

10 Dvouvýběrové parametrické testy

V předchozí kapitole jsme se zaměřili na situace, kdy jsme porovnávali vybraný parametr jednoho náhodného výběru (ať už parametr μ nebo σ^2 , za předpokladu, že náhodný výběr pochází z normálního rozdělení, parametr ρ za předpokladu, že náhodný výběr pochází z dvourozměrného normálního rozdělení, nebo parametr p , za předpokladu, že data pochází z alternativního rozdělení) s konkrétní hodnotou získanou například z literatury.

V této kapitole se zaměříme na situaci, kdy vzájemně porovnááme dva různé navzájem nezávislé náhodné výběry, popisující stejný znak (výšku, hmotnost novorozence, délku femuru, největší šířku mozkovny, apod.). K tomu nám, za splnění určitých předpokladů, poslouží dvouvýběrové parametrické testy.

Vzhledem k tomu, že náhodný výběr je reprezentantem vybrané populace, umožňují nám dvouvýběrové parametrické testy porovnat prostřednictvím náhodných výběrů a procesem testování hypotéz navzájem dvě populace. Konkrétním příkladem na využití dvouvýběrových testů je porovnávání znaků pohlavního dimorfismu (viz příklad ??).

Příklad 10.1. Příklad aplikace dvouvýběrového parametrického testu

Předpokládejme, že chceme vytvořit studii porovnávající výšku mužů a žen. V rámci studie tedy chceme porovnávat dvě populace, a sice populaci mužů a populaci žen. Sledovaným znakem (vlastností) je výška postavy. Pro první populaci vytvoříme reprezentativní vzorek změřením výšky postavy n_1 mužů. Získáme tak první náhodný výběr X_{11}, \dots, X_{1n_1} . Pro druhou populaci vytvoříme reprezentativní vzorek změřením výšky postavy n_2 žen. Získáme tak druhý náhodný výběr X_{21}, \dots, X_{2n_2} . Poznamenejme, že rozsahy náhodných výběrů n_1 a n_2 se mohou lišit, tj. můžeme mít jiný počet naměřených výšek mužů a jiný počet naměřených výšek žen. Naším cílem by potom mohlo být porovnání střední hodnoty výšky mužů se střední hodnotou výšky žen pomocí procesu testování hypotéz. ★

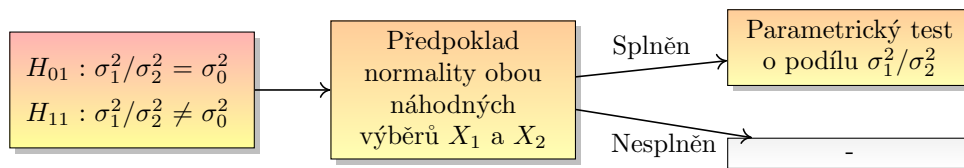
Dalším příkladem využití dvouvýběrových testů je například porovnání téhož znaku (šířka nadočnicového oblouku, délka levé klíční kosti, interorbitální šířka, apod.) u dvou populací (například malajské a čínské populace, německé a bantuské populace, či řecké a indické populace, apod.), nebo u dvou různých sexuálních orientací (heterosexuální populace a homosexuální populace).

Metody uvedené v této kapitole jsou převážně rozšířením teorie uvedené v kapitole ?? na situaci dvou náhodných výběrů. Konkrétně si zde představíme test o podílu rozptylů σ_1^2/σ_2^2 dvou náhodných výběrů pocházejících z normálních rozdělení $N(\mu_1, \sigma_1^2)$ a $N(\mu_2, \sigma_2^2)$, test o rozdílu středních hodnot $\mu_1 - \mu_2$ dvou náhodných výběrů pocházejících z normálních rozdělení $N(\mu_1, \sigma_1^2)$ a $N(\mu_2, \sigma_2^2)$, kde rozptyly σ_1^2 a σ_2^2 jsou neznámé, ale shodné, tj. $\sigma_1^2 = \sigma_2^2 = \sigma^2$, a test o rozdílu středních hodnot $\mu_1 - \mu_2$ dvou náhodných výběrů pocházejících z normálních rozdělení $N(\mu_1, \sigma_1^2)$ a $N(\mu_2, \sigma_1^2)$, kde rozptyly σ_1^2 a σ_2^2 jsou neznámé a různé. Dále si představíme test o rozdílu dvou korelačních koeficientů $\rho_1 - \rho_2$ dvou náhodných výběrů z dvourozměrných normálních rozdělení $N_2(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$ a $N_2(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$ a test o rozdílu pravděpodobností $p_1 - p_2$ dvou náhodných výběrů z alternativních rozdělení $\text{Alt}(p_1)$ a $\text{Alt}(p_2)$.

Metody uvedené v této kapitole, jsou, stejně jako metody uvedené v kapitole ??, tzv. parametrickými metodami. To znamená, že stejně, jako v předchozí kapitole předpokládáme, že sledovaný znak zkoumaný ať už v první nebo v druhé populaci, pochází z nějakého známého rozdělení (normálního nebo alternativního) s příslušnými parametry $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho_1, \rho_2, p_1, p_2)$ a na hladině významnosti α testujeme potom hypotézy o těchto vybraných parametrech.

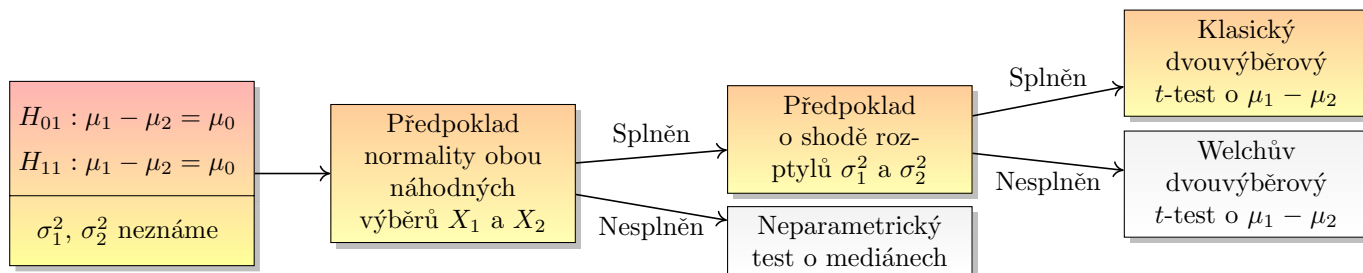
V případě testu o podílu rozptylů σ_1^2/σ_2^2 předpokládáme, že znak sledovaný v dané populaci pochází z normálního rozdělení a tedy také reprezentativní vzorky naměřených hodnot tohoto znaku v první i v druhé populaci pochází z normálního rozdělení. Proto je potřeba před samotným testem o podílu rozptylů σ_1^2/σ_2^2 provést test normality pro každý náhodný výběr zvlášť. Pokud náhodný výběr, reprezentující ať už první nebo druhou populaci, nepochází z normálního rozdělení, nemůžeme parametrický test provést. Neparametrickou alternativu testu o podílu rozptylů si v tomto textu neuvádíme. Pokud naopak oba náhodné výběry splňují předpoklad normality, provedeme parametrický test o podílu rozptylů σ_1^2/σ_2^2 (viz sekce 10.1). Pro lepší orientaci uvádíme rozhodovací cestu k použití parametrického testu o podílu rozptylů v závislosti na splnění předpokladu normality na obrázku 1.

V případě testu o rozdílu středních hodnot $\mu_1 - \mu_2$, kde rozptyly σ_1^2 a σ_2^2 jsou neznámé, ale shodné, musíme rovněž nejprve otestovat normalitu obou náhodných výběrů. Pokud alespoň jeden z náhodných výběrů nepochází z normálního rozdělení, nemůžeme parametrický test provést a musíme použít neparametrický test o mediánech (viz kapitola ??). Pokud naopak oba náhodné výběry splňují předpoklad normality, pokračujeme cestou parametrického testování. Druhým předpokladem, který musíme ověřit, je předpoklad shody obou rozptylů σ_1^2 a σ_2^2 . K tomu



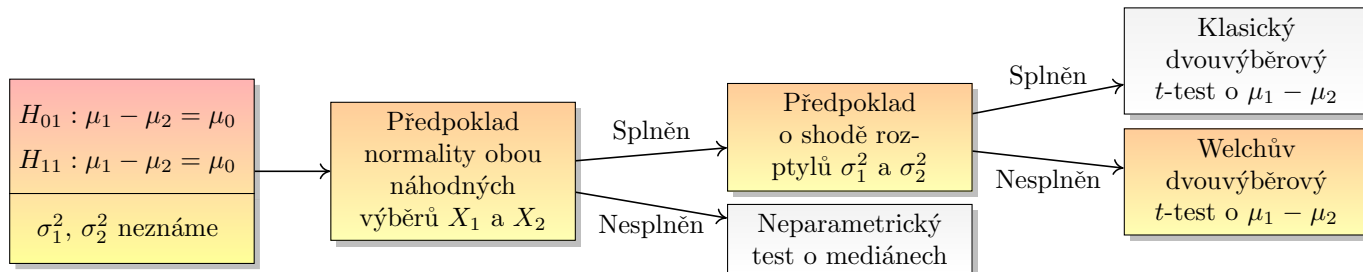
Obrázek 1: Rozhodovací cesta k použití parametrického testu o podílu rozptylů σ_1^2/σ_2^2

využijeme výše zmíněný test o podílu rozptylů. Pokud na základě testu o podílu rozptylů dojdeme k závěru, že rozptyly jsou shodné, tj. $\sigma_1^2 = \sigma_2^2$, použijeme na otestování hypotézy o rozdílu středních hodnot klasický dvouvýběrový t -test o rozdílu středních hodnot $\mu_1 - \mu_2$ (viz sekce 10.2). Pro lepší orientaci uvádíme rozhodovací cestu k použití klasického dvouvýběrového t -testu na obrázku 2.



Obrázek 2: Rozhodovací cesta k použití klasického dvouvýběrového t -testu o rozdílu středních hodnot $\mu_1 - \mu_2$

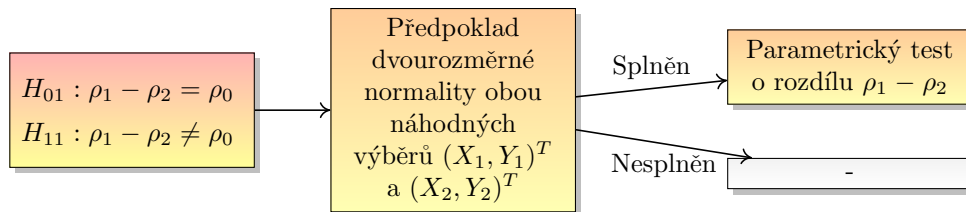
Pokud na základě testu o podílu rozptylů dojdeme k závěru, že rozptyly nejsou shodné, tj. $\sigma_1^2 \neq \sigma_2^2$, použijeme na otestování hypotézy o rozdílu středních hodnot Welchův dvouvýběrový t -test o rozdílu středních hodnot $\mu_1 - \mu_2$ (viz sekce 10.3). Pro lepší orientaci uvádíme rozhodovací cestu k použití Welchova dvouvýběrového t -testu na obrázku 3.



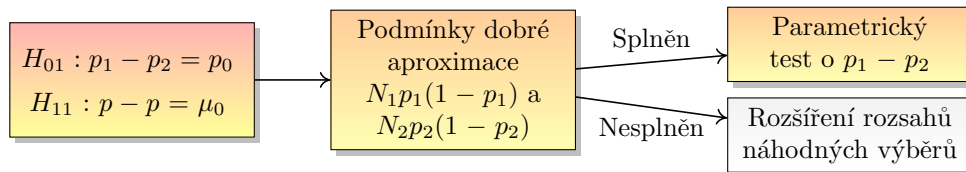
Obrázek 3: Rozhodovací cesta k použití Welchova dvouvýběrového t -testu o rozdílu středních hodnot $\mu_1 - \mu_2$

V případě testu o rozdílu dvou korelačních koeficientů $\rho_1 - \rho_2$ předpokládáme, že oba náhodné výběry pochází z dvourozměrných normálních rozdělení. Proto před samotným testem o podílu rozptylů musíme otestovat dvourozměrnou normalitu zvlášť pro každý náhodný výběr. Pokud alespoň jeden z obou náhodných výběrů nesplňuje předpoklad dvourozměrné normality, nemůžeme parametrický test o rozdílu dvou korelačních koeficientů provést. Neparametrickou alternativu testu o rozdílu rozdílu korelačních koeficientů si v tomto textu neuvádíme. Na obrázku 4 je zobrazena rozhodovací cesta k použití parametrického testu o rozdílu korelačních koeficientů v závislosti na splnění předpokladů dvourozměrné normality obou náhodných výběrů.

Konečně, v případě testu o rozdílu dvou pravděpodobností $p_1 - p_2$ předpokládáme, že oba náhodné výběry pochází z alternativních rozdělení. Tento předpoklad však nijak netestujeme, neboť charakter alternativních dat je zřejmý na první pohled. Test o rozdílu dvou pravděpodobností $p_1 - p_2$ je však asymptotickým testem, což znamená, že závěry testování se stávají spolehlivými s dostatečně velkým rozsahem obou náhodných výběrů. Zda jsou rozsahy obou náhodných výběrů k otestování nulové hypotézy dostačující, ověříme podmínkou dobré aproximace (viz sekce 10.5). V případě, že podmínka dobré aproximace není splněna, je třeba před samotným testem rozšířit oba datové soubory, tak aby rozsah každého z nich byl dostačující. Pro lepší orientaci uvádíme rozhodovací cestu k použití testu o rozdílu pravděpodobností $p_1 - p_2$ na obrázku 5.



Obrázek 4: Rozhodovací cesta k použití parametrického testu o rozdílu korelačních koeficientů $\rho_1 - \rho_2$



Obrázek 5: Rozhodovací cesta k použití testu o rozdílu pravděpodobností $p_1 - p_2$

10.1 Test o podílu rozptylů σ_1^2 a σ_2^2

Nechť X_{11}, \dots, X_{1n_1} je náhodný výběr z $N(\mu_1, \sigma^2)$, a X_{21}, \dots, X_{2n_2} je na něm nezávislý náhodný výběr z rozdělení $N(\mu_2, \sigma^2)$, přičemž $n_1 \geq 2$, $n_2 \geq 2$. Na hladině významnosti α testujeme jednu z následujících tří hypotéz oproti příslušné alternativní hypotéze.

$$\begin{array}{lll} H_{01} : \sigma_1^2/\sigma_2^2 = \sigma_0^2 & \text{oproti} & H_{11} : \sigma_1^2/\sigma_2^2 \neq \sigma_0^2 \quad (\text{oboustranná alt.}) \\ H_{02} : \sigma_1^2/\sigma_2^2 \leq \sigma_0^2 & \text{oproti} & H_{12} : \sigma_1^2/\sigma_2^2 > \sigma_0^2 \quad (\text{pravostranná alt.}) \\ H_{03} : \sigma_1^2/\sigma_2^2 \geq \sigma_0^2 & \text{oproti} & H_{13} : \sigma_1^2/\sigma_2^2 < \sigma_0^2 \quad (\text{levostranná alt.}) \end{array}$$

Test nazýváme dvouvýběrový F -test o podílu rozptylů σ_1^2/σ_2^2 . Testovací statistika má tvar

$$F_W = \frac{S_1^2}{S_2^2}, \quad (10.1)$$

kde S_1^2 je výběrový rozptyl prvního náhodného výběru a S_2^2 je výběrový rozptyl druhého náhodného výběru. Za platnosti nulové hypotézy pochází statistika F_W z Fisherova-Snedecorova F -rozdělení o $n_1 - 1$ a $n_2 - 1$ stupních volnosti, tj.

$$F_W = \frac{S_1^2}{S_2^2} \stackrel{H_0}{\sim} F_{n_1-1, n_2-1}.$$

Kritický obor podle zvolené alternativní hypotézy má tvar

$$\begin{array}{ll} H_{11} : \sigma_1^2/\sigma_2^2 \neq \sigma_0^2 & W = (0; F_{n_1-1, n_2-1}(\alpha/2)) \cup (F_{n_1-1, n_2-1}(1 - \alpha/2); \infty) \\ H_{12} : \sigma_1^2/\sigma_2^2 > \sigma_0^2 & W = (F_{n_1-1, n_2-1}(1 - \alpha); \infty) \\ H_{13} : \sigma_1^2/\sigma_2^2 < \sigma_0^2 & W = (0; F_{n_1-1, n_2-1}(\alpha)) \end{array}$$

kde $F_{n_1-1, n_2-1}(\alpha/2)$, $F_{n_1-1, n_2-1}(1 - \alpha/2)$, $F_{n_1-1, n_2-1}(\alpha)$, $F_{n_1-1, n_2-1}(1 - \alpha)$ jsou kvantily F rozdělení o $n_1 - 1$ a $n_2 - 1$ stupních volnosti, jejichž hodnoty získáme pomocí softwaru \mathbb{R} a implementované funkce $\text{qf}()$.

Interval spolehlivosti má podle zvolené alternativní hypotézy jeden z následujících tvarů

$$\begin{array}{ll} H_{11} : \sigma_1^2/\sigma_2^2 \neq \sigma_0^2 & (d, h) = \left(\frac{s_1^2/s_2^2}{F_{n_1-1, n_2-1}(1 - \alpha/2)}; \frac{s_1^2/s_2^2}{F_{n_1-1, n_2-1}(\alpha/2)} \right) \\ H_{12} : \sigma_1^2/\sigma_2^2 > \sigma_0^2 & (d, \infty) = \left(\frac{s_1^2/s_2^2}{F_{n_1-1, n_2-1}(1 - \alpha)}; \infty \right) \\ H_{13} : \sigma_1^2/\sigma_2^2 < \sigma_0^2 & (0, h) = \left(0; \frac{s_1^2/s_2^2}{F_{n_1-1, n_2-1}(\alpha)} \right) \end{array}$$

Poznámka: Protože parametry σ_1^2 i σ_2^2 jsou z definice větší než 0, je i jejich podíl σ_1^2/σ_2^2 vždy větší než 0. Proto pravostranný interval spolehlivosti omezíme zdola hodnotou 0, namísto minus nekonečnem.

p -hodnota má v závislosti na zvolené alternativní hypotéze jeden z následujících tvarů

$$\begin{array}{ll} H_{11} : \sigma_1^2/\sigma_2^2 \neq \sigma_0^2 & p\text{-hodnota} = 2 \min\{\Pr(F_W \leq f_W), \Pr(F_W > f_W)\} \\ H_{12} : \sigma_1^2/\sigma_2^2 > \sigma_0^2 & p\text{-hodnota} = \Pr(F_W > f_W) = 1 - \Pr(F_W \leq f_W) \\ H_{13} : \sigma_1^2/\sigma_2^2 < \sigma_0^2 & p\text{-hodnota} = \Pr(F_W \leq f_W) \end{array}$$

kde F_W je náhodná veličina, f_W je realizace testovací statistiky F_W (viz vzorec 10.1), tedy konkrétní číslo, a $\Pr(F_W \leq f_W)$ je distribuční funkce F rozdělení o $n_1 - 1$ a $n_2 - 1$ stupních volnosti, jejíž hodnotu získáme pomocí \mathbb{R} a implementované funkce $\text{pf}()$.

Příklad 10.2. Test o podílu rozptylů

Mějme datový soubor 16-anova-head.txt a proměnnou bigo.W popisující šířku dolní čelisti v mm (viz sekce ??). Předpokládejme, že náhodná veličina X , popisující šířku dolní čelisti u mužů orientovaných jinak než výlučně heterosexuálně, pochází z normálního rozdělení $N(\mu_1, \sigma_1^2)$, a že náhodná veličina Y , popisující šířku dolní čelisti u žen orientovaných jinak než výlučně heterosexuálně, pochází z normálního rozdělení $N(\mu_2, \sigma_2^2)$. Na hladině významnosti $\alpha = 0.05$ otestujte hypotézu o shodě rozptylů σ_1^2 a σ_2^2 .

