpodobnosti výskytů. Stiskem tlačítka *Provést test* se test vypočítá. Výsledky se zapíší do polí v dialogovém okně a toto okno zůstane otevřené. V poli *Chi2 statistika* se zobrazí hodnota vypočítané testovací statistiky s χ^2 rozdělením, v poli *Kritická hodnota* se zobrazí $(1 - \alpha/2)$ - kvantil χ^2 rozdělení, který se porovnává s testovací statistikou. Je-li kritická hodnota větší, H_0 o shodě pravděpodobností se zamítne na hladině α. V poli *P-hodnota* se zobrazí *p*-hodnota, tj. hodnota hladiny významnosti, na které by byla hypotéza H_0 právě zamítnuta. V poli *Závěr* se zobrazí výsledek testu ve slovní formulaci: Shoda poměrů se zamítá, resp. nezamítá. Tlačítkem *Výstup do protokolu* se výsledek testu zapíše do protokolu, stiskem tlačítka *Zavřít* se okno uzavře.



Obrázek 10 Dialogový panel pro Multinomický test

Příklad

Viz též příklad v odst. 0. Při náhodném dopadu na stůl v homogenním gravitačním poli průměrné hospody zůstane prázdná krabička od sirek ležet na své největší stěně (A=*ab*), střední stěně (B=*ac*), nebo nejmenší stěně (C=*bc*). Na základě experimentu chceme testovat hypotézu H_0, že pravděpodobnost, že krabička zůstane ležet na určité stěně, je úměrná podílu plochy této stěny S_i (resp. prostorovým úhlem vymezeným touto stěnou s vrcholem v těžišti) a čtverce potenciální energie E_i dané polohy, $P_i \approx S_i/E_i^2$. Rozměry krabičky jsou $a = 47$mm, $b = 35$mm a $c = 15$mm. Po dosazení rozměrů krabičky získáme tedy poměr teoretických pravděpodobností $ab/c^2 : ac/b^2 : bc/a^2$, což odpovídá pravděpodobnostem $P_A = 0.89991$, $P_B = 0.07084$, $P_C = 0.02925$, neboť $P_A + P_B + P_C = 1$. Přitom ze 1651 zkušebních hodů zůstala krabička ve 1495 případech na stěně A, ve 115 případech na stěně B a ve 41 případě na straně C. Test chceme provést opět na hladině α = 0.05. Data budou tedy mít následující tvar:

Počty	**Pravděp**
1495	0.8999082056

115	0.0708382552
41	0.0292535391

Otevřeme dialogové okno *Multinomický test*. V poli *Hladina významnosti* zadáme 0.05. V poli *Počet výskytů Xi* vybereme sloupec *Počty*, v poli *Pravděpodobnosti* vybereme sloupec *Pravděp*. Po stisku tlačítka *Provést test* zjistíme, že experimentální výsledky souhlasí na hladině významnosti $\alpha = 0.05$ s našimi vypočtenými pravděpodobnostmi tří poloh krabičky, a nevyvracejí naši teorii (to ovšem ještě neznamená, že ji definitivně potvrzují).

Obrázek 11 Krabička od sirek

Protokol

Multinomický test shody	Název modulu.
Název úlohy :	Název úlohy z dialogového okna.
Počet tříd K	Počet tříd *K*.
Počty výskytů Ni	Pozorované počty výskytů jevu A_i.
Teoretické výskyty N*Pi	Ideální hodnoty počtů výskytů při platnosti H_0.
Skutečné podíly Ni/N	Skutečně pozorované počty výskytů v jednotlivých skupinách.
Předpokládané podíly Pi	Zadané hodnoty předpokládaných pravděpodobností.
Hypotéza H0	PRi = Pi
Hypotéza HA	PRi <> Pi
Hladina významnosti	Zadaná hladina významnosti, obyčejně se pracuje na hladině 0.05.
Stupně volnosti	Počet stupňů volnosti.
Statistika Chi2	Testační kritérium vypočítané z dat.
Kritická hodnota	Teoretická maximální přijatelná hodnota kritéria při platnosti H_0.
p-hodnota	Vypočítaná *p*-hodnota.
Závěr	Slovně vyjádřený závěr: Shoda poměrů se nezamítá, resp. zamítá.

