- Parametrické testy
- Neparametrické testy

7. Standardní statistické testy

Cíl kapitoly

V předchozí kapitole jsme si vysvětlili východiska a konkrétní postupy užívané při testování statistických hypotéz. Uvedli jsme zde pouze obecný postup, který je nutno dodržet při provádění statistického testu hypotézy. Následující kapitola Vám přináší konkrétnější pohled na danou problematiku v podobě některých významných statistických testů. Vzhledem k široké paletě používaných testů se zaměříme pouze na některé z nich. Bude se jednat především na tzv. parametrické testy, tedy testy, kdy předmětem hypotézy je některá z charakteristik (parametr) zkoumaného souboru – střední hodnota (průměr) nebo rozptyl.

Způsob studia

Metody statistického testování jsou v kapitole prezentovány na příkladu konkrétních testů statistických hypotéz. Jelikož se jedná o výpočtově náročnější operace, je vhodné využít možností, které nabízí výpočetní technika. Proto jsme kapitolu doplnili o přílohu, která naznačuje možnosti při testování hypotéz v prostředí programu MS Excel.

Časová zátěž

6 hodin (3. a 4. týden v prosinci)

Statistické testy tvoří poměrně samostatnou disciplínu teorie statistiky. Jak jsme již uvedli jejich konkrétní podoba je do jisté míry standardizována. Pro testy vybraných vlastností statistických souborů jsou používány dva základní druhy těchto standardních testů testy **parametrické** a **neparametrické**.

7.1 Parametrické testy

Parametrické statistické testy jsou užívány k rozhodování o pravdivosti, či nepravdivosti hypotézy týkající se konkrétní vlastnosti statistického souboru. Takovou typickou vlastností může být například hypotéza o hodnotě aritmetického průměru souboru, hypotéza o jeho rozptylu, či o jiných obdobných popisných statistických veličinách. Princip parametrických testů je založen na předpokladu, že rozdělení základního souboru, z něhož byl výběrový soubor pořízen, je určitého konkrétního typu.

V dalším textu se zaměříme pouze na omezený výsek těchto testů. Budeme se zabývat možnostmi používanými při testování hypotéz o základních dvou charakteristikách statistického souboru – jeho průměru a rozptylu. Ve všech dále používaných typech testů také vycházíme z jednotného předpokladu, že rozdělení, z něhož je výběrový soubor pořízen je normální. Je nutno zmínit, že tento předpoklad je poměrně silný a v velké spoustě praktických příkladů nesplnitelný.

Máme-li pochybnosti o normalitě rozdělení výběrového souboru, je možné tento předpoklad ověřit vhodným statistickým testem. Pro tyto potřeby jsou využívány tzv. neparametrické statistické testy hypotéz, o nichž se zmiňujeme v podkapitole 4.2.

V následující kapitole se seznámíte se čtyřmi základními typy parametrických testů vycházejících z předpokladu normálního rozdělení. Nejprve se seznámíme s dvěma testy o charakteristikách jednoho výběrového souboru – test hypotézy o střední hodnotě a rozptylu. Stěžejní pozornost pak budeme věnovat statistickým testům hypotéz o shodě či odlišnosti těchto charakteristik ve dvou souborech podobného typu (tzv. dvouvýběrové testy). V tomto případě mají hypotézy charakter předpokladu, že ve dvou obdobných statistických souborech se průměr, resp. rozptyl rovná stejné (konkrétní) hodnotě.

7.1.1 Test hypotéz o střední hodnotě μ (Studentovo t-rozdělení)

Pomocí tohoto typu statistického testu ověřujeme, zda na základě údajů z náhodného výběru platí předpoklad, že **průměr základního souboru se rovná určité hodnotě**.

Nulovou hypotézu testu tedy formulujeme ve tvaru $H_0 : \mu = \mu_0$.
Alternativní hypotéza je v případě dvoustranného testu $H_1 : \mu \neq \mu_0$ (v případě jednostranného má H_1 podobu $H_1 : \mu > \mu_0$ resp. $H_1 : \mu < \mu_0$).

Konkrétní **podoba testového kritéria** se odvíjí od rozsahu výběru. V případě, že vycházíme z výběrového souboru **velkého rozsahu** (což je v mnoha případech jednou z postačujících podmínek pro normalitu tohoto souboru) a známe rozptyl hodnot v základním souboru použijeme kritérium ve tvaru

$$U = \frac{\overline{x} - \mu_0}{\sigma}\sqrt{n},$$

kde n je počet prvků výběrového souboru, $\overline{x}$ je jeho aritmetický průměr a σje známá směrodatná odchylka základního souboru.

Neznáme-li hodnotu směrodatné odchylky základního souboru, musíme jej nahradit tzv. výběrovou směrodatnou odchylkou , který definujeme stejně jako jsme již uvedli v kapitole 5.3.1.

Při výběrových souborech **malého rozsahu** a za předpokladu alespoň přibližně normálního rozdělení, musíme použít testové kritérium ve tvaru:

$$t = \frac{\overline{x} - \mu_0}{s'_x}\sqrt{n},$$

která má rozdělení t (Studentovo) s $\nu = n - 1$ stupni volnosti.

Kritický obor je při velkém výběru a **jednostranné** variantě testu tvořen hodnotami menšími než kritická hodnota, kterou je kvantil u_α (resp. $u_{1-\alpha}$) normovaného normálního rozdělení. Při **dvoustranném** testu tvoří kritický obor hodnoty nacházející se mezi kritickými hodnotami $u_{\alpha/2}$ a $u_{1-\alpha/2}$. Pro test s předpokladem **malého výběru** nalezneme kritický obor analogicky. Pouze místo kvantilů normovaného normálního rozdělení použijeme kvantilů t-rozdělení s příslušným počtem stupňů volnosti.

