6. cvičení: Úlohy o dvou nezávislých náhodných výběrech z normálního a alternativního rozložení.

Příklad 1.: Jsou dány dva nezávislé náhodné výběry o rozsazích $n_1 = 25$, $n_2 = 10$, první pochází z rozložení $N(\mu_1, \sigma_1^2)$, druhý z rozložení $N(\mu_2, \sigma_2^2)$, kde parametry μ_1, μ_2, σ_1^2, σ_2^2 neznáme. Byly vypočteny realizace výběrových rozptylů: $s_1^2 = 1{,}7482$, $s_2^2 = 1{,}7121$. Na hladině významnosti 0,05 testujte hypotézu, že neznámé rozptyly σ_1^2 a σ_2^2 jsou shodné proti oboustranné alternativě. Test proveďte pomocí intervalu spolehlivosti.

Výsledek: Testujeme H_0: $\frac{\sigma_1^2}{\sigma_2^2} = 1$ proti H_1: $\frac{\sigma_1^2}{\sigma_2^2} \neq 1$

Protože číslo 1 leží v intervalu (0,28; 2,76), H_0 nezamítáme na hladině významnosti 0,05.

Příklad 2.: Výrobce limonád chtěl zjistit, zda změna technologie výroby se projeví v prodeji limonád. Proto sledoval po 14 náhodně vybraných dnů před zavedením nových limonád tržby v určitém regionu a zjistil, že za den utržil v průměru 39 600 Kč se směrodatnou odchylkou 5 060 Kč. Po zavedení nových limonád prověřil stejným způsobem tržby v 11 náhodně vybraných dnech v témž regionu a zjistil průměrný příjem 41 200 Kč se směrodatnou odchylkou 4 310 Kč. Předpokládejte, že tržby za starý typ limonád se řídí rozložením $N(\mu_1, \sigma_1^2)$ a tržby za nový typ limonád se řídí rozložením $N(\mu_2, \sigma_2^2)$.

a) Na hladině významnosti 0,05 testujte hypotézu H_0: $\frac{\sigma_1^2}{\sigma_2^2} = 1$ proti H_1: $\frac{\sigma_1^2}{\sigma_2^2} \neq 1$.

b) Na hladině významnosti 0,05 testujte hypotézu H_0: $\mu_1 - \mu_2 = 0$ proti H_1: $\mu_1 - \mu_2 \neq 0$.

Výsledky:

ad a)

Protože testové kritérium 1,3783 se nerealizuje v kritickém oboru

W = $\left(0; 0{,}3077\right\rangle \cup \left\langle 3{,}5832; \infty\right)$, nelze na hladině významnosti 0,05 zamítnout hypotézu o shodě rozptylů.

ad b)

Protože testové kritérium -0,8363 se nerealizuje v kritickém oboru

W = $\left(-\infty; -2{,}0687\right\rangle \cup \left\langle 2{,}0687; \infty\right)$, na hladině významnosti 0,05 nelze zamítnout hypotézu o shodě středních hodnot.

Příklad 3.: Bylo vylosováno 11 stejně starých selat téhož plemene. Šesti z nich byla předepsána výkrmná dieta č. 1 a zbylým pěti výkrmná dieta č. 2. Průměrné denní přírůstky v Dg za dobu půl roku jsou následující:

dieta č. 1: 62, 54, 55, 60, 53, 58

dieta č. 2: 52, 56, 49, 50, 51.

Zjištěné hodnoty považujeme za realizace dvou nezávislých náhodných výběrů pocházejících z rozložení $N(\mu_1, \sigma_1^2)$ a $N(\mu_2, \sigma_2^2)$.

a) Sestrojte 95% empirický interval spolehlivosti pro podíl rozptylů.

b) Za předpokladu, že data pocházejí z rozložení $N(\mu_1, \sigma^2)$ a $N(\mu_2, \sigma^2)$, sestrojte 95% empirický interval spolehlivosti pro rozdíl středních hodnot $\mu_1 - \mu_2$.

Pro usnadnění výpočtů máte k dispozici následující číselné charakteristiky:

$m_1 = 57$, $m_2 = 51{,}6$, $s_1^2 = 12{,}8$, $s_2^2 = 7{,}3$.