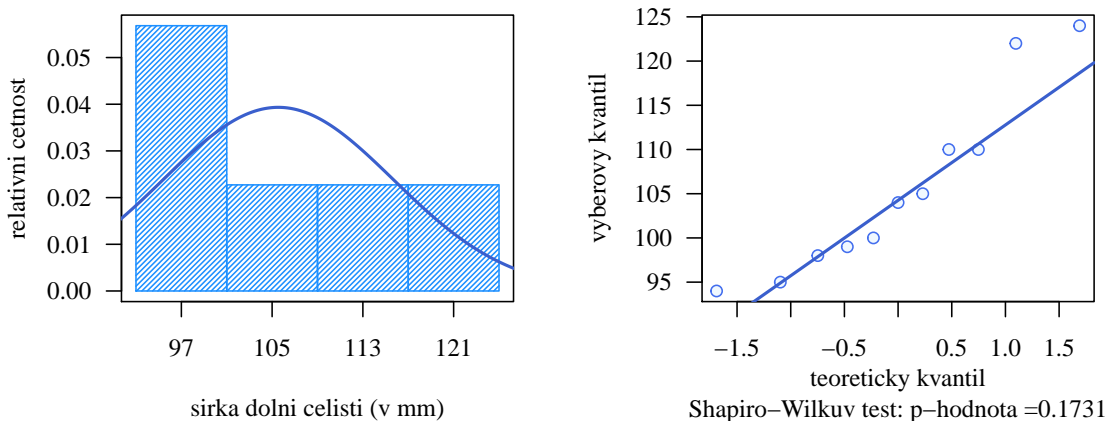
Řešení příkladu 10.2

Pomocí příkazu read.delim() načteme datový soubor a příkazem na.omit() odstraníme ze souboru NA hodnoty. Pomocí operátoru [] vybereme z datové tabulky údaje o šířce dolní čelisti (bigo.W) u mužů (sex == 'm'), resp. u žen (sex == 'f'), orientovaných jinak než výlučně heterosexuálně (sexor == 'sa'). Příkazem length() zjistíme rozsahy obou náhodných výběrů.

```
1 data <- read.delim('00-Data//16-anova-head.txt')
2 data <- na.omit(data)
3 #head(data, n = 3)
4 bigo.Wms <- data[data$sex == 'm' & data$sexor == 'sa', 'bigo.W']
5 bigo.Wfs <- data[data$sex == 'f' & data$sexor == 'sa', 'bigo.W']
6
7 n1 <- length(bigo.Wms) # 11
8 n2 <- length(bigo.Wfs) # 23
```

Datový soubor obsahuje celkem 11 údajů o šířce dolní čelisti u mužů orientovaných jinak než heterosexuálně a 23 údajů o šířce dolní čelisti u žen orientovaných jinak než heterosexuálně. Řešení příkladu vede na test o podílu rozptylů. Před samotným testováním hypotézy ze zadání musíme nejprve ověřit splnění předpokladu normality každého náhodného výběru.

Na námi zvolené hladině významnosti $\alpha = 0.05$ testujeme nulovou hypotézu H_{01} : *Náhodný výběr šířek dolní čelisti mužů pochází z normálního rozdělení.* oproti alternativní hypotéze H_{11} : *Náhodný výběr šířek dolní čelisti mužů nepochází z normálního rozdělení.* Jelikož náhodný výběr sestává z 11 pozorování, což je méně než 30, provedeme test normality Shapiro-Wilkovým testem. Normalitu ověříme též graficky vykreslením kvantilového diagramu (viz sekce ??) a pomocí histogramu superponovaného křivkou normálního rozdělení (viz sekce ??). Datový soubor rozdělíme na základě Sturjerova pravidla (viz sekce ??) do čtyř ekvidistatních intervalů s šířkou 8 mm prostřednictvím stanovených hranic 93, 101, ..., 125 (viz obrázek ??).

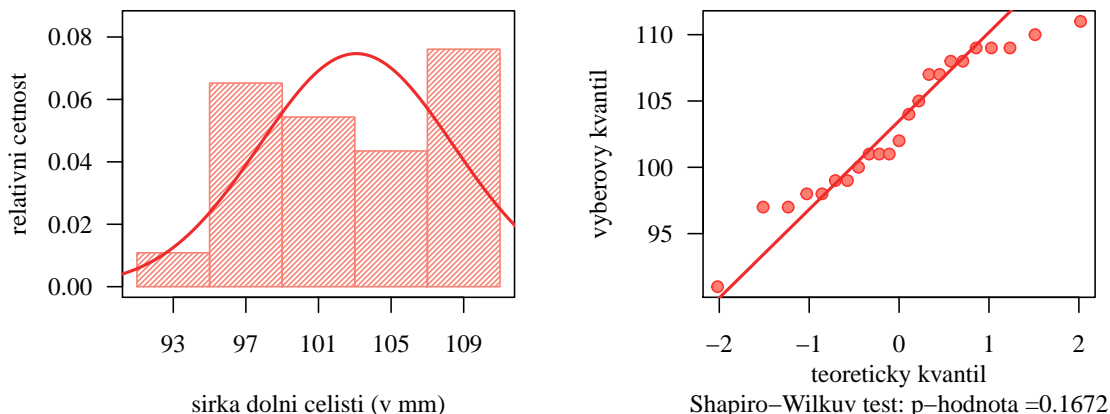


Obrázek 6: Histogram a kvantilový diagram šířky dolní čelisti mužů

Protože p -hodnota = 0.1731 je větší než 0.05, nulovou hypotézu nezamítáme na hladině významnosti $\alpha = 0.05$. Náhodný výběr výšek lebky žen tedy pochází z normálního rozdělení. Při pohledu na histogram bychom možná o normálním rozdělení dat mohli pochybovat. Nezapomeňme však, že náhodný výběr obsahuje pouze 11 pozorování. Při tak malém počtu hodnot je obtížné posuzovat normalitu hodnot graficky, proto se spoléháme na výsledek

Shapiro-Wilcoxonova testu.

Analogicky testujeme na hladině významnosti $\alpha = 0.05$ nulovou hypotézu H_{02} : *Náhodný výběr šířek dolní čelisti žen pochází z normálního rozdělení*, oproti alternativní hypotéze H_{12} : *Náhodný výběr šířek dolní čelisti žen nepochází z normálního rozdělení*. Kvůli nízkému rozsahu náhodného výběru otestujeme normalitu Shapiro-Wilkovým testem. V rámci histogramu rozdělíme soubor do pěti ekvidistatních intervalů s šířkou 4 mm prostřednictvím stanovených hranic 91, 95, ..., 111 (viz obrázek 7).



Obrázek 7: Histogram a kvantilový diagram šířky dolní čelisti žen

Protože p -hodnota = 0.0.1672 je větší než 0.05, nulovou hypotézu nezamítáme na hladině významnosti $\alpha = 0.05$. Náhodný výběr šířky dolní čelisti žen také pochází z normálního rozdělení.

Jelikož oba náhodné výběry splňují předpoklad normality, můžeme k testování hypotézy ze zadání použít parametrický test o podílu rozptylů. Řešení si nyní uvedeme v posloupnosti šesti kroků.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : Rozptyl šířky dolní čelisti mužů orientovaných jinak než heterosexuálně a žen orientovaných jinak než heterosexuálně je shodný.

H_1 : Rozptyl šířky dolní čelisti mužů orientovaných jinak než heterosexuálně a žen orientovaných jinak než heterosexuálně není shodný.

- **matematická formulace** nulové a alternativní hypotézy

H_0 : $\sigma_1^2 = \sigma_2^2 \rightarrow \sigma_1^2/\sigma_2^2 = \sigma_0^2$, kde $\sigma_0^2 = 1$

H_1 : $\sigma_1^2 \neq \sigma_2^2 \rightarrow \sigma_1^2/\sigma_2^2 \neq \sigma_0^2$, kde $\sigma_0^2 = 1$ (oboustranná alternativa)

2. Volba hladiny významnosti

- Hladinu významnosti volíme v souladu se zadáním jako $\alpha = 0.05$.

3. Testování kritickým oborem

- **Testovací statistika**

$$F_W = \frac{S_1^2}{S_2^2} = \frac{10.14262^2}{5.342055^2} = \frac{102.8727}{28.53755} = 3.60482 \doteq 3.6048$$

- **Kritický obor**

$$\begin{aligned} W &= (0; F_{n_1-1, n_2-1}(\alpha/2)) \cup (F_{n_1-1, n_2-1}(1 - \alpha/2); \infty) \\ &= (0; F_{10, 22}(0.025)) \cup (F_{10, 22}(0.975); \infty) \\ &= (0; 0.2950131) \cup (2.699813; \infty) \end{aligned}$$

```

9 alpha <- 0.05
10 s1 <- sd(bigo.Wms)
11 s2 <- sd(bigo.Wfs)
12 Fw <- s1^2 / s2^2 # 3.60482

```

```

13 qf(alpha/2, n1 - 1, n2 - 1) # 0.2950131
14 qf(1 - alpha/2, n1 - 1, n2 - 1) # 2.699813

```

- **Závěr testování**

Protože realizace testovací statistiky $f_W = 3.6048$ náleží do kritického oboru, tj. $f_W \in W$, H_0 zamítáme na hladině významnosti $\alpha = 0.05$.

4. Testování intervalem spolehlivosti

- **Interval spolehlivosti**

$$\begin{aligned}
 (d, h) &= \left(\frac{s_1^2/s_2^2}{F_{n_1-1, n_2-1}(1-\alpha/2)}; \frac{s_1^2/s_2^2}{F_{n_1-1, n_2-1}(\alpha/2)} \right) \\
 &= \left(\frac{10.14262^2/5.342055^2}{F_{106, 214}(0.975)}; \frac{10.14262^2/5.342055^2}{F_{106, 214}(0.025)} \right) \\
 &= \left(\frac{3.60482}{2.699813}; \frac{3.60482}{0.2950131} \right) \\
 &= (1.335211; 12.21919) \doteq (1.3352; 12.2192)
 \end{aligned}$$

- **Závěr testování**

Protože $\sigma_0^2 = 1$ nenáleží do Waldova 95% empirického oboustranného intervalu spolehlivosti, tj. $\sigma_0^2 = 1 \notin IS$, H_0 zamítáme na hladině významnosti $\alpha = 0.05$.

5. Testování p -hodnotou

- **p -hodnota**

$$\begin{aligned}
 p\text{-hodnota} &= 2 \min\{\Pr(F_W \leq f_W), \Pr(F_W > f_W)\} \\
 &= 2 \min\{\Pr(F_W \leq 3.60482), 1 - \Pr(F_W \leq 3.60482)\} \\
 &= 2 \min\{0.9941918, 0.005808181\} \\
 &= 2 \times 0.005808181 \\
 &= 0.01161636 \doteq 0.01162
 \end{aligned}$$

```

15 p.val <- 2*min(pf(Fw, n1 - 1, n2 - 1), 1 - pf(Fw, n1 - 1, n2 - 1)) # 0.01161636

```

- **Závěr testování**

Protože p -hodnota = 0.01162 je menší než $\alpha = 0.05$, H_0 zamítáme na hladině významnosti $\alpha = 0.05$.

6. **Interpretace výsledků:** Na základě všech tří způsobů testování zamítáme hypotézu o shodě rozptylů σ_1^2 a σ_2^2 . Mezi rozptylem šířky dolní čelisti mužů orientovaných jinak než heterosexuálně a rozptylem šířky dolní čelisti žen orientovaných jinak než heterosexuálně existuje statisticky významný rozdíl.

Poznámka: K otestování nulové hypotézy o podílu rozptylů můžeme využít funkci `var.test()`. Vstupními parametry budou nejprve dva vektory reprezentující náhodné výběry, tj. `bigo.Wms` a `bigo.Wfs`, dále hodnota hladiny významnosti α zadaná prostřednictvím koeficientu spolehlivosti $1 - \alpha$ nastavením hodnoty argumentu `conf.level = 0.95` a nakonec typ zvolené alternativní hypotézy (oboustranná), zadaný pomocí argumentu `alternative == 'two.sided'`.

```
16 var.test(bigo.Wms, bigo.Wfs, conf.level = 0.95, alternative = 'two.sided')
```

```
      F test to compare two variances

data:  bigo.Wms and bigo.Wfs
F = 3.6048, num df = 10, denom df = 22, p-value = 0.01162
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 1.335211 12.219186
sample estimates:
ratio of variances
      3.60482
```

17
18
19
20
21
22
23
24
25
26
27

Součástí výstupu je hodnota testovací statistiky $F = 3.6048$, počty stupňů volnosti Fisherova rozdělení $\text{num df} = 10$ a $\text{denom df} = 22$, hranice intervalu spolehlivosti 1.335211 a 12.219186 a p -hodnota $p\text{-value} = 0.01162$. Jediné, co musíme stanovit zvlášť, jsou dolní a horní hranice kritického oboru.



Příklad 10.3. Test o podílu rozptylů

Mějme datový soubor 11-two-samples-means-skull.txt a proměnnou (skull.H), popisující basion-bregmatickou výšku lebky v mm (viz sekce ??). Předpokládejme, že náhodná veličina X , popisující výšku lebky u žen, pochází z normálního rozdělení $N(\mu_1, \sigma_1^2)$, a že náhodná veličina Y , popisující výšku lebky u mužů, pochází z normálního rozdělení $N(\mu_2, \sigma_2^2)$. Na hladině významnosti $\alpha = 0.01$ otestujte nulovou hypotézu, že rozptyl výšky lebky u mužů je větší nebo roven výšce u žen.

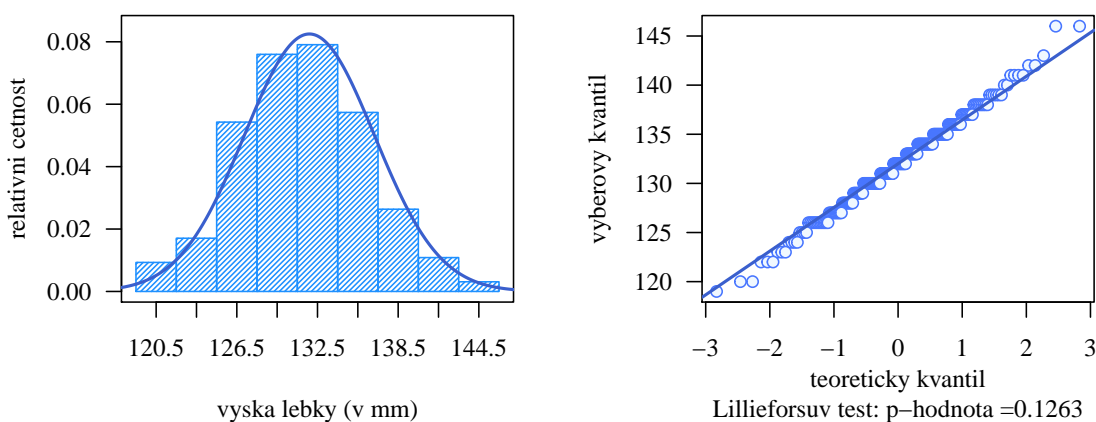
Řešení příkladu 10.3

Pomocí příkazu `read.delim()` načteme datový soubor a příkazem `na.omit()` odtráhneme ze souboru NA hodnoty. Pomocí operátoru `[]` vybereme z datové tabulky údaje o basion-bregmatické výšce lebky (skull.H) u mužů (`sex == 'm'`), resp. u žen (`sex == 'f'`). Dále zjistíme rozsahy obou náhodných výběrů.

```
28 data <- read.delim('00-Data//11-two-samples-means-skull.txt')
29 data <- na.omit(data)
30 #head(data, n = 3)
31 skull.Hm <- data[data$sex == 'm', 'skull.H']
32 skull.Hf <- data[data$sex == 'f', 'skull.H']
33
34 n1 <- length(skull.Hm) # 215
35 n2 <- length(skull.Hf) # 107
```

Datový soubor obsahuje celkem 215 údajů o basion-bregmatické výšce lebky u mužů a 107 údajů o basion-bregmatické výšce lebky u žen. Řešení příkladu vede na test o podílu rozptylů. Před testováním hypotézy ze zadání musíme ověřit předpoklad normality pro oba náhodné výběry.

Jelikož není uvedeno jinak, zvolíme pro test normality hladinu významnosti $\alpha = 0.05$. Nejprve testujeme nulovou hypotézu H_{01} : *Náhodný výběr výšek lebky mužů pochází z normálního rozdělení.* oproti alternativní hypotéze H_{11} : *Náhodný výběr výšek lebky mužů nepochází z normálního rozdělení.* Protože rozsah náhodného výběru $n = 215$ je větší než 30, zvolíme na test normality Lillieforsův test. Normalitu ověříme graficky pomocí kvantilového diagramu a histogramu superponovaného křivkou normálního rozdělení (viz obrázek ??). Datový soubor rozdělíme na základě Sturgerova pravidla do devíti ekvidistantních intervalů s šířkou 3 mm prostřednictvím stanovených hranic 119, 121, ..., 146.

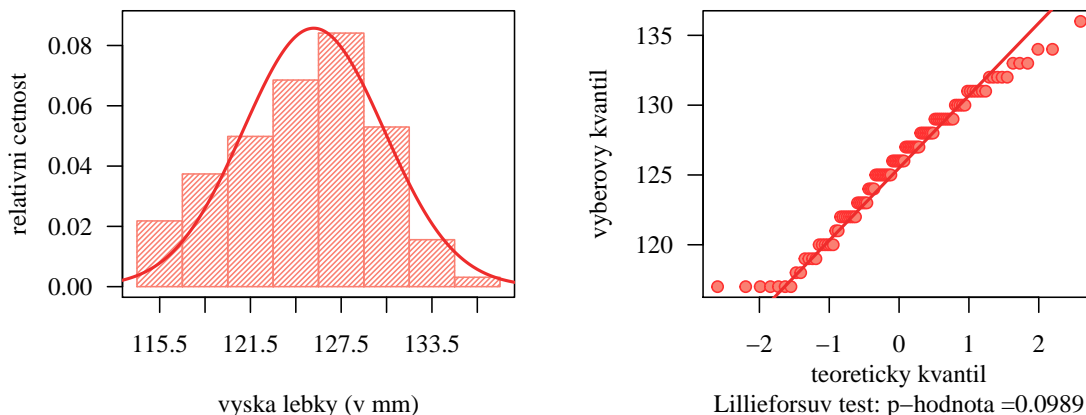


Obrázek 8: Histogram a kvantilový diagram výšky lebky mužů

Protože p -hodnota = 0.1263 je větší než 0.05, nulovou hypotézu nezamítáme na hladině významnosti $\alpha = 0.05$. Náhodný výběr výšek lebky mužů tedy pochází z normálního rozdělení.

Analogicky testujeme na hladině významnosti $\alpha = 0.05$ nulovou hypotézu H_{02} : *Náhodný výběr výšek lebky žen pochází z normálního rozdělení.* oproti alternativní hypotéze H_{12} : *Náhodný výběr výšek lebky žen nepochází z*

normálního rozdělení. Protože rozsah náhodného výběru $n = 107$ je větší než 30, zvolíme na test normality opět Lillieforsův test. V rámci histogramu rozdělíme soubor do osmi ekvidistančních intervalů s šířkou 3 mm prostřednictvím stanovených hranic 114, 117, ..., 138 (viz obrázek 9).