Konkrétní příklady uvedeného typu statistického testu si nastudujte z učebnice Seger, Hindls, Hronová: *Statistika v hospodářství*. Příklad je uveden na stranách 165–166.

7.1.2 Test hypotéz o rozptylu σ (test χ^2)

Druhým významným parametrem, který je obvykle podrobován testování je rozptyl. Při tomto testování opět vycházíme z předpokladu, že z normálně rozděleného výběrového souboru byl proveden náhodný výběr o rozsahu n pozorování. Testovanou hypotézou je předpoklad, že rozptyl základního souboru je roven určité hodnotě σ_0^2. Vycházíme tedy z nulové hypotézy ve tvaru $H_0 : \sigma^2 = \sigma_0^2$. Alternativní hypotéza je v případě dvoustranného testu $H_1 : \sigma^2 \neq \sigma_0^2$ (v případě jednostranného má H_1 podobu $H_1 : \sigma^2 < \sigma_0^2$ resp. $H_1 : \sigma^2 > \sigma_0^2$).

Jako testové kritérium použijeme statistiku χ^2 (čti „chí-kvadrát“):

$$\chi^2 = \frac{(n-1)s_x'^2}{\sigma_0^2},$$

kde $s_x'^2$ je výběrový rozptyl definovaný již v kapitole 5.3.1. Tato statistika má rozdělení s $\nu = n - 1$ stupni volnosti.

Kritický obor je při **jednostranné** variantě testu tvořen hodnotami menšími než kritická hodnota, kterou je kvantil χ^2_α (resp. $\chi^2_{1-\alpha}$) rozdělení chí-kvadrát. Při **dvoustranném** testu tvoří kritický obor hodnoty nacházející se mezi kritickými hodnotami $\chi^2_{\alpha/2}$ a $\chi^2_{1-\alpha/2}$.

Konkrétní příklad uvedeného typu statistického testu si nastudujte z učebnice Seger, Hindls, Hronová: *Statistika v hospodářství*. Příklad je uveden na stranách 169–170.

7.1.3 Testy pro nezávislé výběry ze dvou normálních rozdělení

Tento typ testů patří k jedněm z nejčastěji využívaných, ať již v průmyslových, marketingových či jiných konjunkturálních průzkumech. Výhodou těchto testů je možnost porovnávat různé dvě či více různých situací, průzkumů, či výsledků. Na rozdíl od výše uvedených testů tak můžeme srovnávat úsudky o dvou základních souborech. Předpokladem těchto testů je nezávislost testovaných souborů, což je v praxi obvykle zajištěno, neboť v každém ze souborů se nacházejí jiné jednotky.

a) Parametrické testy o shodě či odlišnosti středních hodnot

Rozlišujeme základní tři skupiny testů týkajících se hypotéz o aritmetickém průměru (resp. střední hodnotě, jelikož pracujeme s náhodnými výběry a z nich pocházejícími náhodnými veličinami) statistického souboru. Testy o střední hodnotě se rozlišují především dle předpokladů, které můžeme učinit o rozptylu tohoto souboru. Na jejich základě dělíme statistické testy o střední hodnotě na

- **Dvouvýběrový *z*-test na střední hodnotu** – známe rozptyly v obou souborech
- **Dvouvýběrový *t*-test s rovností rozptylů** – neznáme rozptyly, ale předpokládáme, že jsou shodné
- **Dvouvýběrový *t*-test s nerovností rozptylů** – neznáme rozptyly a

považujeme je za rozdílné

- **Párový *t*-test na střední hodnotu** – speciální test, kdy máme k dispozici dva stejně velké výběry, přičemž odpovídající hodnoty jsou spárované

Podoba konkrétních testových kritérií u jednotlivých variant testů je nedílnou součástí standardních statistických učebnic. Jelikož je princip testování naprosto identický (liší se pouze v použitém testovém kritériu), omezíme se v dalším textu pouze na jeden příklad statistického testu o střední hodnotě – dvouvýběrového t-testu s předpokladem shodných rozptylů. Testová kritéria pro ostatní typy výše uvedených testů naleznete například v učebnici v učebnici Seger, Hindls, Hronová: *Statistika v hospodářství* na stranách 163–181.

Dvouvýběrový t-test s rovností rozptylů (Studentův t-test)

Studentův t-test je používán na test hypotézy, že průměr základního souboru μ se rovná určité hodnotě μ_0. Hodnota μ_0 je vypočítána z náhodného výběru malého rozsahu o kterém předpokládáme, že má normální rozdělení. Výchozím předpokladem t-testu tohoto typu je, že rozptyly v obou posuzovaných souborech jsou shodné, ačkoli neznáme jejich hodnotu.

Předpoklad shodnosti rozptylů je samostatnou statistickou hypotézou. Je zřejmé, že ji lze ověřit na základě adekvátního statistického testu. Jedná se o druhý typ testu, kterým se budeme zabývat dále v této kapitole.

Testovací kritérium pro hypotézu $H_0: \ \mu_1 = \mu_2$ (za předpokladu $\sigma_1^2 \sim \sigma_2^2$) má podobu

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

kde n_1 je počet prvků prvního výběrového souboru a n_2 je počet prvků druhého výběrového souboru a $\overline{x_1}$ a $\overline{x_2}$ jsou výběrové průměry získané z těchto souborů. Počet stupňů volnosti je dán

$$\nu = (n_1 - 1) + (n_2 - 1) = n_1 + n_2 - 2$$

(tedy jako počet prvků v příslušném statistickém souboru snížený o jedničku).

Kritickými hodnotami při konstrukci kritického oboru jsou kvantily t-rozdělení odpovídající hodnotám $t_{\alpha/2}$ a $t_{1-\alpha/2}$, kde α je příslušná hladina významnosti.