Výsledky:

ad a) $0{,}1872 < \frac{\sigma_1^2}{\sigma_2^2} < 12{,}9541$ s pravděpodobností aspoň 0,95.
ad b) 0,99 Dg < $\mu_1 - \mu_2$ < 9,81 Dg s pravděpodobností aspoň 0,95.

Příklad 4.: Pro údaje z 3. příkladu testujte na hladině významnosti 0,05 hypotézu, že
a) rozptyly hmotnostních přírůstků selat při obou výkrmných dietách jsou shodné
b) obě výkrmné diety mají stejný vliv na hmotnostní přírůstky selat.
Testy proveďte všemi třemi způsoby.

Výsledky:

ad a) Na hladině významnosti 0,05 testujeme hypotézu H_0: $\frac{\sigma_1^2}{\sigma_2^2} = 1$ proti H_1: $\frac{\sigma_1^2}{\sigma_2^2} \neq 1$.
Testování pomocí kritického oboru:

Protože realizace testového kritéria 1,7534 nepatří do kritického oboru
$W = (0; 01354\rangle \cup \langle 9{,}3645; \infty)$, nelze na hladině významnosti 0,05 zamítnout hypotézu o shodě rozptylů.

Testování pomocí intervalu spolehlivosti:
Protože interval (0,1872; 12,9541) obsahuje číslo 1, H_0 nezamítáme na hladině významnosti 0,05.

Testování pomocí p-hodnoty:
Protože p-hodnota = 0,6064 je větší než hladina významnosti 0,05, H_0 nezamítáme na hladině významnosti 0,05.

ad b) Na hladině významnosti 0,05 testujeme hypotézu H_0: $\mu_1 - \mu_2 = 0$ proti H_1: $\mu_1 - \mu_2 \neq 0$.

Testování pomocí kritického oboru:
Protože testové kritérium 2,7712 patří do kritického oboru $W = (-\infty; -2{,}2622\rangle \cup \langle 2{,}2622; \infty)$, H_0 zamítáme na hladině významnosti 0,05. Znamená to, že s rizikem omylu nejvýše 5% jsme prokázali rozdíl v účinnosti obou výkrmných diet.

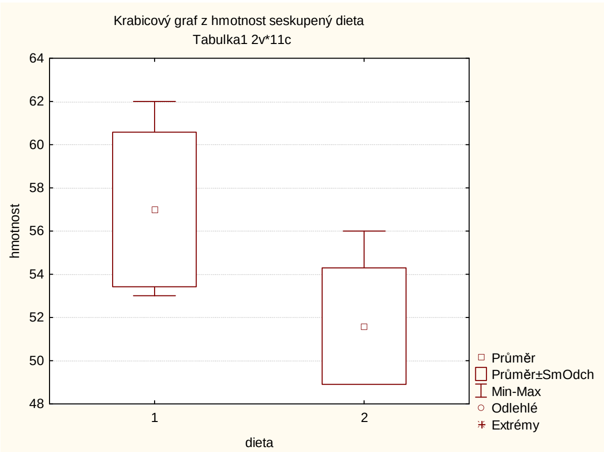
Testování pomocí intervalu spolehlivosti:
Protože interval (0,99; 9,81) neobsahuje nulu, H_0 zamítáme na hladině významnosti 0,05.

Testování pomocí p-hodnoty:
Protože p-hodnota 0,0218 je menší než hladina významnosti 0,05, H_0 zamítáme na hladině významnosti 0,05.

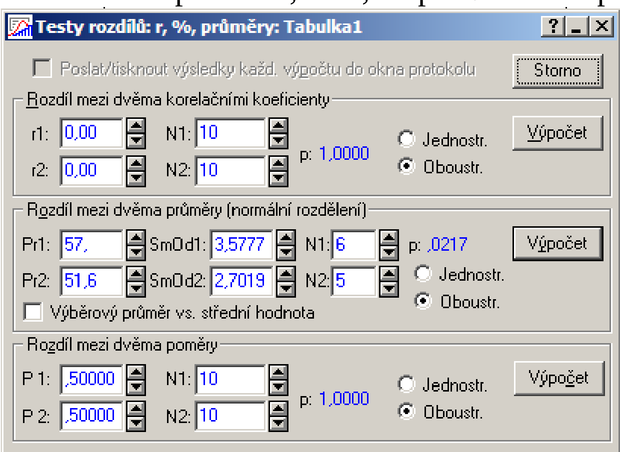
Výpočet pomocí systému STATISTICA
Provedeme dvouvýběrový t-test současně s testem o shodě rozptylů:
Statistika – Základní statistiky a tabulky – t-test, nezávislé, dle skupin – OK, Proměnné – Závislé proměnné hmotnost, Grupovací proměnná dieta – OK.