Obrázek 9: Histogram a kvantilový diagram výšky lebky žen

Protože p -hodnota = 0.0989 je větší než 0.05, nulovou hypotézu nezamítáme na hladině významnosti $\alpha = 0.05$. Náhodný výběr výšek lebky žen tedy pochází z normálního rozdělení.

Jelikož oba náhodné výběry pochází z normálních rozdělení, můžeme k testování hypotézy ze zadání použít parametrický test o podílu rozptylů. Zde je vhodné upozornit, že v zadání máme uvedené přesné znění nulové hypotézy. Zbývá tedy vhodně zvolit hypotézu alternativní.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : Rozptyl výšky lebky mužů je větší nebo roven rozptylu výšky lebky žen.

H_1 : Rozptyl výšky lebky mužů je menší než rozptyl výšky lebky žen.

- **matematická formulace** nulové a alternativní hypotézy

H_0 : $\sigma_1^2 \geq \sigma_2^2 \rightarrow \sigma_1^2/\sigma_2^2 \geq \sigma_0^2$, kde $\sigma_0^2 = 1$

H_1 : $\sigma_1^2 < \sigma_2^2 \rightarrow \sigma_1^2/\sigma_2^2 < \sigma_0^2$, kde $\sigma_0^2 = 1$ (levostranná alternativa)

2. Volba hladiny významnosti

- Hladinu významnosti volíme ze zadání jako $\alpha = 0.01$.

3. Testování kritickým oborem

- **Testovací statistika**

$$F_W = \frac{S_1^2}{S_2^2} = \frac{4.835494^2}{4.653256^2} = \frac{23.382}{21.65279} = 1.079861 \doteq 1.0799$$

```
36 alpha <- 0.01
37 s1 <- sd(skull.Hm)
38 s2 <- sd(skull.Hf)
39 Fw <- s1^2 / s2^2 # 1.079861
```

- Kritický obor

$$\begin{aligned} W &= (0; F_{n_1-1, n_2-1}(\alpha)) \\ &= (0; F_{214, 107}(0.01)) \\ &= (0; 0.683192) \doteq (0; 0.6832) \end{aligned}$$

```
40 qf(alpha, n1 - 1, n2 - 1) # 0.683192
```

- Závěr testování

Protože realizace testovací statistiky $f_W = 1.0799$ nenáleží do kritického oboru, tj. $f_W \notin W$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

4. Testování intervalem spolehlivosti

- Interval spolehlivosti

$$\begin{aligned} (0, h) &= \left(0; \frac{s_1^2/s_2^2}{F_{n_1-1, n_2-1}(\alpha)} \right) \\ &= \left(0; \frac{4.835494^2/4.653256^2}{F_{214, 106}(0.01)} \right) \\ &= \left(0; \frac{1.079861}{0.683192} \right) \\ &= (0; 1.580611) \doteq (0; 1.5806) \end{aligned}$$

```
41 HH <- (s1^2 / s2^2) / qf(alpha, n1 - 1, n2 - 1) # 1.580611
```

- Závěr testování

Protože $\sigma_0^2 = 1$ náleží do Waldova 99% empirického pravostanného intervalu spolehlivosti, tj. $\sigma_0^2 = 1 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

5. Testování p -hodnotou

- p -hodnota

$$\begin{aligned} p\text{-hodnota} &= \Pr(F_W \leq f_W) \\ &= \Pr(F_W \leq 0.92060453) \\ &= 0.3165 \end{aligned}$$

```
42 p.val <- pf(Fw, n1 - 1, n2 - 1) # 0.3165
```

- Závěr testování

Protože p -hodnota = 0.3165 je větší než $\alpha = 0.01$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

6. Interpretace výsledků: Rozptyl výšky lebky mužů není statisticky významně vyšší než rozptyl výšky lebky žen.

Poznámka: K otestování nulové hypotézy o podílu rozptylů můžeme využít funkci `var.test()`, kde hodnotu argumentu `conf.level` nastavíme na hodnotu 0.99 a typ alternativní hypotézy zvolíme pomocí argumentu `alternative = 'less'` jako levostranný.

```
43 var.test(skull.Hm, skull.Hf, conf.level = 0.99, alternative = 'less')
```

```
      F test to compare two variances
data:  skull.Hm and skull.Hf
F = 1.0799, num df = 214, denom df = 106, p-value = 0.6685
alternative hypothesis: true ratio of variances is less than 1
99 percent confidence interval:
 0.000000 1.580611
sample estimates:
ratio of variances
      1.079861
```

44
45
46
47
48
49
50
51
52
53
54

Součástí výstupu je hodnota testovací statistiky $F = 1.0799$, počty stupňů volnosti Fisherova rozdělení $\text{num df} = 214$ a $\text{denom df} = 106$, hranice 99% Waldova empirického pravostranného intervalu spolehlivosti 0 a 1.5806 a p -hodnota $p\text{-value} = 0.6685$. Jediné, co musíme stanovit zvlášť, je horní hranice kritického oboru. ★

Příklad 10.4. Test o podílu rozptylů

Mějme datový soubor 18-more-samples-variances-clavicle.txt a proměnnou cla.L popisující největší délku klíční kosti z pravé strany v mm (viz sekce ??). Předpokládejme, že náhodná veličina X , popisující největší délku klíční kosti z pravé strany u anglické populace, pochází z normálního rozdělení $N(\mu_1, \sigma_1^2)$, a že náhodná veličina Y , popisující největší délku klíční kosti z pravé strany u indické populace z Amritsaru, pochází z normálního rozdělení $N(\mu_2, \sigma_2^2)$. Na hladině významnosti $\alpha = 0.10$ zjistíte, zda je rozptyl největší délky klíční kosti z pravé strany u anglické populace statisticky významně větší než u indické populace z Amritsaru.

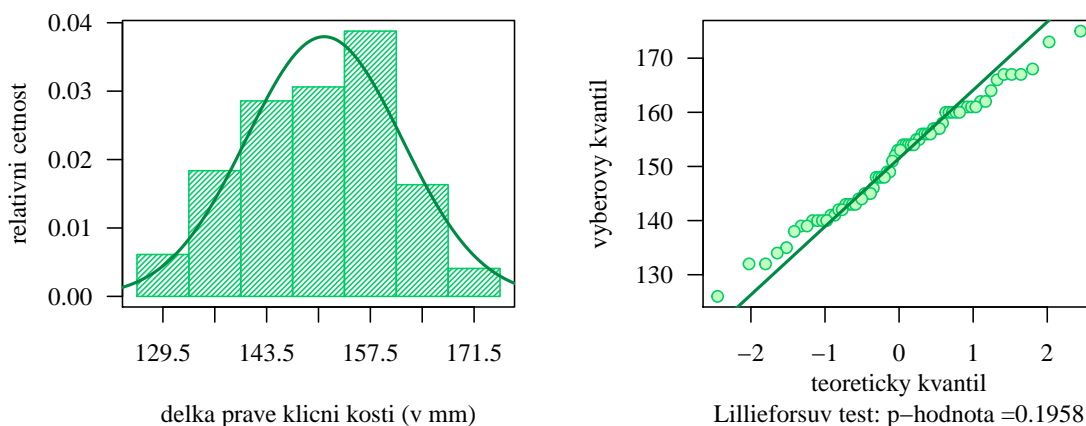
Řešení příkladu 10.4

Pomocí příkazu read.delim() načteme datový soubor a příkazem na.omit() odtráhneme ze souboru NA hodnoty. Pomocí operátoru [] vybereme z datové tabulky údaje o největší délce klíční kosti z pravé strany (cla.L) u jedinců anglické populace (pop == 'eng'), resp. indické populace z Amritsaru (pop == 'ind1'). Nakonec zjistíme rozsahy obou náhodných výběrů.

```
55 data <- read.delim('00-Data//18-more-samples-variances-clavicle.txt')
56 data <- na.omit(data)
57 #head(data, n = 3)
58 cla.Le <- data[data$pop == 'eng', 'cla.L']
59 cla.Li <- data[data$pop == 'ind1', 'cla.L']
60
61 n1 <- length(cla.Le) # 70
62 n2 <- length(cla.Li) # 120
```

Datový soubor obsahuje celkem 70 údajů o největší délce klíční kosti z pravé strany u jedinců anglické populace a 120 údajů o největší délce klíční kosti z pravé strany u jedinců indické populace z Amritsaru. Řešení příkladu vede na test o podílu rozptylů. Před testováním nulové hypotézy ze zadání musíme ověřit normalitu obou náhodných výběrů.

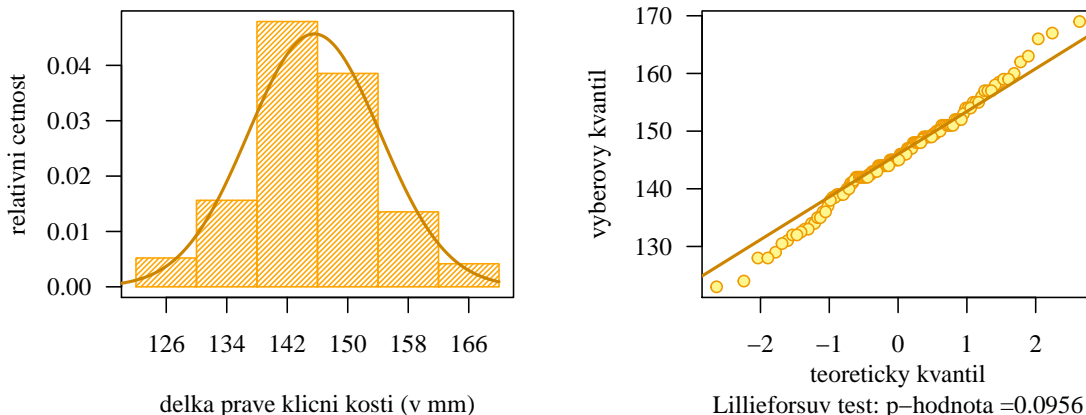
Hladinu významnosti α pro test normality stanovíme standartně, t. $\alpha = 0.05$. Nejprve testujeme hypotézu H_{01} : Náhodný výběr největších délek klíčních kostí z pravé strany u jedinců anglické populace pochází z normálního rozdělení. oproti alternativní hypotéze H_{11} : Náhodný výběr největších délek klíčních kostí z pravé strany u jedinců anglické populace nepochází z normálního rozdělení. Protože rozsah náhodného výběru je větší než 30, ověříme předpoklad normality Lillieforsovým testem. Grafické ověření provedeme na základě kvantilového diagramu a histogramu superponovaného křivkou normálního rozdělení (viz obrázek ??). Datový soubor rozdělíme do devíti ekvidistantních intervalů s šířkou 3 mm stanovením hranic 126, 133, ..., 175.



Obrázek 10: Histogram a kvantilový diagram délky pravé klíční kosti u mužů anglické populace

Jelikož p -hodnota = 0.1958 je větší než 0.05, nulovou hypotézu nezamítáme na hladině významnosti $\alpha = 0.05$. Náhodný největších délek klíčních kostí z pravé strany u jedinců anglické populace pochází z normálního rozdělení.

Analogicky testujeme na hladině významnosti $\alpha = 0.05$ nulovou hypotézu H_{02} : *Náhodný výběr největších délek klíčních kostí z pravé strany u jedinců indické populace z Amritsaru pochází z normálního rozdělení.* oproti alternativní hypotéze H_{12} : *Náhodný výběr největších délek klíčních kostí z pravé strany u jedinců indické populace z Amritsaru nepochází z normálního rozdělení.* Protože rozsah náhodného výběru $n = 120$ je větší než 30, ověříme předpoklad normality opět Lillieforsovým testem. V rámci histogramu rozdělíme data do osmi ekvidistančních intervalů s šířkou 8 mm prostřednictvím stanovených hranic 122, 130, ..., 170 (viz obrázek E10-Test-ss-cla.LAI).



Obrázek 11: Histogram a kvantilový diagram délky pravé klíční kosti u mužů indické populace

Protože p -hodnota = 0.0956 je větší než 0.05, nulovou hypotézu o normalitě dat nezamítáme na hladině významnosti $\alpha = 0.05$. Náhodný výběr největších délek klíčních kostí z pravé strany u jedinců indické populace z Amritsaru pochází z normálního rozdělení.

Jelikož oba náhodné výběry pochází z normálních rozdělení, můžeme k testování hypotézy ze zadání použít parametrický test o podílu rozptylů. Zde je vhodné upozornit, že v zadání příkladu se nepíše nic o znění nulové hypotézy. Ze zadání víme, že se snažíme prokázat, že rozptyl u anglické populace je větší než rozptyl u indické populace z Amritsaru. Jak bylo zmíněno v úvodní části kapitoly ??, tvrzení, jehož platnost se snažíme dokázat, je vždy součástí alternativní hypotézy. Ze zadání tedy známe tvar alternativní hypotézy, zatímco znění nulové hypotézy vhodně doplníme.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : *Rozptyl největší délky klíční kosti z pravé strany u anglické populace je menší nebo roven rozptylu u indické populace z Amritsaru.*

H_1 : *Rozptyl největší délky klíční kosti z pravé strany u anglické populace je větší než rozptyl u indické populace z Amritsaru.*

- **matematická formulace** nulové a alternativní hypotézy

$H_0 : \sigma_1^2 \leq \sigma_2^2 \rightarrow \sigma_1^2 / \sigma_2^2 \leq \sigma_0^2$, kde $\sigma_0^2 = 1$

$H_1 : \sigma_1^2 > \sigma_2^2 \rightarrow \sigma_1^2 / \sigma_2^2 > \sigma_0^2$, kde $\sigma_0^2 = 1$ (pravostranná alternativa)

2. Volba hladiny významnosti

- Hladinu významnosti volíme jako $\alpha = 0.10$ (viz zadání příkladu).

3. Testování kritickým oborem

- **Testovací statistika**

$$F_W = \frac{S_1^2}{S_2^2} = \frac{10.51007^2}{8.733432^2} = \frac{110.4616}{76.27283} = 1.448243 \doteq 1.4482$$

```
63 alpha <- 0.10
64 s1 <- sd(c1a.Le)
65 s2 <- sd(c1a.Li)
66 Fw <- s1^2 / s2^2 # 1.448242
```

- Kritický obor

$$\begin{aligned} W &= (F_{n_1-1, n_2-1}(1-\alpha); \infty) \\ &= (F_{69, 119}(0.90); \infty) \\ &= (1.307459; \infty) \doteq (1.3075; \infty) \end{aligned}$$

```
67 qf(1 - alpha, n1 - 1, n2 - 1) # 1.307459
```

- Závěr testování

Protože realizace testovací statistiky $f_W = 1.4482$ náleží do kritického oboru, tj. $f_W \in W$, H_0 zamítáme na hladině významnosti $\alpha = 0.10$.

4. Testování intervalem spolehlivosti

- Interval spolehlivosti

$$\begin{aligned} (d, \infty) &= \left(\frac{s_1^2/s_2^2}{F_{n_1-1, n_2-1}(\alpha); \infty} \right) \\ &= \left(\frac{10.51007^2/8.733432^2}{F_{69, 119}(0.90); \infty} \right) \\ &= \left(\frac{1.448243}{1.307459}; \infty \right) \\ &= (1.107678; \infty) \doteq (1.1077; \infty) \end{aligned}$$

```
68 (DH <- (s1^2 / s2^2) / qf(1 - alpha, n1 - 1, n2 - 1)) # 1.107677
```

```
[1] 1.107677
```

69

- Závěr testování

Protože $\sigma_0^2 = 1$ nenáleží do Waldova 90% empirického jednostranného intervalu spolehlivosti, tj. $\sigma_0^2 = 1 \in IS$, H_0 zamítáme na hladině významnosti $\alpha = 0.10$.

5. Testování p -hodnotou

- p -hodnota

$$\begin{aligned} p\text{-hodnota} &= \Pr(F_W > f_W) = 1 - \Pr(F_W \leq f_W) \\ &= 1 - \Pr(F_W \leq 1.307459) \\ &= 0.03845745 \doteq 0.0385 \end{aligned}$$

- Závěr testování

Protože p -hodnota = 0.0385 je menší než $\alpha = 0.10$, H_0 zamítáme na hladině významnosti $\alpha = 0.1$.

6. **Interpretace výsledků:** Za základě všech tří typů testování zamítáme nulovou hypotézu na hladině významnosti $\alpha = 0.10$. Rozptyl největší délky klíční kosti z pravé strany u anglické populace je statisticky významně větší než rozptyl u indické populace z Amritsaru.

```
70 p.val <- 1 - pf(Fw, n1 - 1, n2 - 1) # 0.03845745
```

Poznámka: K otestování nulové hypotézy o podílu rozptylů můžeme využít funkci `var.test()`, kde hodnotu argumentu `conf.level` nastavíme na hodnotu 0.90 a typ alternativní hypotézy zvolíme pomocí argumentu `alternative = 'greater'` jako pravostranný.

```
71 var.test(cla.Le, cla.Li, conf.level = 0.90, alternative = 'greater')
```

```
      F test to compare two variances
data:  cla.Le and cla.Li
F = 1.4482, num df = 69, denom df = 119, p-value = 0.03846
alternative hypothesis: true ratio of variances is greater than 1
90 percent confidence interval:
 1.107677      Inf
sample estimates:
ratio of variances
      1.448242
```

72
73
74
75
76
77
78
79
80
81
82

Součástí výstupu je hodnota testovací statistiky $F = 1.4482$, počty stupňů volnosti Fisherova rozdělení $\text{num df} = 69$ a $\text{denom df} = 119$, hranice 90% Waldova empirického levostranného intervalu spolehlivosti 1.1077 a Inf a p -hodnota $p\text{-value} = 0.0385$. Jediné, co musíme stanovit zvlášť, je dolní hranice kritického oboru. ★

Příklad 10.5. Test o podílu rozptylů

Mějme datový soubor 19-more-samples-correlations-skull.txt a proměnnou intorb.B popisující interorbitální šířku v mm (viz sekce ??). Na hladině významnosti $\alpha = 0.01$ zjistěte, zda je rozptyl interorbitální šířky mužů malajské populace statisticky významně větší než rozptyl interorbitální šířky mužů čínské populace.

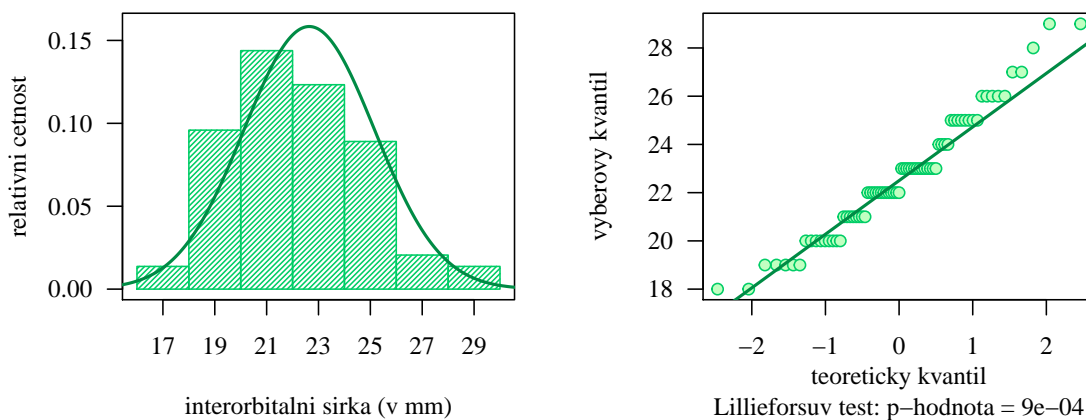
Řešení příkladu 10.5

Pomocí příkazu read.delim() načteme datový soubor a odstraníme ze souboru chybějící hodnoty. Pomocí operátoru [] vybereme z datové tabulky údaje o interorbitální šířce (intorb.B) u mužů malajské populace (pop == 'mal'), resp. u mužů čínské populace (pop == 'cin'). Nakonec zjistíme rozsahy obou náhodných výběrů.

```
83 data <- read.delim('00-Data//19-more-samples-correlations-skull.txt')
84 data <- na.omit(data)
85 intorb.BM <- data[data$pop == 'mal', 'intorb.B']
86 intorb.BC <- data[data$pop == 'cin', 'intorb.B']
87
88 n1 <- length(intorb.BM) # 73
89 n2 <- length(intorb.BC) # 19
```

Datový soubor obsahuje celkem údaje o interorbitální šířce 73 mužů malajské populace a 19 mužů čínské populace. Řešení příkladu vede na test o podílu rozptylů. Před testováním hypotézy ze zadání musíme ověřit normalitu obou náhodných výběrů. Hladinu významnosti zvolíme pro tento účel $\alpha = 0.05$.