Příklad 7.1

V určitém podniku byla zjišťována náhodně u souboru 10 mužů a 5 žen výše jejich hrubého měsíčního příjmu. Výsledky jsou v následující tabulce (v tis. Kč).

M	13	11	19	15	22	20	14	17	14	15
Ž	9	12	8	10	16					

Rozhodněte na základě těchto výsledků zda lze v tomto podniku hovořit o diskriminaci žen nižší mzdou. Předpokládejte normálně rozložené platy.

Řešení

Jak jsme již uvedli, existuje těsná souvislost mezi statistickými testy a intervaly spolehlivosti. Proto uvedeme dva způsoby řešení následujícího příkladu.

A. Řešení s využitím *t*-testu s předpokladem stejných rozptylů

Pro uvedené zadání lze bez větších obav předpokládat, že rozptyly obou výběrových souborů – souboru žen a mužů ve stejném podniku – jsou stejné. Vychází z představy, že v daném podniku jsou platy mužů i žen rozvrstveny přibližně stejnoměrně od nejnižších po nejvyšší. Zajímá nás pouze hypotéza, zda ženy v průměru vydělávají méně než muži.

1) Formulace hypotézy

Hypotézu o platové diskriminaci žen můžeme vyvrátit na základě hypotézy formulované jako „neexistuje rozdíl mezi platy mužů a žen“. Pro zjednodušení ji tedy budeme formulovat následovně:

H_0: platy jsou stejné $\Rightarrow \Delta = 0$ $(\Delta = \mu_1 - \mu_2)$
H_1: $\Delta \neq 0$ – oboustranná varianta
$\Delta < 0$ – levostranná
$\Delta > 0$ – pravostranná

2) Volba testového kritéria

Použijeme t-statistiku (jedná se o malý výběr – do 100). Proto testové kritérium bude vycházet z t-rozdělení. Jak jsme uvedli použijeme t-testu s předpokladem stejných rozptylů. Proto použijeme kritéria:

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

3) Volba hladiny významnosti α

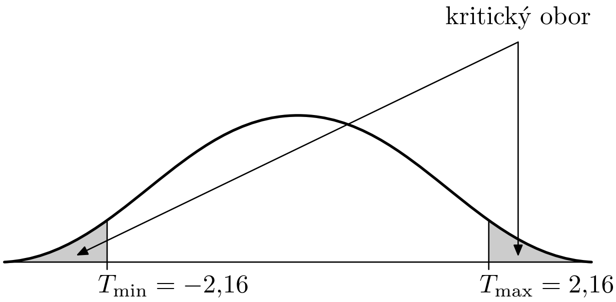
Budeme volit obvyklou $\alpha = 0{,}05$ tj. 5% hladinu významnosti. Výsledek testu ve formě zamítnutí nulové hypotézy tedy bude mít spolehlivost 95%. Jinými slovy pravděpodobnost chyby I. druhu (zamítneme hypotézu, která ve skutečnosti platí) bude pětiprocentní.

4) Sestrojení kritického oboru

a) Pro **oboustrannou variantu** testu bude kritický obor sestrojen na základě kvantilu $t_{\alpha/2}$ a $t_{1-\alpha/2}$ Studentova t-rozdělení. Pro určení hraničních hodnot využijeme vlastností t-rozdělení a postačí nám určit pouze jeden z kvantilů. Druhý nalezneme jako hodnotu s opačným znaménkem prvního z nich (neboť t-rozdělení je symetrické kolem nuly). Obvykle tedy hledáme pouze horní mez kritického oboru (horní kvantil). V našem případě jej nalezneme jako

$$T_{\max} = t_{0{,}975} = 2{,}16.$$

Příslušný počet stupňů volnosti určíme jako $(n_1 - 1) + (n_2 - 1) = (10 - 1) + (5 - 1) = 13$. Ve prospěch zamítnutí hypotézy H_0 budou tedy hovořit hodnoty testového kritéria **nižší než** $-2{,}16$ **a vyšší než** 2,16.

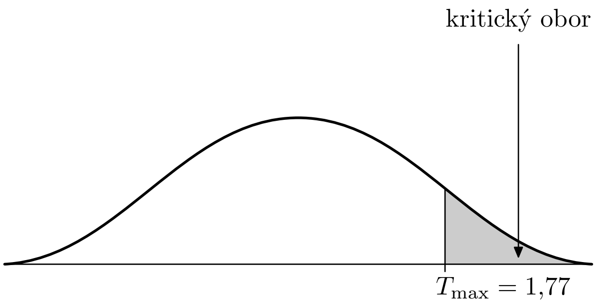


Obrázek 7.1: Kritický obor oboustranné varianty t-testu.

b) Pro **jednostrannou variantu** testu (pravostrannou) postačí výpočet jednoho kvantilu. V tomto případě hledáme $t_{1-\alpha}$ kvantil t-rozdělení. Pro 5% hladinu významnosti jej určíme (pro stejný počet stupňů volnosti jako u oboustranné varianty) jako

$$T_{\max} = t_{0{,}95} = 1,77.$$

Ve prospěch zamítnutí hypotézy H_0 budou tedy hovořit hodnoty testového kritéria vyšší než 1,77.



Obrázek 7.2: Kritický obor jednostranné varianty t-testu.