Normální test, 1 výběr

Menu:	QCExpert	Testování	Testy	Normální test, 1 výběr

Tento test je určen pro testování shody měřených normálně rozdělených dat s danou střední hodnotou. Hypotéza H_0 zde zní: střední hodnota μ sledovaného procesu je shodná s danou hodnotou x_0. Test se provede na základě aritmetického průměru x_0 a směrodatné odchylky s, které se získají z naměřených dat. Často jsou v praxi známy již jen průměr a směrodatná odchylka, proto test nevyžduje původní data. Jsou-li původní data k dispozici, doporučuje se před provedením testu ověřit jejich normalitu, případně použít *t*-testu v modulu *Základní statistika*.

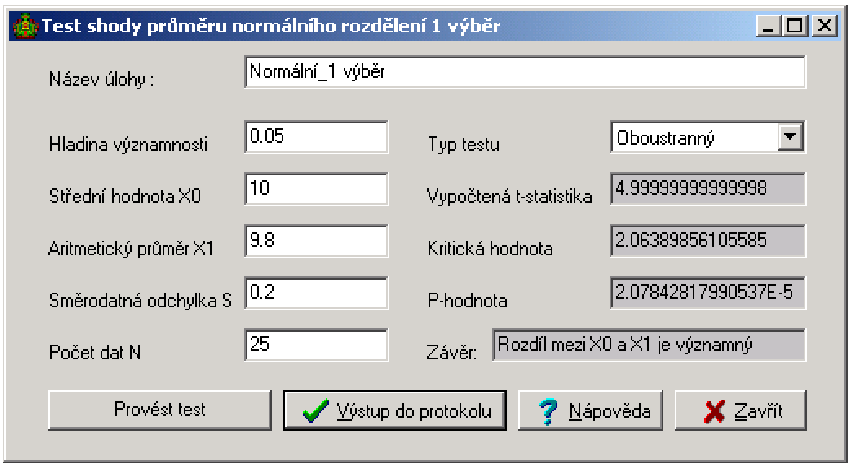
Pro testování se používá *t*-testu, při němž se porovnává *t*-statistika T_1 s kritickou hodnotou T.

$$T_1 = \frac{x_0 - \mu}{s}\sqrt{n};\quad T = t_{n-1}\left(1 - \alpha/2\right)$$

$t_n(\alpha)$ označuje α-kvantil Studentova rozdělení s n stupni volnosti. H_0 se zamítne, je-li $|T_1|>T$. Při jednostranné variantě se použije kvantil $T = t_{n-1}(1 - \alpha)$.

Parametry

V poli *Hladina významnosti* se zapíše požadovaná hladina významnosti testu α, v dalších polích se zapíše předpokládaná střední hodnota, skutečný aritmetický průměr dat a jejich směrodatná odchylka a počet dat, z nich byly průměr a směrodatná odchylka vypočítány. Dále zvolíme typ testu. Při jednostranném testu testujeme, zda je skutečná střední hodnota μ menší, resp. větší, než daná hodnota x_0, u oboustranného testu testujeme, zda je μ rozdílná od x_0 bez ohledu na to, zda je větší, nebo menší. Po stisknutí tlačítka *Provést test* se vypočítá příslušná *t*-statistika a kritická hodnota této statistiky, *p*-hodnota a v poli *Závěr* je uveden slovní závěr o významnosti rozdílu mezi μ a x_0. Okno po výpočtu zůstane otevřené. Zápis do protokolu lze provést tlačítkem *Výstup do protokolu*, tlačítko *Zavřít* zavře okno bez zápisu do protokolu.



Obrázek 12 Dialogový panel pro normální test pro jeden výběr

Protokol

t-test jednovýběrový	Název modulu.
Název úlohy :	Název úlohy z dialogového okna.
Střední hodnota X0	Zadaná testovaná střední hodnota .
Průměr dat X1	Zadaný aritmetický průměr dat.
Směrodatná odchylka dat S	Zadaná směrodatná odchylka dat.
Počet stupňů volnosti	$n - 1$
Vypočtená t-statistika	Hodnota t-statistiky T_1.
Kritická hodnota T	Kritický kvantil *t*-rozdělení pro danou hladinu α.
p-hodnota	Vypočtená p-hodnota.
Závěr	Slovní vyjádření závěru testu.

Normální test, 2 výběry

Menu:	QCExpert	Testování	Testy	Normální test, 2 výběry

Tento test je určen pro testování shody středních hodnot dvou procesů na základě dvou výběrů měřených normálně rozdělených dat se známým aritmetickým průměrem a směrodatnou odchylkou.