Protože rozsah náhodného výběru interorbitálních šířek mužů malajské populace je větší než 30, použijeme na ověření předpokladu normality Lillieforsův test. Rozdělení datového souboru vizualizujeme pomocí kvantilového diagramu a histogramu (viz obrázek 12). Datový soubor rozdělíme do sedmi ekvidistatních intervalů s šířkou 2 mm stanovením hranic 16, 18, ..., 30 mm.

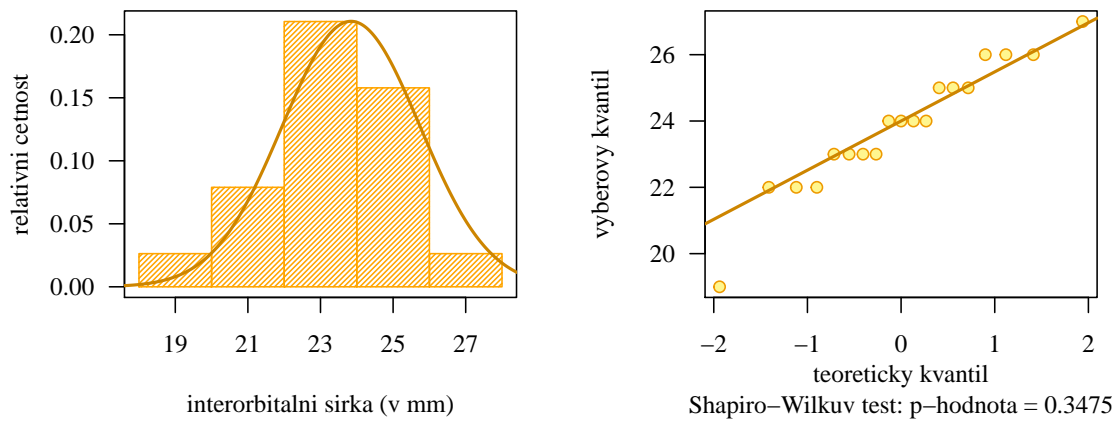


Obrázek 12: Histogram a kvantilový diagram interorbitální šířky mužů malajské populace

Jelikož p -hodnota = 9×10^{-4} je menší než 0.05, nulovou hypotézu zamítáme na hladině významnosti $\alpha = 0.05$. Také histogram ukazuje porušení předpokladu normality zejména vyšikmením naměřených hodnot doleva oproti předpokládanému rozdělení. V kvantilovém diagramu vidíme odklon hodnot od referenční přímky v pravém horním rohu. Náhodný výběr interorbitálních šířek mužů malajské populace nepochází z normálního rozdělení.

Dále otestujeme hypotézu o normalitě náhodného výběru interorbitálních šířek mužů čínské populace. Protože rozsah náhodného výběru $n = 19$ je menší než 30, ověříme předpoklad normality Shapiro-Wilkovým testem. V rámci histogramu rozdělíme data do pěti ekvidistatních intervalů s šířkou 2 mm prostřednictvím stanovených hranic 18, 20, ..., 28 (viz obrázek 13).

Protože p -hodnota = 0.3475 je větší než 0.05, nulovou hypotézu o normalitě dat nezamítáme na hladině významnosti $\alpha = 0.05$. Z obrázku 13 vidíme, že histogram věrně kopíruje tvar křivky normálního rozdělení. Kvantilový diagram ukazuje těsnou příchýlnost bodů k referenční přímce vyjma krajního levého bodu. Odchýlení tohoto bodů však



Obrázek 13: Histogram a kvantilový diagram interorbitální šířky mužů čínské populace

podle Shapiro-Wilkova testu nenarušuje fatálně předpoklad normality náhodného výběru. Náhodný výběr interorbitálních šířek u mužů čínské populace pochází z normálního rozdělení.

Jelikož náhodný výběr interorbitálních šířek mužů malajské populace nesplňuje předpoklad normality, nemůžeme provést parametrický test o podílu rozptylů. ★

10.2 Klasický dvouvýběrový t -test o rozdílu středních hodnot $\mu_1 - \mu_2$

Nechť X_{11}, \dots, X_{1n_1} je náhodný výběr z $N(\mu_1, \sigma_1^2)$, a X_{21}, \dots, X_{2n_2} je na něm nezávislý náhodný výběr z rozdělení $N(\mu_2, \sigma_2^2)$, přičemž $n_1 \geq 2$, $n_2 \geq 2$ a σ_1^2 a σ_2^2 jsou neznámé ale shodné rozptyly, tj. $\sigma_1^2 = \sigma_2^2$. Nechť μ_0 je konstanta. Na hladině významnosti α testujeme jednu z následujících tří hypotéz oproti příslušné alternativní hypotéze.

$$\begin{array}{lll} H_{01} : \mu_1 - \mu_2 = \mu_0 & \text{oproti} & H_{11} : \mu_1 - \mu_2 \neq \mu_0 \quad (\text{oboustranná alt.}) \\ H_{02} : \mu_1 - \mu_2 \leq \mu_0 & \text{oproti} & H_{12} : \mu_1 - \mu_2 > \mu_0 \quad (\text{pravostranná alt.}) \\ H_{03} : \mu_1 - \mu_2 \geq \mu_0 & \text{oproti} & H_{13} : \mu_1 - \mu_2 < \mu_0 \quad (\text{levostranná alt.}) \end{array}$$

Test nazýváme klasický dvouvýběrový t -test o rozdílu středních hodnot $\mu_1 - \mu_2$. Testovací statistika má tvar

$$T_W = \frac{(M_1 - M_2) - \mu_0}{S_* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}. \quad (10.2)$$

kde M_1 je výběrový průměr a n_1 je rozsah prvního náhodného výběru, M_2 je výběrový průměr a n_2 je rozsah druhého náhodného výběru, dále S_* je aritmetický průměr výběrových rozptylů S_1^2 a S_2^2 (viz kapitola ??), kde S_1^2 je výběrový rozptyl prvního náhodného výběru a S_2^2 je výběrový rozptyl druhého náhodného výběru a konečně μ_0 je konstanta z nulové hypotézy. Za platnosti nulové hypotézy pochází statistika T_W ze Studentova rozdělení o $n_1 + n_2 - 2$ stupních volnosti, tj.

$$T_W = \frac{(M_1 - M_2) - \mu_0}{S_* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \stackrel{H_0}{\sim} t_{n_1+n_2-2}.$$

Kritický obor podle zvolené alternativní hypotézy má tvar

$$\begin{array}{ll} H_{11} : \mu_1 - \mu_2 \neq \mu_0 & W = (-\infty; t_{n_1+n_2-2}(\alpha/2)) \cup \langle t_{n_1+n_2-2}(1-\alpha/2); \infty \rangle \\ H_{12} : \mu_1 - \mu_2 > \mu_0 & W = \langle t_{n_1+n_2-2}(1-\alpha); \infty \rangle \\ H_{13} : \mu_1 - \mu_2 < \mu_0 & W = (-\infty; t_{n_1+n_2-2}(\alpha)) \end{array}$$

kde $t_{n_1+n_2-2}(\alpha/2)$, $t_{n_1+n_2-2}(1-\alpha/2)$, $t_{n_1+n_2-2}(\alpha)$ a $t_{n_1+n_2-2}(1-\alpha)$ jsou kvantily Studentova rozdělení o $n_1 + n_2 - 2$ stupních volnosti, jejichž hodnoty získáme pomocí \mathbb{R} a implementované funkce $\text{qt}()$.

Interval spolehlivosti má podle zvolené alternativní hypotézy jeden z následujících tvarů

$$\begin{array}{ll} H_{11} : \mu_1 - \mu_2 \neq \mu_0 & (d, h) = \left(m_1 - m_2 - s_* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} t_{n_1+n_2-2}(1-\alpha/2); m_1 - m_2 - s_* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} t_{n_1+n_2-2}(\alpha/2) \right) \\ H_{12} : \mu_1 - \mu_2 > \mu_0 & (d, \infty) = \left(m_1 - m_2 - s_* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} t_{n_1+n_2-2}(1-\alpha); \infty \right) \\ H_{13} : \mu_1 - \mu_2 < \mu_0 & (-\infty, h) = \left(-\infty; m_1 - m_2 - s_* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} t_{n_1+n_2-2}(\alpha) \right) \end{array}$$

p -hodnota má v závislosti na zvolené alternativní hypotéze jeden z následujících tvarů

$$\begin{array}{ll} H_{11} : \mu \neq \mu_0 & p\text{-hodnota} = 2 \min\{\Pr(T_W \leq t_W), \Pr(T_W > t_W)\} \\ H_{12} : \mu > \mu_0 & p\text{-hodnota} = \Pr(T_W > t_W) = 1 - \Pr(T_W \leq t_W) \\ H_{13} : \mu < \mu_0 & p\text{-hodnota} = \Pr(T_W \leq t_W) \end{array}$$

kde T_W je náhodná veličina, t_W je realizace testovací statistiky T_W (viz vzorec 10.2), tedy konkrétní číslo, a $\Pr(T_W \leq t_W)$ je distribuční funkce Studentova rozdělení o $n_1 + n_2 - 2$ stupních volnosti, jejíž hodnotu získáme pomocí \mathbb{R} a implementované funkce $\text{pt}()$.

Příklad 10.6. Test o rozdílu středních hodnot

Mějme datový soubor 01-one-sample-mean-skull-mf.txt a proměnnou skull.B popisující největší šířku mozkovny v mm (viz sekce ??). Předpokládejme, že náhodná veličina X , popisující největší šířku mozkovny u mužů, pochází z normálního rozdělení $N(\mu_1, \sigma_1^2)$, a že náhodná veličina Y , popisující největší šířku mozkovny u žen, pochází z normálního rozdělení $N(\mu_2, \sigma_2^2)$. Na hladině významnosti $\alpha = 0.05$ otestujte hypotézu o shodě středních hodnot největší šířky mozkovny u mužů a u žen.

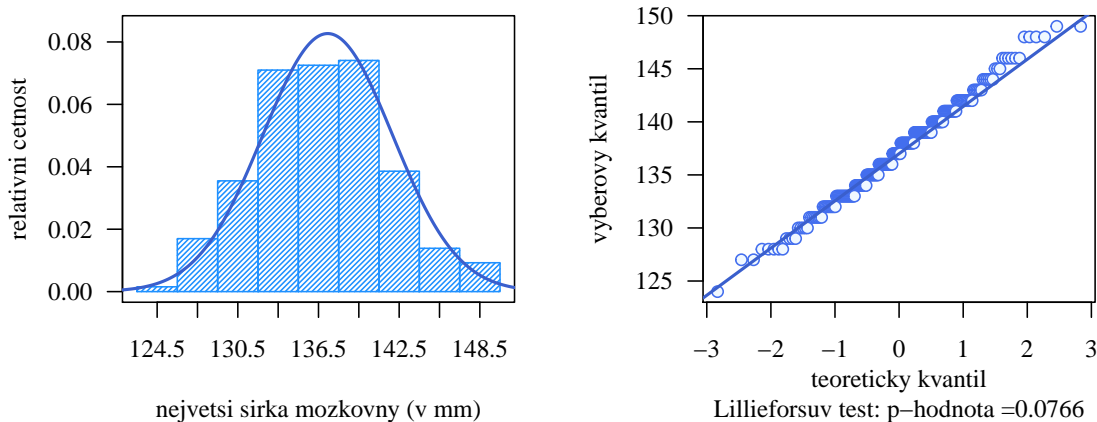
Řešení příkladu 10.6

Příkazem `read.delim()` načteme datový soubor a příkazem `na.omit()` z něj odtráhneme NA hodnoty. Pomocí operátoru `[]` vybereme z datové tabulky údaje o největší šířce mozkovny (skull.B) u mužů (`sex == 'm'`), resp. u žen (`sex == 'f'`) a zjistíme rozsahy obou náhodných výběrů.

```
90 data <- read.delim('00-Data//01-one-sample-mean-skull-mf.txt')
91 data <- na.omit(data)
92 #head(data, n = 3)
93 skull.Bm <- data[data$sex == 'm', 'skull.B']
94 skull.Bf <- data[data$sex == 'f', 'skull.B']
95
96 n1 <- length(skull.Bm) # 216
97 n2 <- length(skull.Bf) # 109
```

Datový soubor obsahuje celkem 216 údajů o největší šířce mozkovny u mužů a 107 údajů o největší šířce mozkovny u žen. Řešení příkladu vede na test o rozdílu středních hodnot. Před testováním nulové hypotézy ze zadání musíme ověřit splnění předpokladu normality obou náhodných výběrů, abychom zjistili, zda použít parametrický nebo neparametrický test. V případě splnění předpokladu normality musíme provést test o shodě rozptylů, abychom určili, zda máme použít klasický dvouvýběrový parametrický t -test nebo Welchův dvouvýběrový t -test. Jelikož není stanoveno jinak, zvolíme pro testy normality i případný test o podílu rozptylů hladinu významnosti $\alpha = 0.05$.

Nejprve testujeme hypotézu H_{01} : *Náhodný výběr největších šířek mozkovny u mužů pochází z normálního rozdělení.* oproti alternativní hypotéze H_{11} : *Náhodný výběr největší šířky mozkovny u mužů nepochází z normálního rozdělení.* Jelikož náhodný výběr sestává z 215 pozorování, což je více než 30, provedeme test normality Lillieforsovým testem. Normalitu ověříme též graficky vykreslením kvantilového diagramu (viz sekce ??) a pomocí histogramu superponovaného křivkou normálního rozdělení (viz sekce ??). Datový soubor rozdělíme na základě Sturjerova pravidla (viz sekce ??) do devíti ekvidistatních intervalů s šířkou 3 mm prostřednictvím stanovených hranic 123, 126, ..., 150 (viz obrázek 14).

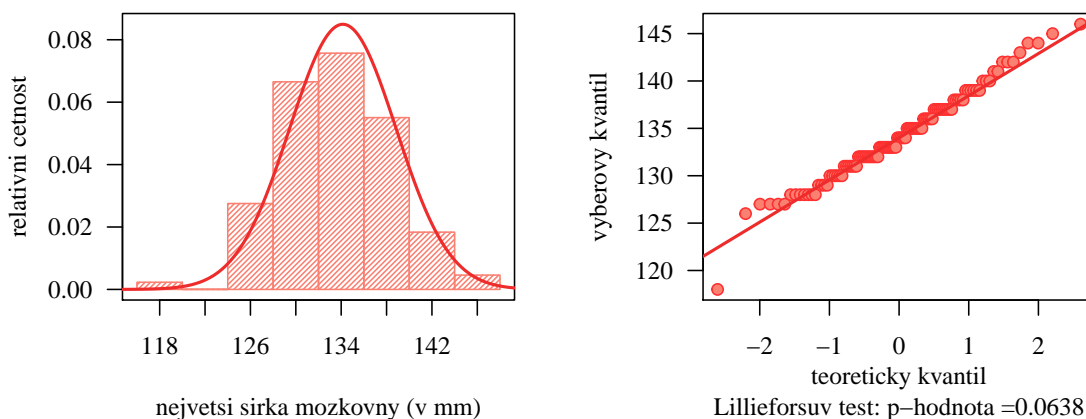


Obrázek 14: Histogram a kvantilový diagram největší šířky mozkovny mužů

Protože p -hodnota = 0.0766 je větší než 0.05, nulovou hypotézu nezamítáme na hladině významnosti $\alpha = 0.05$.

Náhodný výběr největších šířek mozkovny mužů pochází z normálního rozdělení. Vykreslený histogram i kvantilový diagram výsledek testu normality podporují. Histogram kopíruje tvar křivky normálního rozdělení a ve vykresleném kvantilovém diagramu se body pohybují okolo referenční čáry.

Analogicky testujeme hypotézu H_{02} : *Náhodný výběr největších šířek mozkovky žen pochází z normálního rozdělení.* oproti alternativní hypotéze H_{12} : *Náhodný výběr největších šířek mozkovny žen nepochází z normálního rozdělení.* Jelikož rozsah náhodného výběru je dostatečně vysoký, použijeme na otestování hypotézy o normalitě Lillieforsův test. V rámci histogramu rozdělíme soubor do osmi ekvidistatních intervalů s šířkou 4 mm prostřednictvím stanovených hranic 116, 120, ..., 148 (viz obrázek 15).



Obrázek 15: Histogram a kvantilový diagram největší šířky mozkovny žen

Protože p -hodnota = 0.0638 je větší než 0.05, nulovou hypotézu nezamítáme na hladině významnosti $\alpha = 0.05$. Náhodný výběr největší šířky mozkovny žen také pochází z normálního rozdělení. Vykreslený histogram i kvantilový diagram podporují výsledek testování.

Jelikož oba náhodné výběry splňují předpoklad normality, můžeme k testování hypotézy ze zadání použít parametrický test. Zda je vhodné zvolit klasický dvouvýběrový t -test nebo Welchův dvouvýběrový t -test rozhodneme na základě výsledku testu o shodě rozptylů obou náhodných výběrů.

Na hladině významnosti $\alpha = 0.05$ testujeme hypotézu $H_0 : \sigma_1^2/\sigma_2^2 = 1$ (rozptyly σ_1^2 a σ_2^2 jsou shodné) oproti alternativní hypotéze $H_1 : \sigma_1^2/\sigma_2^2 \neq 1$ (rozptyly σ_1^2 a σ_2^2 nejsou shodné). Nulovou hypotézu otestujeme pomocí p -hodnoty získané pomocí funkce `var.test()`. Kompletní postup testování by byl analogický postupu uvedenému v příkladu 10.2.

```
98 var.test(skull.Bm, skull.Bf, alternative = 'two.sided')$p.val
```

```
[1] 0.761025
```

99

Jelikož p -hodnota = 0.761025 je větší než 0.05, nulovou hypotézu o shodě rozptylů nezamítáme na hladině významnosti $\alpha = 0.05$. Na základě výsledků testování tedy předpokládáme, že oba rozptyly jsou shodné.

Protože oba výběry pochází z normálních rozdělení, jejichž rozptyly jsou shodné, použijeme na otestování hypotézy ze zadání klasický dvouvýběrový t -test o rozdílu středních hodnot $\mu_1 - \mu_2$. Řešení si nyní uvedeme v posloupnosti sedmi kroků.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : Střední hodnota největší šířky mozkovny mužů a žen jsou shodné.

H_1 : Střední hodnota největší šířky mozkovny mužů a žen nejsou shodné.

- **matematická formulace** nulové a alternativní hypotézy

$$H_0 : \mu_1 = \mu_2 \quad \rightarrow \quad \mu_1 - \mu_2 = \mu_0, \text{ kde } \mu_0 = 0$$

$$H_1 : \mu_1 \neq \mu_2 \quad \rightarrow \quad \mu_1 - \mu_2 \neq \mu_0, \text{ kde } \mu_0 = 0$$

(oboustranná alternativa)

2. Volba hladiny významnosti

- Hladinu významnosti volíme ze zadání $\alpha = 0.05$.