5) Výpočet hodnoty testového kritéria

T – odhad (rozdíl průměrných platů) / sm. odchylka stř. hodnot obou výběrů

$$T = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{n_1 s_1^2 + n_2 s_2^2}} \cdot \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}} =$$

$$= \frac{16 - 11}{\sqrt{10 \cdot 10{,}6 + 5 \cdot 8}} \cdot \sqrt{\frac{10 \cdot 5(10 + 5 - 2)}{10 + 5}} = \frac{5}{12{,}08} \cdot 6{,}58 = 2{,}72$$

6) Formulace výsledků testu

a) **oboustranná varianta**

Vypočítaná hodnota testového kritéria T (2,72) je vyšší než $T_{\max}$ (2,16)

$\Downarrow$

T padá do kritického oboru

$\Downarrow$

zamítáme hypotézu H_0 $(\Delta = 0)$

$\Downarrow$

na hladině spolehlivosti 5% přijímáme hypotézu H_1 $(\Delta \neq 0)$

$\Downarrow$

existuje rozdíl mezi platy mužů a žen

b) **jednostranná varianta (pravostranná)**

Vypočítaná hodnota testového kritéria T (2,72) je vyšší než $T_{\max}$ (1,77)

$\Downarrow$

T padá do kritického oboru

$\Downarrow$

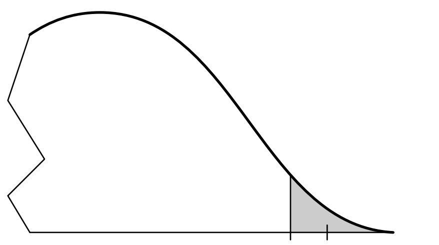
zamítáme hypotézu H_0 $(\Delta = 0)$

$\Downarrow$

na hladině spolehlivosti 5% přijímáme hypotézu H_1 $(\Delta > 0)$

$\Downarrow$

platy mužů jsou vyšší než platy žen



Obrázek 7.3: Výsledek testu.

B. Řešení pomocí konstrukce intervalu spolehlivosti

Při konstrukci intervalu spolehlivosti využijeme následujícího vztahu, který určuje oboustranný interval spolehlivosti pro rozdíl dvou průměrů při hladině spolehlivosti 95%.

$$\mu_1 - \mu_2 = \overline{x_1} - \overline{x_2} \pm t_{0,975} \cdot s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}},$$

kde

$$s_p = \frac{\sum_{i=1}^{n} (x_1 - \overline{x_1}) + \sum_{i=1}^{n} (x_2 - \overline{x_2})}{(n_1 - 1) + (n_2 - 1)}.$$

Po dosazení získáváme

$$\mu_1 - \mu_2 = (16 - 11) \pm 2{,}16 \cdot \sqrt{\frac{146}{13}} \cdot \sqrt{\frac{1}{10} + \frac{1}{5}} = 5 \pm 2{,}16 \cdot 1{,}84 = 5 \pm 4.$$

Lze tedy říci, že s 95% spolehlivostí leží rozdíl platů mezi muži a ženami v daném podniku mezi 1 a 9 tisíci Kč. Lze tedy hypotézu o rovnosti platů (rozdíl=0) zamítnout, neboť tato hodnota neleží uvnitř intervalu spolehlivosti.

b) Parametrické testy o shodě dvou rozptylů

U výše uvedeného prvního typu parametrických testů byl pro konkrétní podobu testové kritéria určující předpoklad o vlastnostech rozptylu ve zkoumaných souborech. Testy hypotéz o rovnosti, či odlišnosti rozptylů ve statistickém souboru tedy tvoří druhou nejvýznamnější skupinu parametrických statistických testů. Vlastnost rovnosti, resp. nerovnosti rozptylů se označují pojmy homo, resp. heteroskedasticita. Rozlišujeme tedy statistické soubory (prvky):

- **homoskedastická** (soubory se stejným rozptylem)
- **heteroskedastická** (soubory s různým rozptylem)

Pro testování hypotéz o rovnosti rozptylů se standardně používá Fisherova F-testu o shodě rozptylů. Vychází z nulové hypotézy ve tvaru $H_0: \ \sigma_1^2 = \sigma_2^2$ a alternativy $H_1: \ \sigma_1^2 \neq \sigma_2^2$.

Symbol σ_1^2 resp. σ_2^2 označuje předpokládaný rozptyl posuzovaných statistických souborů. Pro ověření uvedené hypotézy je nutno z obou souborů vypočítat výběrový rozptyl s'^2_x a na jeho základě spočítat testové kritérium. Testové kritérium má tvar podílu rozptylů příslušných výběrových souborů:

$$F = \frac{s'^2_1}{s'^2_2},$$

kde s'^2_1 a s'^2_2 jsou **výběrové rozptyly**, které lze vypočítat jako:

$$s'^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{n-1}.$$

Hodnoty hranic kritického oboru tvoří příslušné kvantily Fisherova rozdělení F.

Příklad 7.2

V rámci statistického zjišťování bylo prováděno šetření, zda existuje rozdíl mezi výdaji domácností za tzv. bílou techniku v závislosti na počtu dětí. Srovnání bylo provedeno pro dva vzorky domácností – domácnosti se dvěma a čtyřmi dětmi. Výsledky jsou uvedeny v následující tabulce (v tis. Kč za rok)

2 děti	41,2	40,6	39,4	41,5	36,3	37,4
2 děti	38,7	43,1	39,9	35,7	38,3	35,8
4 děti	39,2	43,8	38,9	44,3	41,2	44,1

Rozhodněte na základě těchto výsledků zda lze hovořit o rozdílech v objemu prostředků vynaložených na bílou techniku mezi uvedenými dvěma typy domácností.

Řešení

K ověření hypotézy o shodnosti, či odlišnosti průměrů ve dvou statistických souborech je možno využít ***t*-testu s předpokladem stejných rozptylů**. Jelikož u souboru domácností nelze vycházet z předpokladu stejného rozptylu hodnot, je nutno nejprve ověřit vlastnost shodnosti rozptylů v uvedených dvou souborech.