Proměnná	t-testy; grupováno: dieta (Tabulka1) Skup. 1: 1 Skup. 2: 2										
	Průměr 1	Průměr 2	t	sv	p	Poč.plat 1	Poč.plat. 2	Sm.odch. 1	Sm.odch. 2	F-poměr Rozptyly	p Rozptyly
hmotnost	57,00000	51,60000	2,771222	9	0,021710	6	5	3,577709	2,701851	1,753425	0,606345

Testová statistika pro test shody rozptylů se realizuje hodnotou 1,7534, odpovídající p-hodnota je 0,6063, tedy na hladině významnosti 0,05 nezamítáme hypotézu o shodě rozptylů. (Upozornění: v případě zamítnutí hypotézy o shodě rozptylů je zapotřebí v tabulce t-testu pro nezávislé vzorky dle skupin zaškrtnout volbu Test se samostatnými odhady rozptylu.)
Dále z tabulky plyne, že testová statistika pro test shody středních hodnot se realizuje hodnotou 2,7712, počet stupňů volnosti je 9, odpovídající p-hodnota 0,0217, tedy hypotézu o shodě středních hodnot zamítáme na hladině významnosti 0,05. Znamená to, že s rizikem omylu nejvýše 5 % se prokázalo, že obě výkrmné diety se liší účinností.
Tabulku ještě doplníme krabicovými diagramy. Na záložce Detaily zaškrtneme krabicový graf a vybereme volbu Průměr/SmOdch/Min-Max.



Upozornění: Dvouvýběrový t-test lze v systému STATISTICA provést ještě jiným způsobem, který je vhodný zvláště tehdy, známe-li realizace výběrových průměrů a výběrových směrodatných odchylek.
Statistiky – Základní statistiky a tabulky – Testy rozdílů: r, %, průměry – OK – vybereme Rozdíl mezi dvěma průměry (normální rozdělení) – do políčka Pr1 napíšeme 57, do políčka SmOd1 napíšeme 3,5777, do políčka N1 napíšeme 6, do políčka Pr2 napíšeme 51,6, do políčka SmOd1 napíšeme 2,7019, do políčka N1 napíšeme 5 - Výpočet.



Dostaneme p-hodnotu 0,0217, tedy zamítáme nulovou hypotézu na hladině významnosti 0,05.

Příklad 5.: Při výstupní kontrole bylo náhodně vybráno 150 výrobků vyrobených na ranní směně a rovněž 150 výrobků vyrobených na odpolední směně. U ranní směny bylo zjištěno 16 zmetků a u odpolední 12 zmetků. Sestrojte 95% asymptotického interval spolehlivosti pro rozdíl pravděpodobností vyrobení zmetku v obou směnách.

Výsledek: S pravděpodobností přibližně 0,95 platí, že –0,039 < $\vartheta_1 - \vartheta_2$ < 0,092.

Příklad 6.: Pro údaje z příkladu 5 testujte na asymptotické hladině významnosti 0,05 hypotézu, že pravděpodobnost vyrobení zmetků v obou směnách je táž.

Výsledek: Protože 95% asymptotický interval spolehlivosti pro $\vartheta_1 - \vartheta_2$, jehož meze jsme vypočetli v příkladu 5 a zjistili jsme, že jsou -0,039 a 0,092, obsahuje 0, nezamítáme nulovou hypotézu na asymptotické hladině významnosti 0,05.

Výpočet pomocí systému STATISTICA
Statistiky – Základní statistiky a tabulky – Testy rozdílů: r, %, průměry – OK – vybereme Rozdíl mezi dvěma poměry – do políčka P 1 napíšeme 0,1067, do políčka N1 napíšeme 150, do políčka P 2 napíšeme 0,08, do políčka N2 napíšeme 150 –Výpočet. Dostaneme p-hodnotu 0,4274, tedy nezamítáme nulovou hypotézu na hladině významnosti 0,05.