3. Testování kritickým oborem

- **Aritmetický průměr výběrových rozptylů**

$$\begin{aligned} S_* &= \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}} \\ &= \sqrt{\frac{(216 - 1) \times 4.824642^2 + (109 - 1) \times 4.695991^2}{216 + 109 - 2}} \\ &= \sqrt{\frac{215 \times 23.27717 + 108 \times 22.05233}{323}} \\ &= \sqrt{\frac{7386.243}{323}} &= \sqrt{22.86763} = 4.7820109 \doteq 4.7820 \end{aligned}$$

```
100 s1 <- sd(skull.Bm)
101 s2 <- sd(skull.Bf)
102 sh <- sqrt(((n1 - 1) * s1^2 + (n2 - 1) * s2^2) / (n1 + n2 - 2)) # 4.7820109
```

- **Testovací statistika**

$$\begin{aligned} T_W &= \frac{(M_1 - M_2) - \mu_0}{S_* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \\ &= \frac{(137.1852 - 134.1468) - 0}{4.7820109 \sqrt{\frac{1}{215} + \frac{1}{109}}} \\ &= \frac{3.038396}{4.7820109 \times 0.1174902} \\ &= \frac{3.038396}{0.5618394} = 5.40794 \doteq 5.4079 \end{aligned}$$

```
103 alpha <- 0.05
104 mu0 <- 0
105 m1 <- mean(skull.Bm)
106 m2 <- mean(skull.Bf)
107 tw <- ((m1 - m2) - mu0) / (sh * sqrt(1 / n1 + 1 / n2)) # 5.40794
```

Kritický obor

$$\begin{aligned} W &= (-\infty; t_{n_1+n_2-2}(\alpha/2)) \cup (t_{n_1+n_2-2}(1-\alpha/2); \infty) \\ &= (-\infty; t_{216+109-2}(0.05/2)) \cup (t_{216+109-2}(1-0.05/2); \infty) \\ &= (-\infty; t_{323}(0.025)) \cup (t_{323}(0.975); \infty) \\ &= (-\infty; -1.967336) \cup (1.967336; \infty) \end{aligned}$$

```
108 qt(alpha/2, n1 + n2 - 2) # -1.967336
109 qt(1-alpha/2, n1 + n2 - 2) # 1.967336
```

- **Závěr testování**

Protože realizace testovací statistiky $t_W = 5.4079$ náleží do kritického oboru, tj. $t_W \in W$, H_0 zamítáme na hladině významnosti $\alpha = 0.05$.

4. Testování intervalem spolehlivosti

- **Interval spolehlivosti**

$$\begin{aligned}
 (d, h) &= \left(m_1 - m_2 - s_* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} t_{n_1+n_2-2}(1 - \alpha/2); m_1 - m_2 - s_* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} t_{n_1+n_2-2}(\alpha/2) \right) \\
 &= ((137.1852 - 134.1468) - 4.7820109 \sqrt{\frac{1}{215} + \frac{1}{109}} t_{323}(0.975); (137.1852 - 134.1468) - 4.7820109 \sqrt{\frac{1}{215} + \frac{1}{109}} t_{323}(0.025)) \\
 &= (3.038396 - 4.7820109 \times 0.1174902 \times 1.967336; 3.038396 - 4.7820109 \times 0.1174902 \times (-1.967336)) \\
 &= (1.933069; 4.143723) \doteq (1.9331; 4.1437)
 \end{aligned}$$

```
110 dh <- (m1 - m2) - sh * sqrt(1 / n1 + 1 / n2) * qt(1 - alpha / 2, n1 + n2 - 2) # 1.93307
111 hh <- (m1 - m2) + sh * sqrt(1 / n1 + 1 / n2) * qt(alpha / 2, n1 + n2 - 2) # 4.143723
```

- **Závěr testování**

Protože $\mu_0 = 0$ nenáleží do Waldova 95% empirického oboustranného intervalu spolehlivosti, tj. $\mu_0 = 0 \notin IS$, H_0 zamítáme na hladině významnosti $\alpha = 0.05$.

5. Testování p -hodnotou

- **p -hodnota**

$$\begin{aligned}
 p\text{-hodnota} &= 2 \min\{\Pr(T_W \leq t_W), \Pr(T_W > t_W)\} \\
 &= 2 \min\{\Pr(T_W \leq 5.40794), 1 - \Pr(T_W \leq 5.40794)\} \\
 &= 2 \min\{0.9999999, 6.212613 \times 10^{-8}\} \\
 &= 2 \times (6.212613 \times 10^{-8}) \\
 &= 1.242523 \times 10^{-7} \doteq 1.2425 \times 10^{-7}
 \end{aligned}$$

```
112 p.val <- 2 * min(pt(tw, n1 + n2 - 2), 1 - pt(tw, n1 + n2 - 2)) # 1.242523e-07
```

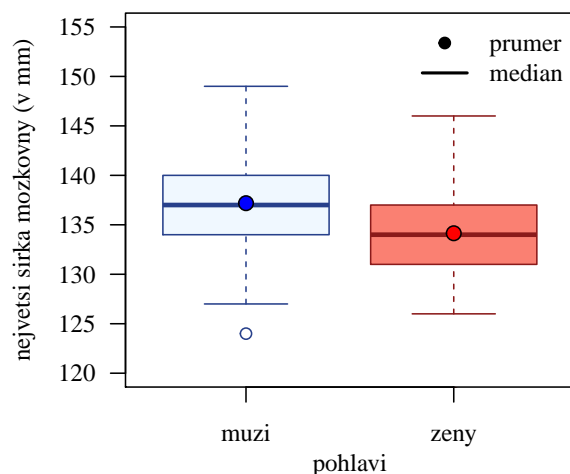
- **Závěr testování**

Protože p -hodnota = 1.2425×10^{-7} je menší než $\alpha = 0.05$, H_0 zamítáme na hladině významnosti $\alpha = 0.05$.

6. Grafická vizualizace výsledků testování

Vhodný grafem porovnávajícím střední hodnoty obou náhodných výběrů je krabicový diagram (obrázek 16).

7. **Interpretace výsledků:** Na základě všech tří způsobů testování zamítáme hypotézu o shodě středních hodnot největší šířky mozkovny mužů a žen. Mezi největší šířkou mozkovny mužů a žen existuje statisticky významný rozdíl.



Obrázek 16: Krabicový diagram největší šířky mozkovny mužů a žen

```
113 t.test(skull.Bm, skull.Bf, conf.level = 0.95, alternative = 'two.sided', var.equal = T)
```

Poznámka: K otestování nulové hypotézy o rozdílu středních hodnot můžeme využít funkci `t.test()`. Vstupními parametry budou nejprve dva vektory reprezentující náhodné výběry, tj. `skull.Bm` a `skull.Bf`, dále hodnota hladiny významnosti α zadaná prostřednictvím koeficientu spolehlivosti $1 - \alpha$ nastavením hodnoty argumentu `conf.level = 0.95`, typ zvolené alternativní hypotézy (oboustranná) zadaný pomocí argumentu `alternative = 'two.sided'` a nakonec argument `var.equal == T` volbu klasického t -testu založeného na předpokladu shody rozptylů σ_1^2 a σ_2^2 .

```
Two Sample t-test
data: skull.Bm and skull.Bf
t = 5.4079, df = 323, p-value = 1.243e-07
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 1.933070 4.143723
sample estimates:
mean of x mean of y
 137.1852  134.1468
```

114
115
116
117
118
119
120
121
122
123
124

Součástí výstupu je hodnota testovací statistiky $t = 5.4079$, počet stupňů volnosti Studentova rozdělení $\text{num df} = 323$, hranice intervalu spolehlivosti `1.933070` a `4.143723` a p -hodnota `p-value = 1.243e-07`. Jediné, co musíme stanovit zvlášť, jsou dolní a horní hranice kritického oboru.

★

Příklad 10.7. Test o rozdílu středních hodnot

Mějme datový soubor 15-anova-means-skull.txt a proměnnou upface.H popisující výšku horní části tváře v mm (viz sekce ??). Předpokládejme, že náhodná veličina X , popisující výšku horní části tváře čínské populace, pochází z normálního rozdělení $N(\mu_1, \sigma_1^2)$, a že náhodná veličina Y , popisující výšku horní části tváře malajské populace, pochází z normálního rozdělení $N(\mu_2, \sigma_2^2)$. Na hladině významnosti $\alpha = 0.05$ proveďte, zda je střední hodnota výšky horní části tváře čínské populace statisticky významně větší než střední hodnota výšky horní části tváře malajské populace.

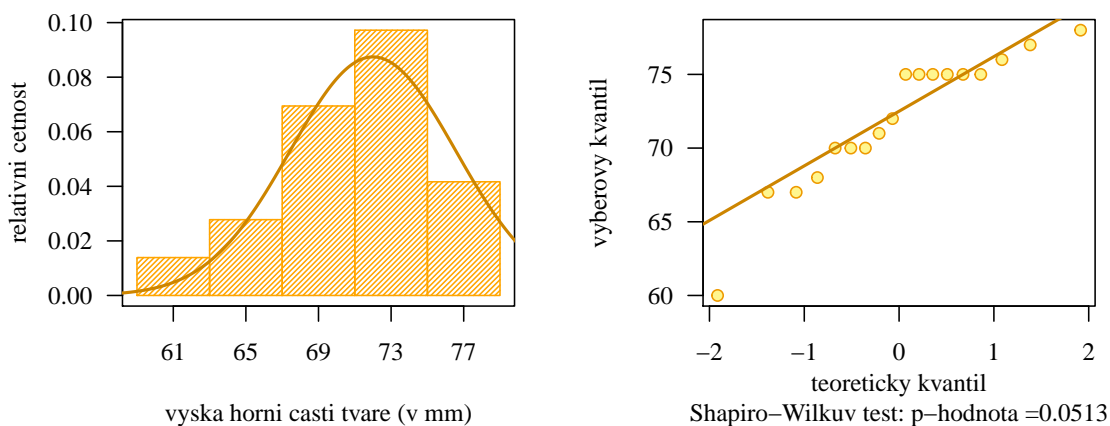
Řešení příkladu 10.7

Příkazem read.delim() načteme datový soubor a příkazem na.omit() z něj odtráhneme NA hodnoty. Pomocí operátoru [] vybereme z datové tabulky naměřené hodnoty výšky horní části tváře (upface.H) čínské populace (pop == 'cin'), resp. malajské populace (pop == 'mal') a zjistíme rozsahy obou náhodných výběrů.

```
125 data <- read.delim('00-Data//15-anova-means-skull.txt')
126 data <- na.omit(data)
127 #head(data, n = 3)
128 upface.Hc <- data[data$pop == 'cin', 'upface.H']
129 upface.Hm <- data[data$pop == 'mal', 'upface.H']
130
131 n1 <- length(upface.Hc) # 18
132 n2 <- length(upface.Hm) # 69
```

Datový soubor obsahuje naměřené hodnoty výšky horní části tváře u 18 jedinců čínské populace a 69 jedinců malajské populace. Řešení příkladu vede na test o rozdílu středních hodnot.

Nejprve otestujeme normalitu obou náhodných výběrů. Na hladině významnosti α testujeme hypotézu H_{01} : *Náhodný výběr naměřených výšek horní části tváře čínské populace pochází z normálního rozdělení.* oproti alternativní hypotéze H_{11} : *Náhodný výběr naměřených výšek horní části tváře čínské populace nepochází z normálního rozdělení.* K otestování předpokladu normality použijeme Shapiro-Wilkův test spolu s kvantilovým diagramem a histogramem (viz obrázek 17). Datový soubor rozdělíme do pěti ekvidistatních intervalů s šířkou 4 mm prostřednictvím stanovených hranic 59, 63, ..., 79.

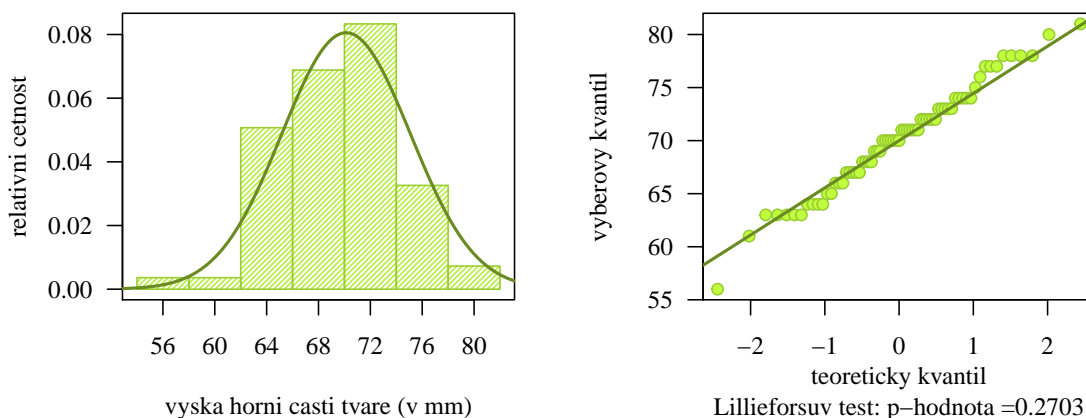


Obrázek 17: Histogram a kvantilový diagram výšky horní části tváře mužů čínské populace

Protože p -hodnota = 0.0513 je větší než 0.05, nulovou hypotézu nezamítáme na hladině významnosti $\alpha = 0.05$. Náhodný výběr výšek horní části tváře čínské populace pochází z normálního rozdělení. Vykreslený histogram vhodně kopíruje tvar křivky normálního rozdělení, naměřené hodnoty jsou pouze mírně vyšikmené doprava. Tentýž jev je viditelný i na kvantilovém diagramu, kde se krajní bod umístěný nalevo dole odchyluje od referenční křivky. Při tak nízkém počtu pozorování však nemusí jít nutně o porušení normálního charakteru dat. Proto se kloníme k

závěru Shapiro-Wilkova testu a data považujeme za normálně rozdělená.

Analogicky testujeme hypotézu H_{02} : *Náhodný výběr naměřených výšek horní části tváře malajské populace pochází z normálního rozdělení.* oproti alternativní hypotéze H_{12} : *Náhodný výběr naměřených výšek horní části tváře malajské populace nepochází z normálního rozdělení.* Vzhledem k rozsahu náhodného výběru použijeme k otestování normality Lillieforsův test. Datový soubor rozdělíme do sedmi ekvidistančních intervalů s šířkou 4 mm prostřednictvím stanovených hranic 54, 58 . . . , 82 (viz obrázek 18).



Obrázek 18: Histogram a kvantilový diagram výšky horní části tváře mužů malajské populace

Protože p -hodnota = 0.2703 je větší než 0.05, nulovou hypotézu nezamítáme na hladině významnosti $\alpha = 0.05$. Náhodný výběr naměřených výšek největší části tváře malajské populace pochází z normálního rozdělení. Vykreslený histogram i kvantilový diagram podporují výsledek testování.

Nyní zbývá rozhodnout o splnění nebo nesplnění předpokladu o shodě rozptylů obou náhodných výběrů. Na hladině významnosti $\alpha = 0.05$ testujeme hypotézu $H_0 : \sigma_1^2/\sigma_2^2 = 1$ (rozptyly σ_1^2 a σ_2^2 jsou shodné) oproti alternativní hypotéze $H_1 : \sigma_1^2/\sigma_2^2 \neq 1$ (rozptyly σ_1^2 a σ_2^2 nejsou shodné). Nulovou hypotézu otestujeme pomocí funkce `var.test()` na základě p -hodnoty (viz příklad 10.2).

```
133 var.test(upface.Hc, upface.Hm, alternative = 'two.sided')$p.val
```

```
[1] 0.7340571
```

134

Jelikož p -hodnota = 0.7341 je větší než 0.05, nulovou hypotézu o shodě rozptylů nezamítáme na hladině významnosti $\alpha = 0.05$. Na základě výsledků testování tedy předpokládáme, že oba rozptyly jsou shodné.

Protože oba náhodné výběry jsou normálně rozdělené, a o jejich rozptylech předpokládáme, že jsou shodné, použijeme na otestování hypotézy ze zadání klasický dvouvýběrový t -test o rozdílu středních hodnot $\mu_1 - \mu_2$. Poznamenejme, že v zadání není žádná zmínka o tvaru nulové hypotézy. Zadaným úkolem je prověřit, zda je střední hodnota výšky horní části tváře u čínské populace statisticky významně větší u malajské populace. Toto tvrzení je zněním alternativní hypotézy, zatímco nulová hypotéza tvoří doplněk tohoto tvrzení.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : *Střední hodnota výšky horní části tváře čínské populace je menší nebo rovna střední hodnotě výšky horní části tváře malajské populace.*

H_1 : *Střední hodnota výšky horní části tváře čínské populace je větší než střední hodnota výšky horní části tváře malajské populace.*

- **matematická formulace** nulové a alternativní hypotézy

$$H_0 : \mu_1 \leq \mu_2 \quad \rightarrow \quad \mu_1 - \mu_2 \leq \mu_0, \text{ kde } \mu_0 = 0$$

$$H_1 : \mu_1 > \mu_2 \quad \rightarrow \quad \mu_1 - \mu_2 > \mu_0, \text{ kde } \mu_0 = 0$$

(pravostranná alternativa)

2. Volba hladiny významnosti

- Hladinu významnosti volíme ze zadání $\alpha = 0.05$.

3. Testování kritickým oborem

- **Aritmetický průměr výběrových rozptylů**

$$\begin{aligned} S_* &= \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}} \\ &= \sqrt{\frac{(18 - 1) \times 4.563281^2 + (69 - 1) \times 4.952459^2}{18 + 69 - 2}} \\ &= \sqrt{\frac{17 \times 20.82353 + 68 \times 24.52685}{85}} \\ &= \sqrt{\frac{2021.826}{85}} = \sqrt{23.78619} \\ &= 4.877109 \doteq 4.8771 \end{aligned}$$

```
135 s1 <- sd(upface.Hc)
136 s2 <- sd(upface.Hm)
137 sh <- sqrt(((n1 - 1) * s1^2 + (n2 - 1) * s2^2) / (n1 + n2 - 2)) # 4.877109
```

- **Testovací statistika**

$$\begin{aligned} T_W &= \frac{(M_1 - M_2) - \mu_0}{S_* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \\ &= \frac{(72 - 70.13043) - 0}{4.877109 \sqrt{\frac{1}{18} + \frac{1}{69}}} \\ &= \frac{1.86957}{4.877109 \times 0.2646664} \\ &= \frac{1.86957}{1.290807} = 1.44837 \doteq 1.4484 \end{aligned}$$

```
138 alpha <- 0.05
139 mu0 <- 0
140 m1 <- mean(upface.Hc)
141 m2 <- mean(upface.Hm)
142 tw <- ((m1 - m2) - mu0) / (sh * sqrt(1 / n1 + 1 / n2)) # 1.448369
```

- **Kritický obor**

$$\begin{aligned} W &= \langle t_{n_1+n_2-2}(1-\alpha); \infty \rangle \\ &= \langle t_{18+69-2}(1-0.05); \infty \rangle \\ &= \langle t_{85}(0.95); \infty \rangle \\ &= \langle 1.662978; \infty \rangle \end{aligned}$$

```
143 qt(1 - alpha, n1 + n2 - 2) # 1.662978
```

- **Závěr testování**

Protože realizace testovací statistiky $t_W = 1.44837$ nenáleží do kritického oboru, tj. $t_W \notin W$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

4. Testování intervalem spolehlivosti

- **Interval spolehlivosti**

$$\begin{aligned}(d, \infty) &= \left(m_1 - m_2 - s_* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} t_{n_1+n_2-2}(1-\alpha); \infty \right) \\ &= \left((72 - 70.13043) - 4.877109 \sqrt{\frac{1}{18} + \frac{1}{69}} t_{85}(0.95); \infty \right) \\ &= (1.869565 - 4.877109 \times 0.2646664 \times 1.662978; \infty) \\ &= (-0.277018; \infty) \doteq (-0.2770; \infty)\end{aligned}$$

```
144 DH <- (m1 - m2) - sh * sqrt(1 / n1 + 1 / n2) * qt(1 - alpha, n1 + n2 - 2) # -0.2770188
```

- **Závěr testování**

Protože $\mu_0 = 0$ náleží do Waldova 95% empirického levostranného intervalu spolehlivosti, tj. $\mu_0 = 0 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

5. Testování p -hodnotou

- **p -hodnota**

$$\begin{aligned}p\text{-hodnota} &= \Pr(T_W \geq t_W) = 1 - \Pr(T_w < t_W) \\ &= 1 - \Pr(T_W \leq 1.44837) \\ &= 0.9999\end{aligned}$$

```
145 p.val <- 1 - pt(tw, n1 + n2 - 2) # 0.07559631
```

- **Závěr testování**

Protože p -hodnota = 0.07556 je větší než $\alpha = 0.05$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

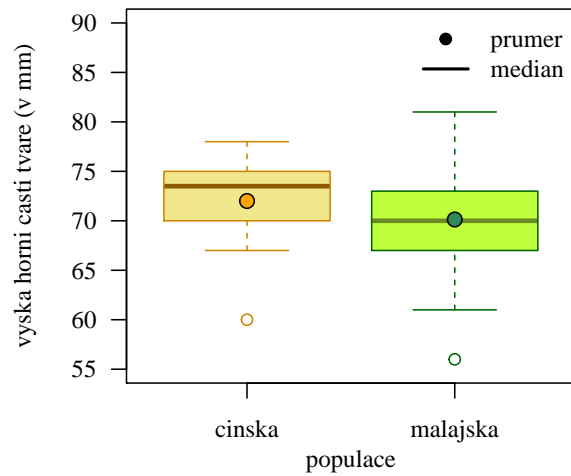
6. Grafická vizualizace výsledků testování

Vhodným grafem porovnávajícím střední hodnoty obou náhodných výběrů je krabicový diagram (viz obrázek 19).