K tomuto ověření slouží **Dvouvýběrový Fischerův *F*-test o shodě dvou rozptylů**.

Pokud tento F-test prokáže shodnost rozptylů, je možno provést výše uvedený t-test. Úloha se tedy rozpadá do dvou statistických testů. V případě, že první test (na shodnost rozptylů) bude hovořit ve prospěch nezamítnutí hypotézy o jejich shodnosti, pak můžeme použít výše uvedený t-test. Pokud bude výsledek prvního testu hovořit ve prospěch zamítnutí hypotézy o stejných rozptylech, nelze tento t-test provést a je nutno použít jiného typu testu (např. Dvouvýběrový t-test s nerovností rozptylů).

1. Dvouvýběrový Fischerův *F*-test o shodě dvou rozptylů.

1) Formulace hypotézy

H_0: rozptyly jsou stejné $\Rightarrow H_0 : \ \sigma_1^2 = \sigma_2^2$
H_1: $H_1 : \ \sigma_1^2 \neq \sigma_2^2$ – oboustranná varianta
$H_1 : \ \sigma_1^2 < \sigma_2^2$ – levostranná
$H_1 : \ \sigma_1^2 > \sigma_2^2$ – pravostranná

2) Volba testového kritéria

Jako testové kritérium použijeme výše uvedenou F-statistiku ve tvaru:

$$F = \frac{{s'}_1^2}{{s'}_2^2}.$$

3) Volba hladiny významnosti α

Stejně jako u výše uvedeného t-testu zvolíme $\alpha = 0{,}05$ tj. 5% hladinu významnosti.

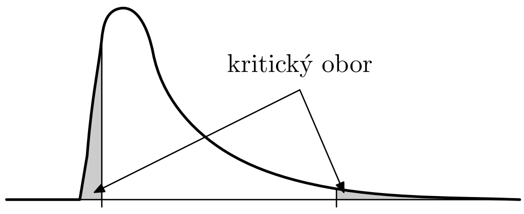
4) Sestrojení kritického oboru

Pro výpočet kritického oboru je nutno spočítat hodnoty kvantilů F-rozdělení příslušejících dané hladině významnosti. Kvantily lze nalézt v tabulkách nebo spočítat (např. pomocí Excelu – viz příloha této kapitoly).

Pro určení příslušných kvantilů F-rozdělení je nutno znát počet stupňů volnosti. Fischerovo rozdělení je charakterizováno dvěma stupni volnosti ν_1 a ν_2. Jejich hodnoty jsou dány $\nu_1 = n_1 - 1$, $\nu_2 = n_2 - 1$.

a) **pro oboustrannou variantu testu:**

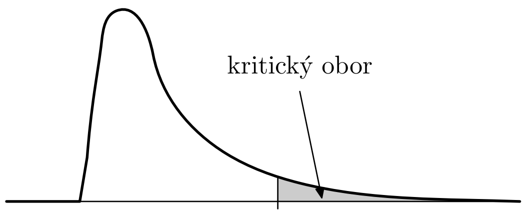
$$F_{\min} = F_{0{,}025} = 0{,}247, \qquad F_{\max} = F_{0{,}975} = 6{,}568$$



Obrázek 7.4: Kritický obor oboustranné varianty F-testu.

b) **pro jednostrannou (pravostrannou) variantu testu:**

$$F_{\max} = F_{0{,}95} = 4{,}70$$



Obrázek 7.5: Kritický obor jednostranné varianty F-testu.

Pozn. uvedené grafy příliš neodpovídají skutečnému tvaru F-rozdělení. Jsou uvedeny spíše jako ilustrační.

5) Výpočet hodnoty testového kritéria

$$F = \frac{s'^2_1}{s'^2_2} = \frac{5{,}726}{6{,}198} = 0{,}924.$$

6) Formulace výsledků testu

a) **oboustranná varianta**

Vypočítaná hodnota testového kritéria F (0,924) je vyšší než $F_{\min}$ (0,247), ale menší než $F_{\max}$ (6,58)

$\Downarrow$

F padá do oboru přijetí

$\Downarrow$

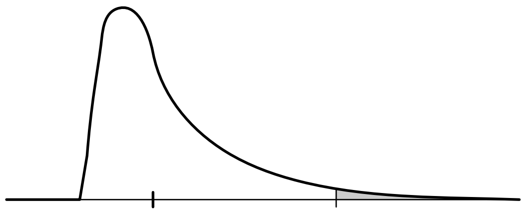
nezamítáme hypotézu H_0 ($\sigma_1^2 = \sigma_2^2$)

$\Downarrow$

na hladině významnosti 5% se neprokázal rozdíl mezi rozptyly v obou typech domácností

b) **jednostranná varianta (pravostranná)**

Vypočítaná hodnota testového kritéria F (0,94) je nižší než $F_{\max}$ (4,70)
$\Downarrow$
F padá do oboru přijetí
$\Downarrow$
nezamítáme hypotézu H_0 ($\Delta = 0$)
$\Downarrow$
na hladině významnosti 5% se neprokázal rozdíl mezi rozptyly v obou typech domácnost



Obrázek 7.6: Výsledek F-testu .

F-test tedy neprokázal rozdílnost rozptylů v obou základních souborech (typech domácností). Je proto možno přistoupit k ověření hypotézy o shodných průměrných výdajích za bílou techniku v rodinách se 2 resp. 4 dětmi.