Hypotéza H_0 zde zní: střední hodnota μ_1 prvního sledovaného procesu je shodná s střední hodnota μ_2 druhého sledovaného procesu. Test se provede na základě aritmetických průměrů x_1, x_2 a směrodatných odchylek s_1 a s_2, které se vypočítají z naměřených dat. Počet dat prvního výběru je n_1, druhého výběru n_2. Často jsou v praxi známy již jen průměry a směrodatné odchylky, proto test nevyžaduje původní data. Jsou-li původní data k dispozici, doporučuje se použít modulu *Porovnání dvou výběrů*.

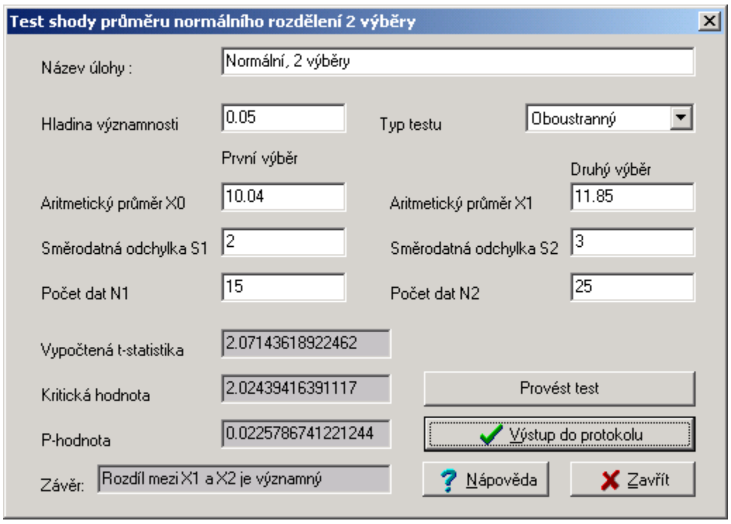
Pro testování se používá t-testu, při němž se porovnává t-statistika T_1 s kritickou hodnotou T.

$$T_1 = \frac{|x_2 - x_1|}{\sqrt{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}} \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}}; \quad T = t_{n_1+n_2-2}(1 - \alpha/2)$$

$t_n(\alpha)$ označuje α-kvantil Studentova rozdělení s n stupni volnosti. H_0 se zamítne, je-li $|T_1| > T$. Při jednostranné variantě se použije kvantil $T = t_{n1 + n2 - 2}(1 - \alpha)$.

Parametry

V poli *Hladina významnosti* se zapíše požadovaná hladina významnosti testu α, v dalších polích se zapíše skutečný aritmetický průměr prvního a druhého výběru *X1* a *X2* a jejich směrodatné odchylky *S1* a *S2* a počty dat *N1* a *N2*, z nichž byly průměry a směrodatné odchylky vypočítány. Dále zvolíme typ testu. Při jednostranném testu testujeme, zda je první střední hodnota μ_1 menší, resp. větší, než druhá μ_2, u oboustranného testu testujeme, zda je μ_1 rozdílná od μ_2 bez ohledu na to, zda je větší, nebo menší. Po stisknutí tlačítka *Provést test* se vypočítá příslušná t-statistika a kritická hodnota této statistiky, p-hodnota a v poli *Závěr* je uveden slovní závěr o významnosti rozdílu. Okno po výpočtu zůstane otevřené. Zápis do protokolu lze provést tlačítkem *Výstup do protokolu*, tlačítko *Zavřít* zavře okno bez zápisu do protokolu.



Obrázek 13 Dialogový panel pro normální test pro dva výběry

Protokol

t-test dvouvýběrový	Název modulu.
Název úlohy	Název úlohy z dialogového okna.
Průměr dat X1	Zadaný aritmetický průměr prvního výběru.
Směrodatná odchylka dat S1	Zadaná směrodatná odchylka prvního výběru.

Průměr dat X2	Zadaný aritmetický průměr druhého výběru.
Směrodatná odchylka dat S2	Zadaná směrodatná odchylka druhého výběru.
Počet stupňů volnosti	$n_1 + n_2 - 2$.
Vypočtená t-statistika	Hodnota t-statistiky T_1.
Kritická hodnota T	Kritický kvantil *t*-rozdělení pro danou hladinu α.
p-hodnota	Vypočtená *p*-hodnota.
Závěr	Slovní vyjádření závěru testu.