7. Interpretace výsledků:

Na základě všech tří způsobů testování nezamítáme hypotézu o tom, že střední hodnota výšky horní části tváře čínské populace je menší nebo rovna střední hodnotě výšky horní části tváře malajské populace. Testování tedy neprokázalo, že by výška horní části tváře čínské populace byla statisticky významně vyšší než u malajské populace.

Poznámka: K otestování nulové hypotézy o rozdílu středních hodnot můžeme využít funkci `t.test()` s argumenty `conf.level = 0.95` pro volbu hladiny významnosti $\alpha = 0.05$, `alternative = 'greater'` pro volbu pravostranné alternativní hypotézy a `var.equal = T` pro volbu klasického t -testu.



Obrázek 19: Krabicový diagram výšky horní části tváře mužů čínské a malajské populace

```
146 t.test(upface.Hc, upface.Hm, conf.level = 0.95, alternative = 'greater', var.equal = T)
```

```

Two Sample t-test
data: upface.Hc and upface.Hm
t = 1.4484, df = 85, p-value = 0.0756
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 -0.2770188      Inf
sample estimates:
mean of x mean of y
 72.00000  70.13043

```

147
148
149
150
151
152
153
154
155
156
157

★

Příklad 10.8. Test o rozdílu středních hodnot

Mějme datový soubor 19-more-samples-correlations-skull.txt a proměnnou nose.B popisující šířku nosu v mm (viz sekce ??). Předpokládejme, že náhodná veličina X , popisující *šířku nosu německé populace*, pochází z normálního rozdělení $N(\mu_1, \sigma_1^2)$, a že náhodná veličina Y , popisující *šířku nosu bantuské populace*, pochází z normálního rozdělení $N(\mu_2, \sigma_2^2)$. Na hladině významnosti $\alpha = 0.01$ zjistěte, zda šířka nosu německé populace je statisticky významně menší než šířka nosu bantuské populace.

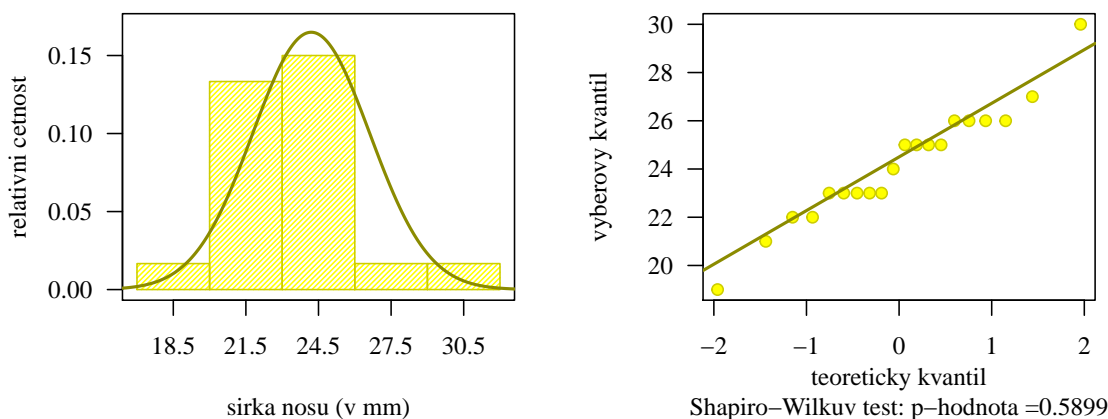
Řešení příkladu 10.8

Příkazem read.delim() načteme datový soubor a příkazem na.omit() odstraníme NA hodnoty. Pomocí operátoru [] vybereme z datové tabulky údaje o šířce nosu (nose.B) německé (pop == 'nem'), resp. bantuské (pop == 'ban') populace.

```
158 data <- read.delim('00-Data//19-more-samples-correlations-skull.txt')
159 data <- na.omit(data)
160 #head(data, n = 3)
161 nose.Bn <- data[data$pop == 'nem', 'nose.B']
162 nose.Bb <- data[data$pop == 'ban', 'nose.B']
163
164 n1 <- length(nose.Bn) # 20
165 n2 <- length(nose.Bb) # 14
```

Datový soubor obsahuje naměřené hodnoty šířek nosu 20 jedinců německé populace a 14 jedinců bantuské populace.

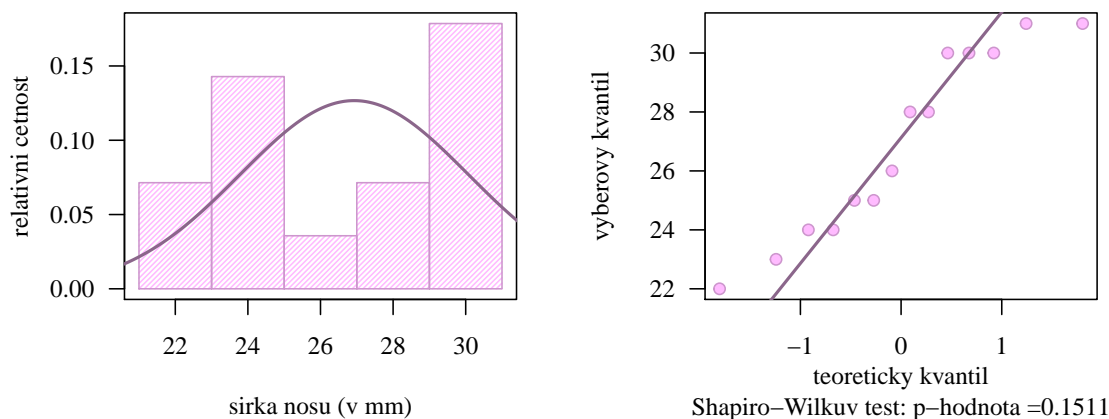
Nejprve otestujeme předpoklad normality obou náhodných výběrů. Na hladině významnosti α testujeme hypotézu H_{01} : *Náhodný výběr šířky nosu německé populace pochází z normálního rozdělení*, oproti alternativní hypotéze H_{11} : *Náhodný výběr šířky nosu německé populace nepochází z normálního rozdělení*. K otestování předpokladu normality použijeme Shapiro-Wilkův test, kvantilový diagram a histogram. Datový soubor rozdělíme do pěti ekvidistatních intervalů s šířkou 3 mm prostřednictvím stanovených hranic 17, 20, ..., 32 (viz obrázek 20).



Obrázek 20: Histogram a kvantilový diagram šířky nosu mužů německé populace

Protože p -hodnota = 0.5899 je větší než 0.05, nulovou hypotézu nezamítáme na hladině významnosti $\alpha = 0.05$. Náhodný výběr šířky nosu německé populace pochází z normálního rozdělení. Vykreslený histogram dostatečným způsobem kopíruje tvar křivky normálního rozdělení. Splnění předpokladu normality je tentokrát viditelné na kvantilovém diagramu, kde se všechny body pohybují okolo referenční přímky.

Dále testujeme hypotézu H_{02} : *Náhodný výběr šířky nosu bantuské populace pochází z normálního rozdělení*, oproti alternativní hypotéze H_{12} : *Náhodný výběr šířky nosu bantuské populace nepochází z normálního rozdělení*. Předpoklad normality otestujeme Shapiro-Wilkovým testem, kvantilovým diagramem a histogramem. Datový soubor rozdělíme do pěti ekvidistatních intervalů s šířkou 2 mm prostřednictvím stanovených hranic 21, 23, ..., 31 (viz obrázek 21).



Obrázek 21: Histogram a kvantilový diagram šířky nosu mužů bantuské populace

Protože p -hodnota = 0.1511 je větší než 0.05, nulovou hypotézu nezamítáme na hladině významnosti $\alpha = 0.05$. Náhodný výběr šířky nosu bantuské populace pochází z normálního rozdělení. Při pohledu na histogram bychom o normálním rozdělení naměřených hodnot mohli pochybovat. Nesmíme však zapomínat, že pro takto nízký počet pozorování může i nepříliš příznivě vypadající histogram reflektovat normálně rozdělená data. Kvantilový diagram naproti tomu normální charakter dat popisuje uspokojivě.

Zbývá ověřit potenciální shodu rozptylů obou náhodných výběrů. Na hladině významnosti $\alpha = 0.05$ testujeme hypotézu $H_0 : \sigma_1^2/\sigma_2^2 = 1$ oproti alternativní hypotéze $H_1 : \sigma_1^2/\sigma_2^2 \neq 1$. Nulovou hypotézu testujeme pomocí funkce `var.test()` na základě p -hodnoty (viz příklad 10.2).

```
166 var.test(nose.Bn, nose.Bb, alternative = 'two.sided')$p.val
```

```
[1] 0.2877778
```

167

Jelikož p -hodnota = 0.2878 je větší než 0.05, nulovou hypotézu o shodě rozptylů nezamítáme na hladině významnosti $\alpha = 0.05$. Na základě výsledků testování tedy předpokládáme, že rozptyly obou výběrů jsou shodné.

Protože oba náhodné výběry pochází z normálních rozdělení, o jejichž rozptylech předpokládáme, že jsou shodné, použijeme na otestování hypotézy ze zadání klasický dvouvýběrový t -test o rozdílu středních hodnot $\mu_1 - \mu_2$. V zadání není uvedeno znění nulové hypotézy. Zadaným úkolem je prověřit, zda je šířka nosu německé populace statisticky významně menší, než šířka nosu malajské populace. Toto tvrzení je zněním alternativní hypotézy. Nulová hypotéza tvoří potom doplněk tohoto tvrzení.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy
 H_0 : Střední hodnota šířky nosu německé populace je větší nebo rovna střední hodnotě šířky nosu bantuské populace.
 H_1 : Střední hodnota šířky nosu německé populace je menší než střední hodnota šířky nosu bantuské populace.
- **matematická formulace** nulové a alternativní hypotézy
 $H_0 : \mu_1 \geq \mu_2 \rightarrow \mu_1 - \mu_2 \geq \mu_0$, kde $\mu_0 = 0$
 $H_1 : \mu_1 < \mu_2 \rightarrow \mu_1 - \mu_2 < \mu_0$, kde $\mu_0 = 0$
 (levostranná alternativa)

2. Volba hladiny významnosti

- Hladina významnosti $\alpha = 0.01$ (viz zadání příkladu).

3. Testování kritickým oborem

- Aritmetický průměr výběrových rozptylů

$$\begin{aligned} S_* &= \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}} \\ &= \sqrt{\frac{(20 - 1) \times 2.419221^2 + (14 - 1) \times 3.149219^2}{20 + 14 - 2}} \\ &= \sqrt{\frac{19 \times 5.85263 + 13 \times 9.91758}{32}} \\ &= \sqrt{\frac{240.1285}{32}} = \sqrt{7.504016} \\ &= 2.739346 \doteq 2.7393 \end{aligned}$$

```
168 s1 <- sd(nose.Bn)
169 s2 <- sd(nose.Bb)
170 sh <- sqrt(((n1 - 1) * s1^2 + (n2 - 1) * s2^2) / (n1 + n2 - 2)) # 2.739346
```

- Testovací statistika

$$\begin{aligned} T_W &= \frac{(M_1 - M_2) - \mu_0}{S_* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \\ &= \frac{(24.2 - 26.92857) - 0}{2.739346 \sqrt{\frac{1}{20} + \frac{1}{14}}} \\ &= \frac{-2.72857}{2.739346 \times 0.348466} \\ &= \frac{-2.72857}{0.9545689} = -2.85843 \doteq -2.8584 \end{aligned}$$

```
171 alpha <- 0.01
172 mu0 <- 0
173 m1 <- mean(nose.Bn)
174 m2 <- mean(nose.Bb)
175 tw <- ((m1 - m2) - mu0) / (sh * sqrt(1 / n1 + 1 / n2)) # -2.858432
```

- Kritický obor

$$\begin{aligned} W &= (-\infty; t_{n_1+n_2-2}(\alpha)) \\ &= (-\infty; t_{20+14-2}(0.01)) \\ &= (-\infty; t_{32}(0.01)) \\ &= (-\infty; -2.448678) \end{aligned}$$

```
176 qt(alpha, n1 + n2 - 2) # -2.448678
```

- Závěr testování

Protože realizace testovací statistiky $t_W = -2.8584$ náleží do kritického oboru, tj. $t_W \in W$, H_0 zamítáme na hladině významnosti $\alpha = 0.01$.

4. Testování intervalem spolehlivosti

- Interval spolehlivosti

$$\begin{aligned}(d, \infty) &= \left(-\infty; m_1 - m_2 - s_* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} t_{n_1+n_2-2}(1-\alpha); \infty \right) \\ &= \left(-\infty; (24.2 - 26.92857) - 2.739346 \sqrt{\frac{1}{20} + \frac{1}{14}} t_{32}(0.01) \right) \\ &= (-\infty; -2.72857 - 2.739346 \times 0.348466 \times (-2.448678)) \\ &= (-\infty; -0.39113) \doteq (-\infty; -0.3911)\end{aligned}$$

```
177 HH <- (m1 - m2) - sh * sqrt(1 / n1 + 1 / n2) * qt(alpha, n1 + n2 - 2) # -0.3911394
```

- Závěr testování

Protože $\mu_0 = 0$ nenáleží do Waldova 99% empirického pravostranného intervalu spolehlivosti, tj. $\mu_0 = 0 \notin IS$, H_0 zamítáme na hladině významnosti $\alpha = 0.01$.

5. Testování p -hodnotou

- p -hodnota

$$\begin{aligned}p\text{-hodnota} &= \Pr(T_W \leq t_W) \\ &= \Pr(T_W \leq -2.858432) \\ &= 0.003714569 \doteq 0.0037\end{aligned}$$

```
178 p.val <- pt(tw, n1 + n2 - 2) # 0.003714569
```

- Závěr testování

Protože p -hodnota = 0.0037 je menší než $\alpha = 0.01$, H_0 zamítáme na hladině významnosti $\alpha = 0.01$.

6. Grafická vizualizace výsledků testování

Výsledek testování vizualizujeme pomocí krabicového diagramu (viz obrázek 22).

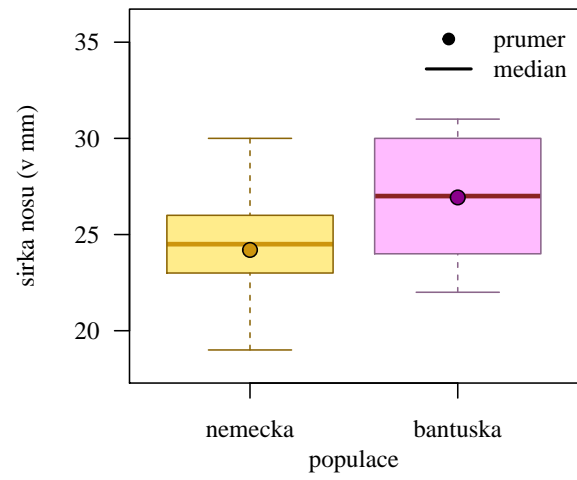
7. **Interpretace výsledků:** Na základě všech tří způsobů testování zamítáme hypotézu o tom, že střední hodnota šířky nosu německé populace je větší nebo rovna střední hodnotě šířky nosu bantuské populace. Testování tedy prokázalo, že šířka nosu německé populace je statisticky významně nižší než šířka nosu bantuské populace.

Poznámka: K otestování nulové hypotézy o rozdílu středních hodnot můžeme využít funkci `t.test()` s argumenty `conf.level = 0.99` pro volbu hladiny významnosti $\alpha = 0.01$, `alternative = 'less'` pro volbu levostranné alternativní hypotézy a `var.equal = T` pro volbu klasického t -testu.

```
179 t.test(nose.Bn, nose.Bb, conf.level = 0.99, alternative = 'less', var.equal = T)
```

```
Two Sample t-test
data: nose.Bn and nose.Bb
t = -2.8584, df = 32, p-value = 0.003715
alternative hypothesis: true difference in means is less than 0
99 percent confidence interval:
 -Inf -0.3911394
sample estimates:
mean of x mean of y
24.20000 26.92857
```

180
181
182
183
184
185
186
187
188
189
190



Obrázek 22: Krabicový diagram šířky nosu mužů německé a bantuské populace



Příklad 10.9. Test o rozdílu středních hodnot

Mějme datový soubor `10-two-samples-means-birth.txt` a proměnnou `weight.W` popisující porodní hmotnost novorozence v `g` (viz sekce ??). Na hladině významnosti $\alpha = 0.05$ zjistíte, zda se porodní porodní hmotnost novorozence bez staršího sourozence a porodní hmotnost novorozence s jedním starším sourozencem liší.