2. Statistický *t*-test s předpokladem stejných rozptylů

1) Formulace hypotézy

H_0: výdaje jsou stejné $\Rightarrow \mu_1 = \mu_2$
H_1: $\mu_1 \neq \mu_2$ – oboustranná varianta
$\mu_1 > \mu_2$ – pravostranná

2) Volba testového kritéria

Použijeme t-statistiku (jedná se o malý výběr – do 100):

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

3) Volba hladiny významnosti α

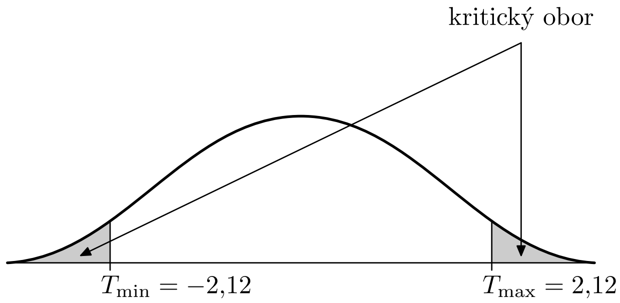
Standardně volíme $\alpha = 0{,}05$ tj. 5% hladinu významnosti.

4) Sestrojení kritického oboru

a) **pro oboustrannou variantu testu:**

$$T_{\min} = t_{0{,}025} = -2{,}12; \quad T_{\max} = t_{0{,}975} = 2{,}12.$$

Počet stupňů volnosti $\nu = (12 - 1) + (6 - 1) = 16$.

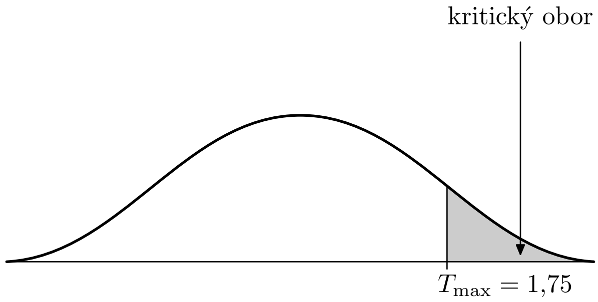


Obrázek 7.7: Kritický obor pro oboustrannou variantu t-testu.

b) **pro jednostrannou variantu testu:**

$$T_{\max} = t_{0,95} = 1,746.$$

Počet stupňů volnosti $\nu = (12 - 1) + (6 - 1) = 16$.



Obrázek 7.8: Kritický obor pro jednostrannou variantu t-testu.

5) Výpočet hodnoty testového kritéria

T – odhad (rozdíl průměrných platů) / sm. odchylka stř. hodnot obou výběrů

$$T = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{n_1 s_1^2 + n_2 s_2^2}} \cdot \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}} =$$
$$= \frac{38{,}99 - 41{,}92}{\sqrt{12 \cdot 5{,}25 + 5 \cdot 5{,}16}} \cdot \sqrt{\frac{12 \cdot 6(12 + 6 - 2)}{12 + 6}} = -\frac{2{,}925}{9{,}694} \cdot 8 = -2{,}414$$

Pozn.: s_1^2 a s_2^2 jsou rozptyly, které se v EXCELu spočítají pomocí funkce VAR.

6) Formulace výsledků testu

a) **oboustranná varianta**

Vypočítaná hodnota testového kritéria T $(-2{,}414)$ je nižší než $T_{\min}$ $(2{,}16)$

$\Downarrow$

T padá do kritického oboru

$\Downarrow$

zamítáme hypotézu H_0 ($\mu_1 = \mu_2$)
$\Downarrow$
na hladině spolehlivosti 5% přijímáme hypotézu H_1 ($\mu_1 \neq \mu_2$)
$\Downarrow$
**existuje rozdíl mezi výdaji domácností za „bílou techniku“
v závislosti na počtu dětí**

b) **jednostranná varianta (pravostranná)**

Vypočítaná hodnota testového kritéria T ($-2{,}414$) je vyšší než $T_{\max}$ (1,746)
$\Downarrow$
T padá do kritického oboru
$\Downarrow$
zamítáme hypotézu H_0 ($\mu_1 = \mu_2$)
$\Downarrow$
na hladině spolehlivosti 5% přijímáme hypotézu H_1 ($\mu_1 > \mu_2$)
$\Downarrow$
**existuje rozdíl mezi výdaji domácností za „bílou techniku“
v závislosti na počtu dětí**

7.2 Neparametrické testy

Druhou významnou skupinou statistických testů jsou neparametrické testy. Na rozdíl od parametrických testů v tomto případě vycházíme z předpokladu, že neznáme rozdělení základního souboru a přesto chceme porovnávat například průměry hodnot ve dvou statistických souborech. Stejně tak se neparametrických testů používá všude tam, kde předmětem statistické hypotézy je přímo tvar rozdělení prvků základního souboru, nebo například nezávislost sledovaných znaků.

liter Nejvýznamnějším typem neparametrických testů jsou tzv. **testy dobré shody**, testy nezávislosti v kombinační tabulce a testy shody úrovně. Popis některých standardních neparametrických testů, včetně příslušných testových kritérií, naleznete v učebnici Seger, Hindls, Hronová *Statistika v hospodářství* na stranách 181–221.

Testy dobré shody se používají, pokud jsou předpoklady normality dat evidentně nesplněné, například z důvodu, že:

- v souboru je příliš mnoho stejných hodnot, nebo
- některé hodnoty evidentně příliš odlehlé, nebo
- rozdělení četností je sice souměrné, ale má tvar písmene „U“.

Příkladem neparametrického testu dobré shody je χ^2 – test dobré shody („chí kvadrát“). Pomocí něj například testujeme zda se hodnoty jednoho výběrového souboru shodují s teoretickým modelem (např. výsledky testů pacienta před a po léčení – model je „že by se měly lišit“).