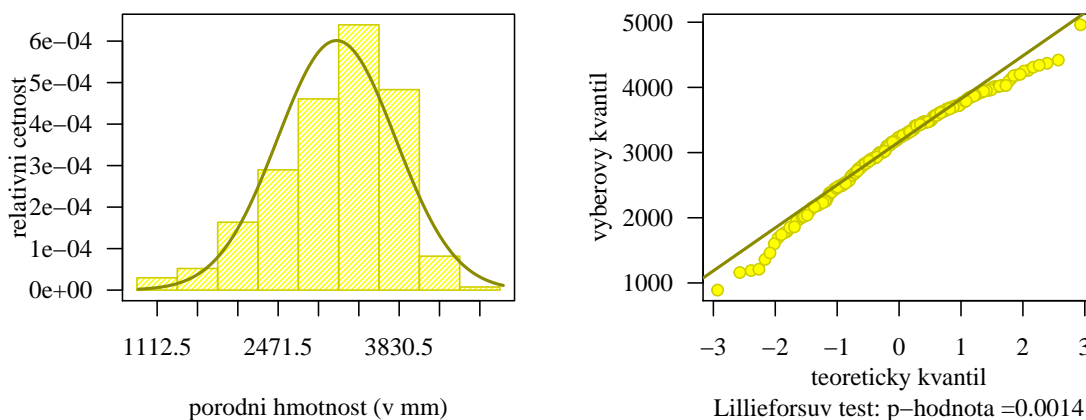
Řešení příkladu 10.9

Příkazem `read.delim()` načteme datový soubor a odstraníme chybějící hodnoty. Pomocí operátoru `[]` vybereme z datové tabulky údaje o porodní hmotnosti (`weight.W`) novorozenců bez staršího sourozence (`o.sib.N == 0`), resp. novorozenců s jedním starším sourozencem (`o.sib.N == 1`) populace.

```
191 data <- read.delim('00-Data//10-two-samples-means-birth.txt')
192 data <- na.omit(data)
193 #head(data, n = 3)
194 weight.W0 <- data[data$o.sib.N == 0, 'birth.W']
195 weight.W1 <- data[data$o.sib.N == 1, 'birth.W']
196
197 n1 <- length(weight.W0) # 297
198 n2 <- length(weight.W1) # 276
```

Datový soubor obsahuje naměřené porodní hmotnosti 297 novorozenců bez staršího sourozence a 276 novorozenců s jedním starším sourozencem.

Nejprve otestujeme předpoklad normality obou náhodných výběrů ($\alpha = 0.05$). Protože rozsah náhodného výběru porodních hmotností novorozenců bez staršího sourozence je větší než 30, použijeme na ověření předpokladu normality náhodného výběru Lillieforsův test v kombinaci s kvantilovým diagramem a histogramem (viz obrázek 23). Datový soubor rozdělíme do devíti ekvidistatních intervalů s šířkou 453 mm prostřednictvím stanovených hranic 886, 1339, ..., 4064 g.



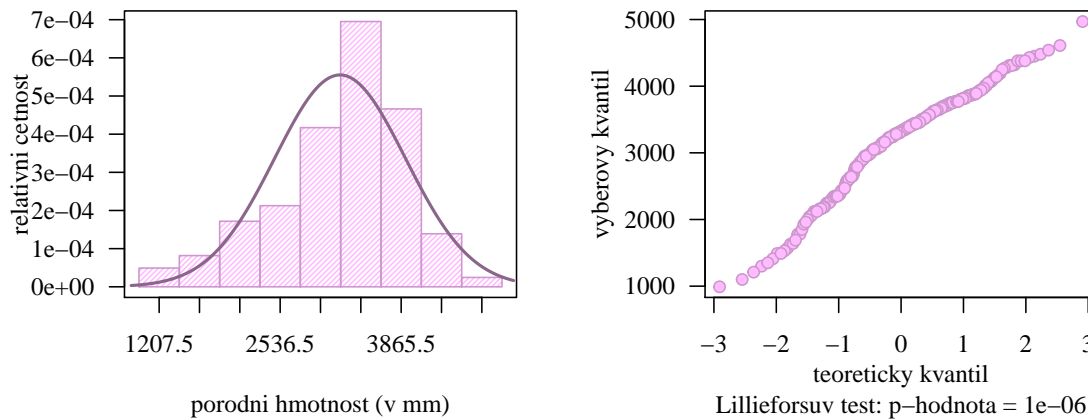
Obrázek 23: Histogram a kvantilový diagram porodní hmotnosti novorozenců bez staršího sourozence

Protože p -hodnota = 0.0014 je menší než 0.05, nulovou hypotézu zamítáme na hladině významnosti $\alpha = 0.05$. Vykreslený histogram je oproti křivce hustoty normálního rozdělení posunutý doprava. Taktéž kvantilový diagram ukazuje odlehlost vykreslených bodů od referenční přímky. Náhodný výběr porodní hmotnost novorozenců bez staršího sourozence nepochází z normálního rozdělení.

Dále testujeme hypotézu o normalitě náhodného výběru porodních hmotností novorozenců s jedním starším sourozencem. Předpoklad normality otestujeme opět Lillieforsovým testem, kvantilovým diagramem a histogramem. Datový soubor rozdělíme do devíti ekvidistatních intervalů s šířkou 443 mm prostřednictvím stanovených hranic 986, 1429, ..., 4973 g (viz obrázek 24).

```
Error in quantile(y, probs, names = FALSE, type = qtype, na.rm = TRUE): object 'weigh.W1' not found
```

199



Obrázek 24: Histogram a kvantilový diagram porodní hmotnosti novorozenců s jedním starším sourozencem

Protože p -hodnota = 1×10^{-6} je menší než 0.05, nulovou hypotézu zamítáme na hladině významnosti $\alpha = 0.05$. Při pohledu na histogram vidíme jednak posunutí naměřených hodnot oproti křivce normálního rozdělení směrem doprava a větší špičatost. Kvantilový digram ukazuje odlehlost hodnot od referenční křivky. Náhodný výběr porodních hmotností novorozenců s jedním starším sourozencem nepochází z normálního rozdělení.

Jelikož ani jeden z náhodných výběrů nepochází z normálního rozdělení, nemůžeme na hypotézu ze zadání použít parametrický test o rozdílu středních hodnot $\mu_1 - \mu_2$. Hypotézu ze zadání bychom testovali některou z metod uvedených v sekci ??.



10.3 Welchův dvouvýběrový t -test o rozdílu středních hodnot $\mu_1 - \mu_2$

Nechť X_{11}, \dots, X_{1n_1} je náhodný výběr z $N(\mu_1, \sigma_1^2)$, a X_{21}, \dots, X_{2n_2} je na něm nezávislý náhodný výběr z rozdělení $N(\mu_2, \sigma_2^2)$, přičemž $n_1 \geq 2$, $n_2 \geq 2$ a σ_1^2 a σ_2^2 jsou neznámé různé rozptyly, tj. $\sigma_1^2 \neq \sigma_2^2$. Nechť μ_0 je konstanta. Na hladině významnosti α testujeme jednu z následujících tří hypotéz oproti příslušné alternativní hypotéze.

$$\begin{array}{lll} H_{01} : \mu_1 - \mu_2 = \mu_0 & \text{oproti} & H_{11} : \mu_1 - \mu_2 \neq \mu_0 \quad (\text{oboustranná alt.}) \\ H_{02} : \mu_1 - \mu_2 \leq \mu_0 & \text{oproti} & H_{12} : \mu_1 - \mu_2 > \mu_0 \quad (\text{pravostranná alt.}) \\ H_{03} : \mu_1 - \mu_2 \geq \mu_0 & \text{oproti} & H_{13} : \mu_1 - \mu_2 < \mu_0 \quad (\text{levostranná alt.}) \end{array}$$

Test nazýváme Welchův dvouvýběrový t -test o rozdílu středních hodnot $\mu_1 - \mu_2$. Testovací statistika má tvar

$$T_W = \frac{(M_1 - M_2) - \mu_0}{S_* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}. \quad (10.3)$$

kde M_1 je výběrový průměr a n_1 je rozsah prvního náhodného výběru, M_2 je výběrový průměr a n_2 je rozsah druhého náhodného výběru, S_* je aritmetický průměr výběrových rozptylů S_1^2 a S_2^2 (viz kapitola ??) a μ_0 je konstanta z nulové hypotézy. Za platnosti nulové hypotézy pochází statistika T_W ze Studentova rozdělení o df stupních volnosti, tj.

$$T_W = \frac{(M_1 - M_2) - \mu_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \stackrel{H_0}{\sim} t_{df},$$

kde

$$df = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1-1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2-1}} \quad (10.4)$$

je Welchova aproximace stupňů volnosti Studentova rozdělení navržená tak, aby zohledňovala různost rozptylů σ_1^2 a σ_2^2 . Kritický obor podle zvolené alternativní hypotézy má tvar

$$\begin{array}{ll} H_{11} : \mu \neq \mu_0 & W = (-\infty; t_{df}(\alpha/2)) \cup (t_{df}(1 - \alpha/2); \infty) \\ H_{12} : \mu > \mu_0 & W = (t_{df}(1 - \alpha); \infty) \\ H_{13} : \mu < \mu_0 & W = (-\infty; t_{df}(\alpha)) \end{array}$$

kde $t_{df}(\alpha/2)$, $t_{df}(1 - \alpha/2)$, $t_{df}(\alpha)$ a $t_{df}(1 - \alpha)$ jsou kvantily Studentova rozdělení o df stupních volnosti, jejichž hodnoty získáme pomocí \mathcal{R} a implementované funkce $qt()$.

Interval spolehlivosti má podle zvolené alternativní hypotézy jeden z následujících tvarů

$$\begin{array}{ll} H_{11} : \mu_1 - \mu_2 \neq \mu_0 & (d, h) = \left(m_1 - m_2 - \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} t_{df}(1 - \alpha/2); m_1 - m_2 - \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} t_{df}(\alpha/2) \right) \\ H_{12} : \mu_1 - \mu_2 > \mu_0 & (d, \infty) = \left(m_1 - m_2 - \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} t_{df}(1 - \alpha); \infty \right) \\ H_{13} : \mu_1 - \mu_2 < \mu_0 & (-\infty, h) = \left(-\infty; m_1 - m_2 - \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} t_{df}(\alpha) \right) \end{array}$$

p -hodnota má v závislosti na zvolené alternativní hypotéze jeden z následujících tvarů

$$\begin{array}{ll} H_{11} : \mu \neq \mu_0 & p\text{-hodnota} = 2 \min\{\Pr(T_W \leq t_W), \Pr(T_W > t_W)\} \\ H_{12} : \mu > \mu_0 & p\text{-hodnota} = \Pr(T_W > t_W) = 1 - \Pr(T_W \leq t_W) \\ H_{13} : \mu < \mu_0 & p\text{-hodnota} = \Pr(T_W \leq t_W) \end{array}$$

kde T_W je náhodná veličina, t_W je realizace testovací statistiky T_W (viz vzorec 10.3), tedy konkrétní číslo, a $\Pr(T_W \leq t_W)$ je distribuční funkce Studentova rozdělení o df stupních volnosti, jejíž hodnotu získáme pomocí \mathcal{R} a implementované funkce $pt()$.

Příklad 10.10. Test o rozdílu středních hodnot

Mějme datový soubor 16-anova-head.txt a proměnnou bigo.W popisující šířku dolní čelisti v mm (viz sekce ??). Předpokládejme, že náhodná veličina X , popisující šířku dolní čelisti u žen orientovaných heterosexuálně, pochází z normálního rozdělení $N(\mu_1, \sigma_1^2)$, a že náhodná veličina Y , popisující šířku dolní čelisti žen orientovaných jinak než výhradně heterosexuálně, pochází z normálního rozdělení $N(\mu_2, \sigma_2^2)$. Na hladině významnosti $\alpha = 0.05$ zjistíte, zda šířka dolní čelisti žen orientovaných heterosexuálně je statisticky významně menší než šířka dolní čelisti žen orientovaných jinak než výhradně heterosexuálně.

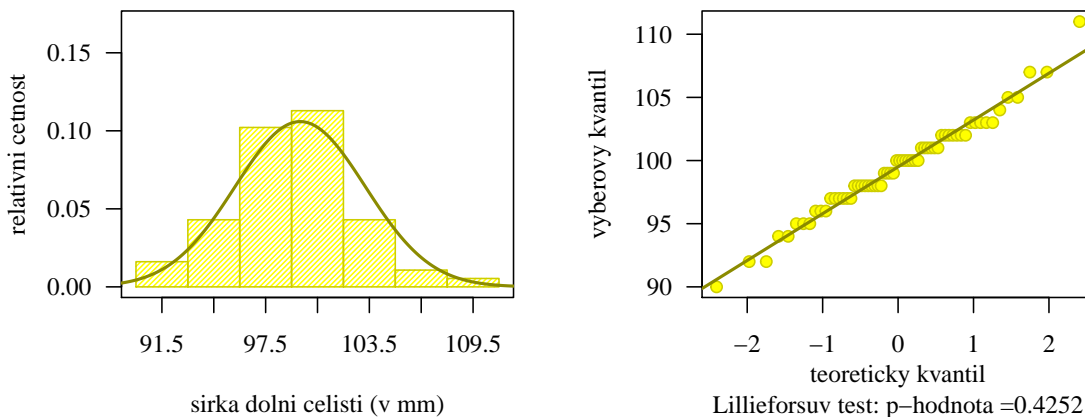
Řešení příkladu 10.10

Příkazem read.delim() načteme datový soubor a příkazem na.omit() odstraníme NA hodnoty. Pomocí operátoru [] vybereme z datové tabulky údaje o šířce dolní čelisti (bigo.W) u žen (sex == 'f') orientovaných heterosexuálně (sexor == 'op'), resp. orientovaných jinak než heterosexuálně (sexor == 'sa') .

```
200 data <- read.delim('00-Data//16-anova-head.txt')
201 data <- na.omit(data)
202 #head(data, n = 3)
203 bigo.Wfo <- data[data$sex == 'f' & data$sexor == 'op', 'bigo.W']
204 bigo.Wfs <- data[data$sex == 'f' & data$sexor == 'sa', 'bigo.W']
205
206 n1 <- length(bigo.Wfo) # 62
207 n2 <- length(bigo.Wfs) # 23
```

Datový soubor obsahuje hodnoty šířek dolní čelisti u 62 žen orientovaných heterosexuálně a 23 žen orientovaných jinak než heterosexuálně.

Řešení příkladu začneme testováním předpokladu normality obou náhodných výběrů. Na hladině významnosti α testujeme hypotézu H_{01} : *Náhodný výběr šířky dolní čelisti u žen orientovaných heterosexuálně pochází z normálního rozdělení.* oproti alternativní hypotéze H_{11} : *Náhodný výběr šířky dolní čelisti u žen orientovaných heterosexuálně nepochází z normálního rozdělení.* K otestování normality náhodného výběru použijeme Lillieforsův test, kvantilový diagram a histogram (viz obrázek 25). Datový soubor rozdělíme na základě Sturgesova pravidla do sedmi třídicích intervalů s ekvidistantní šířkou 3 mm prostřednictvím stanovených hranic 90, 93, ..., 111.

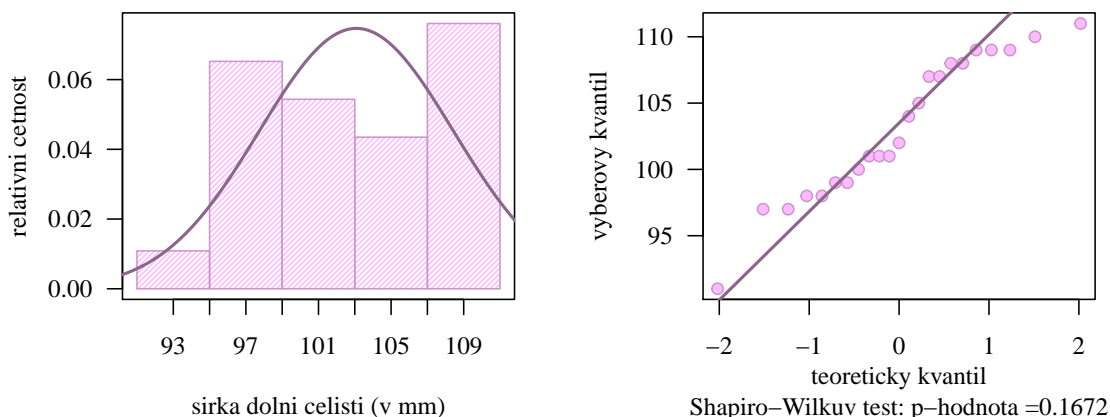


Obrázek 25: Histogram a kvantilový diagram šířky dolní čelisti heterosexuálně orientovaných žen

Protože p -hodnota = 0.4252 je větší než 0.05, nulovou hypotézu nezamítáme na hladině významnosti $\alpha = 0.05$. Náhodný výběr šířky dolní čelisti heterosexuálně orientovaných žen pochází z normálního rozdělení. Histogram, stejně jako kvantilový diagram, předpoklad normality též potvrzují.

Dále testujeme hypotézu H_{02} : *Náhodný výběr šířky dolní čelisti žen orientovaných jinak než heterosexuálně pochází z normálního rozdělení.* oproti alternativní hypotéze H_{12} : *Náhodný výběr šířky dolní čelisti žen orientovaných jinak než heterosexuálně nepochází z normálního rozdělení.* Předpoklad normality otestujeme Shapiro-Wilkovým testem,

kvantilovým diagramem a histogramem. Datový soubor rozdělíme do pěti ekvidistatních intervalů s šířkou 4 mm prostřednictvím stanovených hranic 91, 95 . . . , 111 (viz obrázek 26).



Obrázek 26: Histogram a kvantilový diagram šířky dolní čelisti žen orientovaných jinak než heterosexuálně

Protože p -hodnota = 0.1672 je větší než 0.05, nulovou hypotézu nezamítáme na hladině významnosti $\alpha = 0.05$. Náhodný výběr šířky dolní čelisti žen orientovaných jinak než heterosexuálně pochází z normálního rozdělení.

Zbývá ověřit předpoklad shody rozptylů obou náhodných výběrů. Na hladině významnosti $\alpha = 0.05$ testujeme nulovou hypotézu $H_0 : \sigma_1^2/\sigma_2^2 = 1$ oproti hypotéze $H_1 : \sigma_1^2/\sigma_2^2 \neq 1$. Nulovou hypotézu testujeme pomocí funkce `var.test()` na základě p -hodnoty (viz příklad 10.2).

```
208 var.test(bigo.Wfo, bigo.Wfs, alternative = 'two.sided')$p.val
```

```
[1] 0.0336774
```

209

Jelikož p -hodnota = 0.03368 je menší než 0.05, nulovou hypotézu o shodě rozptylů zamítáme na hladině významnosti $\alpha = 0.05$. Rozptyly obou náhodných výběrů se statisticky významně liší.

Oba náhodné výběry pochází z normálního rozdělení. Na otestování hypotézy ze zadání tedy můžeme použít parametrický test. Protože však není splněn předpoklad shody rozptylů obou náhodných výběrů, použijeme na otestování hypotézy ze zadání namísto klasického dvouvýběrového t -tesu Welchův dvouvýběrový t -test. Zadaným úkolem je prověřit, zda je šířka dolní čelisti heterosexuálně orientovaných žen statisticky významně menší, než šířka dolní čelisti žen orientovaných jinak než heterosexuálně. Toto tvrzení je zněním alternativní hypotézy a nulová hypotéza tvoří doplněk tohoto tvrzení.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : Střední hodnota šířky dolní čelisti heterosexuálně orientovaných žen je větší nebo rovna střední hodnotě šířky dolní čelisti žen orientovaných jinak než heterosexuálně.

H_1 : Střední hodnota šířky dolní čelisti heterosexuálně orientovaných žen je menší než střední hodnota šířky dolní čelisti žen orientovaných jinak než heterosexuálně.

- **matematická formulace** nulové a alternativní hypotézy

$H_0 : \mu_1 \geq \mu_2 \rightarrow \mu_1 - \mu_2 \geq \mu_0$, kde $\mu_0 = 0$

$H_1 : \mu_1 < \mu_2 \rightarrow \mu_1 - \mu_2 < \mu_0$, kde $\mu_0 = 0$

(levostranná alternativa)

2. Volba hladiny významnosti

- Hladina významnosti $\alpha = 0.05$ (viz zadání příkladu).