Shrnutí kapitoly

Konkrétní podoba statistických testů je do jisté míry standardizována. Pro testy vybraných vlastností statistických souborů jsou používány dva základní druhy těchto standardních testů testy parametrické a neparametrické. Parametrické statistické testy jsou užívány k rozhodování o pravdivosti, či nepravdivosti hypotézy týkající se konkrétní vlastnosti statistického souboru. Takovou typickou vlastností může být například hypotéza o hodnotě aritmetického průměru souboru, hypotéza o jeho rozptylu, či o jiných obdobných popisných statistických veličinách. Neparametrické testy (na rozdíl od parametrických) vycházejí z předpokladu, že rozdělení základního souboru není známo. Jsou užívány zejména v podobě tzv. testů shody, pomocí nichž testujeme zda se hodnoty jednoho výběrového souboru shodují s teoretickým modelem.

Nejznámějšími parametrickými testy jsou tzv. dvouvýběrové testy. Tyto testy se zaměřují na testování hypotéz o shodnosti či odlišnosti jednotlivých charakteristik ve dvou statistických souborech. V kapitole jsou popsány dva z těchto testů – dvouvýběrový t-test o odlišnosti středních hodnot ve dvou souborech a Fisherův F-test o shodě či odlišnosti rozptylů ve dvou souborech.

Otázky k zamyšlení

1. Modifikujte statistický test uvedený v příkladu 7.1 pro hladinu významnosti 1%.
2. Vysvětlete souvislost intervalu spolehlivosti a testování statistických hypotéz.
3. Ověřte výsledky t-testu uvedeného v příkladu 7.2 pomocí intervalu spolehlivosti.

POT

Součástí studia předmetu je i vypracování a odevzdání dvou krátkých samostatných prací, které jsou označovány jako POT. Obě samostatné práce mají formu příkladu, který by Vám měl dát možnost otestovat vědomosti nabyté v předchozí části studijní opory. Výsledky obou POTů odevzdáte ve stanovených termínech tutorovi v elektronické podobě (soubor v MS EXCEL + případný doprovodný text).

Termíny odevzdání jednotlivých úkolů jsou následující:

- POT 1 – 4. týden v říjnu
- POT 2 – 2. týden v prosinci

Odevzdání POTů a jejich správné řešení je podmínkou připuštění ke zkoušce z předmětu.

Zadání POT 2

Podnik používá dva stroje pro balení sušenek. Vedoucí střediska je přesvědčen, že variabilita vah sáčků balených na těchto dvou strojích není stejná.

Bylo tedy náhodně vybráno 14 balíčků balených na prvním stroji a 8 balených na druhém stroji. Lze empiricky potvrdit názor vedoucího střediska?

1. stroj	243,2	244,8	253,1	247,5	251	251,7	254	252,8
1. stroj	252,5	250,1	247,3	250,9	253,2	252,7		
2. stroj	250,2	250,1	251,3	249,1	249,9	250,8	251,9	252,2

Za předpokladu, že váha balíčků má normální rozdělení:

a) prokažte shodu středních hodnot vah balíčků na obou strojích,
b) otestujte shodnost či odlišnost rozptylů.

Příloha kapitoly 7

Využití programu MS EXCEL pro statistické testování hypotéz

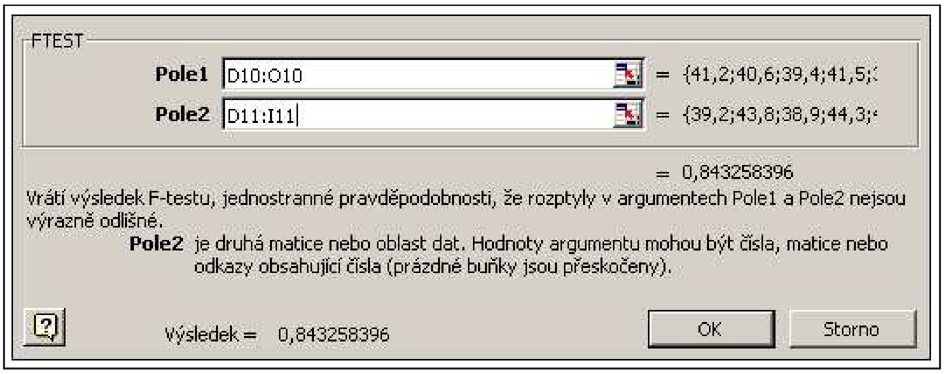
Stejně jako v předchozích případech i v oblasti testování statistických hypotéz je možno využít prostředí programu MS EXCEL k některým pomocným výpočtům. Především je možno pomocí tohoto programu snadno vypočítat kvantily náhodných rozdělení (postup jsme naznačili v příloze kapitoly 4). V případě funkce TINV je však upozornit, že jejím výsledkem je kritická hodnota Studentova rozdělení pro oboustranné testy. Pokud se testuje jednostranná hypotéza, je třeba za „alfa“ dosadit do funkce TINV dvojnásobek hladiny významnosti.

Výpočet p-hodnoty statistického testu

Další oblastí je výpočet výsledků statistických testů v podobě tzv. p-hodnot. p-hodnota udává statistickou významnost vypočítaného testového kritéria. Je tedy alternativním zobrazením hodnoty testového kritéria. Program EXCEL v sobě má zabudovány dvě funkce, které zobrazují hodnoty p-hodnot. Jedná se o t-test (funkce TTEST) a F-test (funkce FTEST).

Funkce TTEST je však možno použít pouze pro párovou variantu Studentova testu. Tedy pro variantu, kdy posuzujeme průměry stejně velkých statistických souborů.

Funkce FTEST je obecnější a je ji možno použít k zobrazení výsledku F-testu dvou libovolných statistických souborů. Výstupem funkce je p-hodnota jednostranné varianty testu. Je to tedy pravděpodobnost, že rozptyly v daném souboru nejsou odlišné. Dostáváme-li hodnoty vyšší než 0,05, resp. 0,01, můžeme zamítnout hypotézu o shodnosti rozptylů na hladině významnosti 5%, resp. 1%.



Obrázek 7.9: Zadání funkce FTEST pro hodnoty z příkladu 7.2.

Využití analytických nástrojů

Nejpřehlednější možností jak v EXCELu provést statistický test nabízí analytické nástroje (naleznete je v nabídce Nástroje/Analýza dat). Program ve verzi 2000 disponuje následujícími vestavěnými standardními testy:

- Analytický nástroj Dvouvýběrový z-test na střední hodnotu.
- Analytický nástroj Dvouvýběrový t-test s rovností rozptylů.

- Analytický nástroj Dvouvýběrový t-test s nerovností rozptylů.
- Analytický nástroj Dvouvýběrový párový t-test na střední hodnotu.
- Analytický nástroj Dvouvýběrový F-test pro rozptyl.

Jejich použití je velmi obdobné jako u analytických nástrojů zmíněných v přílohách předchozích kapitol. Opět je výstupem komplexní přehled výsledků jednotlivých testů v podobě samostatné tabulky. Pro příklady 7.1 a 7.2 uvedené v této kapitole lze využít nástroje „Dvouvýběrový t-test s rovností rozptylů“ pro příklad 7.1 a druhou část příkladu 7.2 a analytického nástroje „Dvouvýběrový F-test pro rozptyl“ pro první část příkladu 7.2.

Tabulky s výstupy pro tyto příklady mají následující podobu:

Dvouvýběrový t-test s rovností rozptylů		
	M	Ž
Stř. hodnota	16	11
Rozptyl	11,77777778	10
Pozorování	10	5
Společný rozptyl	11,23076923	
Hyp. rozdíl stř. hodnot	0	
Rozdíl	13	
t stat	2,723984457	
P(T<=t) (1)	0,008690118	
t krit (1)	1,770931704	
P(T<=t) (2)	0,017380235	
t krit (2)	2,16036824	

Tabulka 7.1: Dvouvýběrový t-test s rovností rozptylů pro test odlišnosti platů mužů a žen (příklad 7.1).

Označení některých významných buněk:

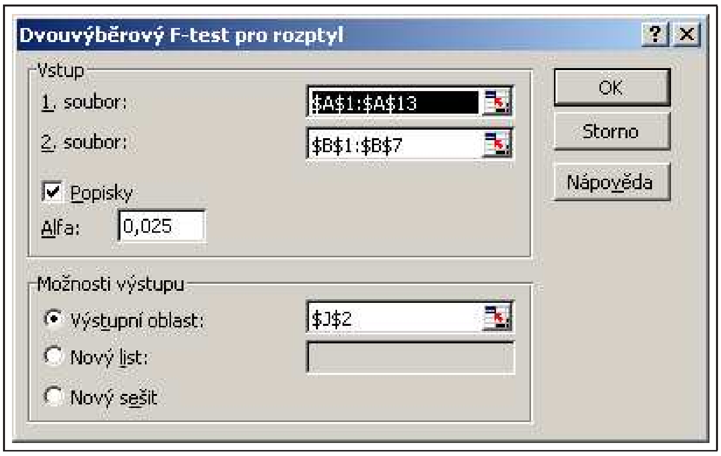
- Stř. hodnota – aritmetický průměr hodnot jednotlivých výběrových souborů
- Rozptyl – výběrový rozptyl souborů
- Pozorování – počet prvků ve výběrových souborech
- Hyp. rozdíl stř. hodnot – testovaná hypotéza
- Rozdíl – počet stupňů volnosti
- t stat – hodnota testového kritéria
- P(T<=t) (1) – p-hodnota pro jednostrannou variantu testu
- t krit (1) – kritická hodnota Studentova rozdělení pro jednostrannou variantu testu
- P(T<=t) (2) – p-hodnota pro oboustrannou variantu testu
- t krit (2) – kritická hodnota Studentova rozdělení pro oboustrannou variantu testu

Pozn. Jednotlivé buňky označují prakticky stejné hodnoty jako v případě t-testu. Pro oboustrannou variantu testu je však nutno v zadání analytického nástroje zadat v políčku „alfa“ hodnotu odpovídající $\alpha/2$, neboť nástroj standardně počítá jednostrannou variantu testu. Dostaneme pak dolní mez

Dvouvýběrový F-test pro rozptyl		
	2 děti	4 děti
Stř. hodnota	38,99166667	41,91667
Rozptyl	5,726287879	6,197667
Pozorování	12	6
Rozdíl	11	5
F	0,923942539	
P(F<=f) (1)	0,421629198	
F krit (1)	0,24727953	

Tabulka 7.2: Dvouvýběrový F-test pro rozptyl pro test „bílé techniky“ (příklad 7.2).

kritického oboru pro oboustranný test. Je proto vhodné se opět orientovat především podle p-hodnoty (v buňce P(F<=f)), jejíž hodnota musí být nižší než 0,05 (pro hladinu významnosti 5%).



Obrázek 7.10: Zadání analytického nástroje pro dvouvýběrový F-test pro rozptyl.

Dvouvýběrový t-test s rovností rozptylů			
		2 děti	4 děti
	Stř. hodnota	38,992	41,917
	Rozptyl	5,726	6,198
	Pozorování	12	6
	Společný rozptyl	5,874	
	Hyp. rozdíl stř. hodnot	0	
	Rozdíl	16	
	t stat	−2,414	
jednostranná	P(T<=t) (1)	0,014	
	t krit (1)	1,746	
oboustranná	P(T<=t) (2)	0,028	
	t krit (2)	2,120	

Tabulka 7.3: Dvouvýběrový t-test s rovností rozptylů pro test „bílé techniky“ (příklad 7.2).