3. Testování kritickým oborem

- Testovací statistika

$$\begin{aligned}
 T_W &= \frac{(M_1 - M_2) - \mu_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \\
 &= \frac{(99.5 - 103.08695) - 0}{\sqrt{\frac{3.766766^2}{62} + \frac{5.342055^2}{23}}} \\
 &= \frac{-3.58696}{\sqrt{1.46961}} \\
 &= \frac{-3.58696}{1.212275} = -2.958867 \doteq -2.9589
 \end{aligned}$$

```

210 alpha <- 0.05
211 mu0 <- 0
212 m1 <- mean(bigo.Wfo)
213 m2 <- mean(bigo.Wfs)
214 tw <- ((m1 - m2) - mu0) / (sqrt(s1 ^ 2 / n1 + s2 ^ 2 / n2)) # -2.958864

```

- Welchova aproximace počtu stupňů volnosti

$$\begin{aligned}
 df &= \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1-1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2-1}} \\
 &= \frac{\left(\frac{3.766766^2}{62} + \frac{5.342055^2}{23}\right)^2}{\frac{\left(\frac{3.766766^2}{62}\right)^2}{62-1} + \frac{\left(\frac{5.342055^2}{23}\right)^2}{23-1}} \\
 &= \frac{(0.2288473 + 1.240763)^2}{\frac{0.2288473^2}{61} + \frac{1.240763^2}{22}} \\
 &= \frac{2.159754}{0.07083549} \\
 &= 30.48972
 \end{aligned}$$

```

215 df <- (s1 ^ 2 / n1 + s2 ^ 2 / n2)^2 / ((s1 ^ 2 / n1) ^ 2 / (n1 - 1) + (s2 ^ 2 / n2) ^ 2 / (n2 - 1)) # 30.48972

```

- Kritický obor

$$\begin{aligned}
 W &= (-\infty; t_{df}(\alpha)) \\
 &= (-\infty; t_{30.48972}(0.05)) \\
 &= (-\infty; -2.448678)
 \end{aligned}$$

```

216 qt(alpha, df) # -1.696393

```


- **Závěr testování**

Protože realizace testovací statistiky $t_W = -2.9589$ náleží do kritického oboru, tj. $t_W \in W$, H_0 zamítáme na hladině významnosti $\alpha = 0.05$.

4. Testování intervalem spolehlivosti

- **Interval spolehlivosti**

$$\begin{aligned} (d, \infty) &= \left(-\infty; m_1 - m_2 - \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} t_{df}(1 - \alpha); \infty \right) \\ &= \left(-\infty; (99.5 - 103.08695) - \sqrt{\frac{3.766766^2}{63} + \frac{5.342055^2}{23}} t_{30.48972}(0.05) \right) \\ &= (-\infty; -3.586957 - 1.212275 \times (-1.696393)) \\ &= (-\infty; -1.530462) \doteq (-\infty; -1.5305) \end{aligned}$$

```
217 HH <- (m1 - m2) - sqrt(s1^2 / n1 + s2^2 / n2) * qt(alpha, df) # -1.530462
```

- **Závěr testování**

Protože $\mu_0 = 0$ nenáleží do Waldova 95% empirického pravostranného intervalu spolehlivosti, tj. $\mu_0 = 0 \notin IS$, H_0 zamítáme na hladině významnosti $\alpha = 0.05$.

5. Testování p -hodnotou

- **p -hodnota**

$$\begin{aligned} p\text{-hodnota} &= \Pr(T_W \leq t_W) \\ &= \Pr(T_W \leq -2.958864) \\ &= 0.002961084 \doteq 0.002961 \end{aligned}$$

```
218 p.val <- pt(tw, df) # 0.002961084
```

- **Závěr testování**

Protože p -hodnota = 0.002961 je menší než $\alpha = 0.05$, H_0 zamítáme na hladině významnosti $\alpha = 0.01$.

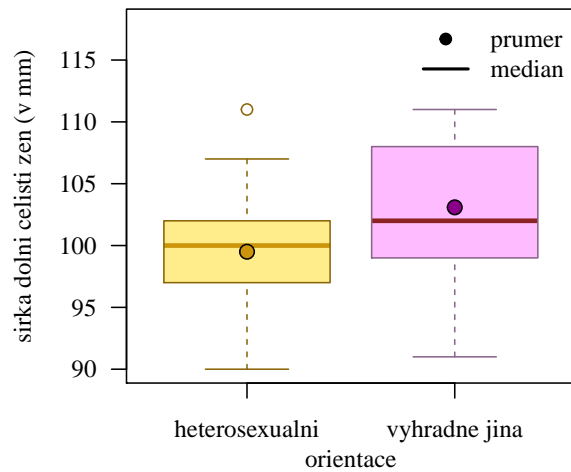
6. Grafická vizualizace výsledků testování

Výsledek testování vizualizujeme pomocí krabicového diagramu (viz obrázek 27).

7. Interpretace výsledků:

Na základě všech tří způsobů testování zamítáme hypotézu o tom, že střední hodnota šířky dolní čelisti heterosexuálně orientovaných žen je větší nebo rovna střední hodnotě šířky dolní čelisti žen orientovaných jinak než heterosexuálně. Šířka dolní čelisti žen orientovaných heterosexuálně je statisticky významně menší než šířka dolní čelisti žen orientovaných jinak než heterosexuálně.

Poznámka: K otestování nulové hypotézy o rozdílu středních hodnot můžeme využít funkci `t.test()`. Vstupními parametry budou nejprve dva vektory reprezentující náhodné výběry, tj. `big0.Wfo` a `big0.Wfs`, dále hodnota hladiny významnosti α zadaná prostřednictvím koeficientu spolehlivosti $1 - \alpha$ nastavením hodnoty argumentu `conf.level = 0.95`, typ zvolené alternativní hypotézy (levostranná) zadaný pomocí argumentu `alternative == 'less'` a nakonec argument `var.equal == F` pro volbu Welchova dvouvýběrového t -testu založeného na předpokladu rozdílu mezi rozptyly σ_1^2 a σ_2^2 .



Obrázek 27: Krabicový diagram šířky dolní čelisti žen orientovaných heterosexuálně a žen orientovaných jinak než heterosexuálně

```
219 t.test(bigo.Wfo, bigo.Wfs, conf.level = 0.95, alternative = 'less', var.equal = F)
```

```

Welch Two Sample t-test

data:  bigo.Wfo and bigo.Wfs
t = -2.9589, df = 30.49, p-value = 0.002961
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
    -Inf -1.530462
sample estimates:
mean of x mean of y
  99.500   103.087

```

220
221
222
223
224
225
226
227
228
229
230

Součástí výstupu je hodnota testovací statistiky $t = -2.9589$, Welchova aproximace počtu stupňů volnosti Studentova rozdělení $\text{num df} = 30.49$, hranice intervalu spolehlivosti $-\text{Inf}$ a -1.530462 a p -hodnota $p\text{-value} = 0.002961$. Jediné, co musíme stanovit zvlášť, je dolní hranice kritického oboru. ★

Příklad 10.11. Test o rozdílu středních hodnot

Mějme datový soubor 13-two-samples-correlations-trunk.txt a proměnnou tru.L popisující délku trupu v mm (viz sekce ??). Předpokládejme, že náhodná veličina X , popisující *délku trupu mužů*, pochází z normálního rozdělení $N(\mu_1, \sigma_1^2)$, a že náhodná veličina Y , popisující *délku trupu žen*, pochází z normálního rozdělení $N(\mu_2, \sigma_2^2)$. Na hladině významnosti $\alpha = 0.01$ zjistěte, zda existuje statisticky významný rozdíl mezi délkou trupu mužů a žen.

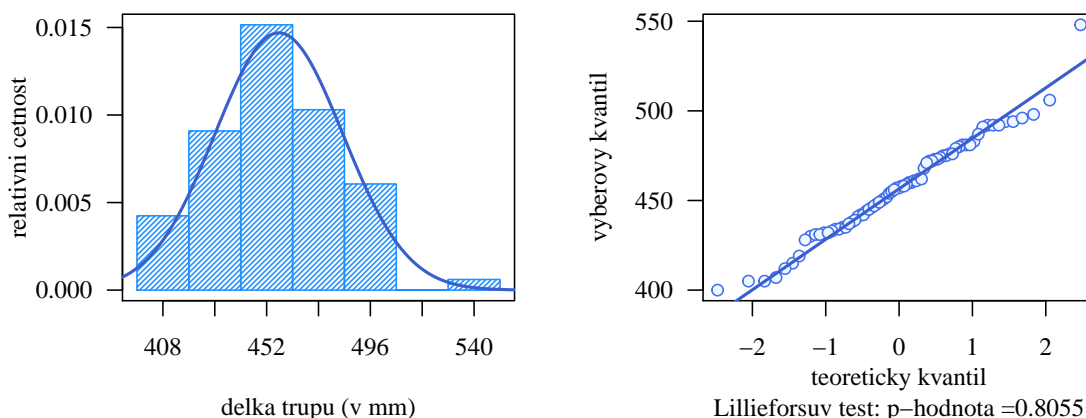
Řešení příkladu 10.11

Příkazem read.delim() načteme datový soubor a příkazem na.omit() odstraníme případné NA hodnoty. Pomocí operátoru [] vybereme z datové tabulky údaje o délce trupu (tru.L) mužů (sex == 'm'), resp. žen (sex == 'f') a zjistíme rozsahy obou náhodných výběrů.

```
231 data <- read.delim('00-Data//13-two-samples-correlations-trunk.txt')
232 data <- na.omit(data)
233 #head(data, n = 3)
234 tru.Lm <- data[data$sex == 'm', 'tru.L']
235 tru.Lf <- data[data$sex == 'f', 'tru.L']
236
237 n1 <- length(tru.Lm) # 75
238 n2 <- length(tru.Lf) # 100
```

Datový soubor obsahuje údaje o délkách trupu 75 mužů a 100 žen. Řešení příkladu vede na test o rozdílu středních hodnot. Nejprve otestujeme normalitu obou náhodných výběrů. Hladinu významnosti pro testy normality zvolíme $\alpha = 0.05$.

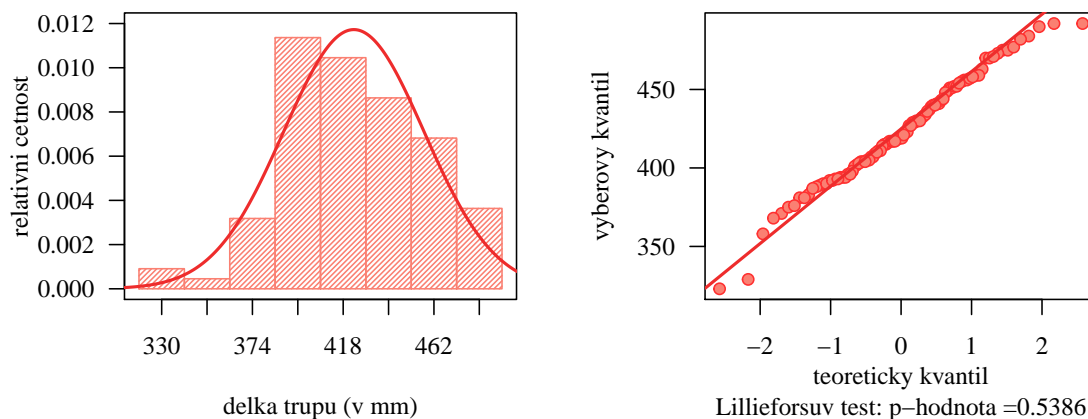
Nejprve testujeme hypotézu H_{01} : *Náhodný výběr délky trupu mužů pochází z normálního rozdělení*. oproti alternativní hypotéze H_{11} : *Náhodný výběr délky trupu mužů nepochází z normálního rozdělení*. Normalitu otestujeme Lillieforsovým testem, kvantilovým diagramem a histogramem (viz obrázek 28). Naměřené hodnoty rozdělíme do sedmi ekvidistatních intervalů s šířkou 22mm prostřednictvím stanovených hranic 397, 419, ..., 154.



Obrázek 28: Histogram a kvantilový diagram délky trupu mužů

Protože p -hodnota = 0.8055 je větší než 0.05, nulovou hypotézu nezamítáme na hladině významnosti $\alpha = 0.05$. Náhodný výběr naměřených délek trupu mužů pochází z normálního rozdělení. Stejně tak je normalita zřetelná z histogramu, kde data kopírují tvar křivky hustoty a kvantilového diagramu, kde se vykreslené body zdržují okolo referenční přímky.

Analogicky testujeme hypotézu H_{02} : *Náhodný výběr délky trupu žen pochází z normálního rozdělení*. oproti alternativní hypotéze H_{12} : *Náhodný výběr délky trupu žen nepochází z normálního rozdělení*. Normalitu otestujeme Lillieforsovým testem. V rámci histogramu rozdělíme soubor do osmi ekvidistatních intervalů s šířkou 22 mm prostřednictvím stanovených hranic 228, 250 ..., 495 (viz obrázek 29).



Obrázek 29: Histogram a kvantilový diagram délky trupu žen

Protože p -hodnota = 0.5386 je větší než 0.05, nulovou hypotézu nezamítáme na hladině významnosti $\alpha = 0.05$. Náhodný výběr délky trupu žen pochází z normálního rozdělení. Vykreslený histogram i kvantilový diagram podporují výsledek testování.

Dále ověříme, zda rozptyly obou výběrů jsou shodné. Na hladině významnosti $\alpha = 0.05$ testujeme hypotézu $H_0 : \sigma_1^2/\sigma_2^2 = 1$ oproti alternativní hypotéze $H_1 : \sigma_1^2/\sigma_2^2 \neq 1$. Nulovou hypotézu otestujeme pomocí p -hodnoty získané příkazem `var.test()` (viz příklad 10.2).

```
239 var.test(tru.Lm, tru.Lf, alternative = 'two.sided')$p.val # 0.04256397
```

```
[1] 0.04256397
```

240

Jelikož p -hodnota = 0.04256 je menší než 0.05, nulovou hypotézu o shodě rozptylů zamítáme na hladině významnosti $\alpha = 0.05$. Na základě výsledků testování tedy předpokládáme, že oba rozptyly jsou shodné.

Protože oba náhodné výběry pochází z normálního rozdělení, použijeme na otestování hypotézy ze zadání parametrický test. Jelikož není splněn předpoklad shody rozptylů obou náhodných výběrů, zvolíme na test hypotézy ze zadání Welchův dvouvýběrový t -test. Zadaným úkolem je zjistit, zda existuje statisticky významný rozdíl mezi délkou trupu mužů a žen. Toto tvrzení je zněním alternativní hypotézy, neboť v nulové hypotéze nikdy nepředpokládáme nerovnost.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : Střední hodnota délky trupu mužů a střední hodnota délky trupu žen jsou shodné.

H_1 : Střední hodnota délky trupu mužů a střední hodnota délky trupu žen nejsou shodné.

- **matematická formulace** nulové a alternativní hypotézy

$H_0 : \mu_1 = \mu_2 \rightarrow \mu_1 - \mu_2 = \mu_0$, kde $\mu_0 = 0$

$H_1 : \mu_1 \neq \mu_2 \rightarrow \mu_1 - \mu_2 \neq \mu_0$, kde $\mu_0 = 0$

(oboustranná alternativa)

2. Volba hladiny významnosti

- Hladinu významnosti volíme ze zadání $\alpha = 0.01$.

3. Testování kritickým oborem

- Testovací statistika

$$\begin{aligned}
 T_W &= \frac{(M_1 - M_2) - \mu_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \\
 &= \frac{(457.1867 - 423.17) - 0}{\sqrt{\frac{27.15555^2}{75} + \frac{34.02288^2}{100}}} \\
 &= \frac{34.0167}{\sqrt{21.40788}} \\
 &= \frac{34.0167}{4.626865} = 7.35199 \doteq 7.35199
 \end{aligned}$$

```

241 alpha <- 0.01
242 mu0 <- 0
243 m1 <- mean(tru.Lm)
244 m2 <- mean(tru.Lf)
245 s1 <- sd(tru.Lm)
246 s2 <- sd(tru.Lf)
247
248 tw <- ((m1 - m2) - mu0) / (sqrt(s1 ^ 2 / n1 + s2 ^ 2 / n2)) # 7.351989

```

- Welchova aproximace počtu stupňů volnosti

$$\begin{aligned}
 df &= \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1-1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2-1}} \\
 &= \frac{\left(\frac{27.15555^2}{75} + \frac{34.02288^2}{100}\right)^2}{\frac{\left(\frac{34.02288^2}{75}\right)^2}{75-1} + \frac{\left(\frac{34.02288^2}{100}\right)^2}{100-1}} \\
 &= \frac{(9.832319 + 11.57556)^2}{\frac{9.832319^2}{74} + \frac{11.57556^2}{99}} \\
 &= \frac{458.2973}{2.659883} \\
 &= 172.2998
 \end{aligned}$$

```

249 df <- (s1 ^ 2 / n1 + s2 ^ 2 / n2)^2 / ((s1 ^ 2 / n1) ^ 2 / (n1 - 1) + (s2 ^ 2 / n2) ^ 2 / (n2 - 1)) # 172.2998

```

Kritický obor

$$\begin{aligned}
 W &= (-\infty; t_{df}(\alpha/2)) \cup (t_{df}(1 - \alpha/2); \infty) \\
 &= (-\infty; t_{172.2998}(0.005)) \cup (t_{172.2998}(0.995); \infty) \\
 &= (-\infty; -2.604664) \cup (2.604664; \infty)
 \end{aligned}$$

```

250 qt(alpha/2, df) # -2.604664
251 qt(1-alpha/2, df) # 2.604664

```

- **Závěr testování**

Protože realizace testovací statistiky $t_W = 7.35199$ náleží do kritického oboru, tj. $t_W \in W$, H_0 zamítáme na hladině významnosti $\alpha = 0.01$.

4. Testování intervalem spolehlivosti

- **Interval spolehlivosti**

$$\begin{aligned}
 (d, h) &= \left(m_1 - m_2 - \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} t_{df}(1 - \alpha/2); m_1 - m_2 - \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} t_{df}(\alpha/2) \right) \\
 &= ((457.1867 - 423.17) - \sqrt{\frac{27.15555^2}{75} + \frac{34.02288^2}{100}} t_{172.2998}(0.995); (457.1867 - 423.17) - \sqrt{\frac{27.15555^2}{75} + \frac{34.02288^2}{100}} t_{172.2998}(0.005)) \\
 &= (34.0167 - 4.626865 \times 2.604664; 34.0167 - 4.626865 \times (-2.604664)) \\
 &= (21.96527; 46.06813) \doteq (21.9653; 46.06813)
 \end{aligned}$$

```

252 dh <- (m1 - m2) - sqrt(s1^2 / n1 + s2^2 / n2) * qt(1 - alpha / 2, n1 + n2 - 2) #
      21.96527
253 hh <- (m1 - m2) + sqrt(s1^2 / n1 + s2^2 / n2) * qt(alpha / 2, n1 + n2 - 2) # 46.06755

```

- **Závěr testování**

Protože $\mu_0 = 0$ nenáleží do Waldova 99% empirického oboustranného intervalu spolehlivosti, tj. $\mu_0 = 0 \notin IS$, H_0 zamítáme na hladině významnosti $\alpha = 0.01$.

5. Testování p -hodnotou

- **p -hodnota**

$$\begin{aligned}
 p\text{-hodnota} &= 2 \min\{\Pr(T_W \leq t_W), \Pr(T_W > t_W)\} \\
 &= 2 \min\{\Pr(T_W \leq 7.35199), 1 - \Pr(T_W \leq 7.35199)\} \\
 &= 2 \min\{0.9999999, 3.774536 \times 10^{-12}\} \\
 &= 2 \times (3.774536 \times 10^{-12}) \\
 &= 7.549072 \times 10^{-12} \doteq 1.5490 \times 10^{-11}
 \end{aligned}$$

```

254 p.val <- 2 * min(pt(tw, df), 1 - pt(tw, df)) # 7.549072e-12

```

- **Závěr testování**

Protože p -hodnota $= 1.5490 \times 10^{-11}$ je menší než $\alpha = 0.01$, H_0 zamítáme na hladině významnosti $\alpha = 0.01$.

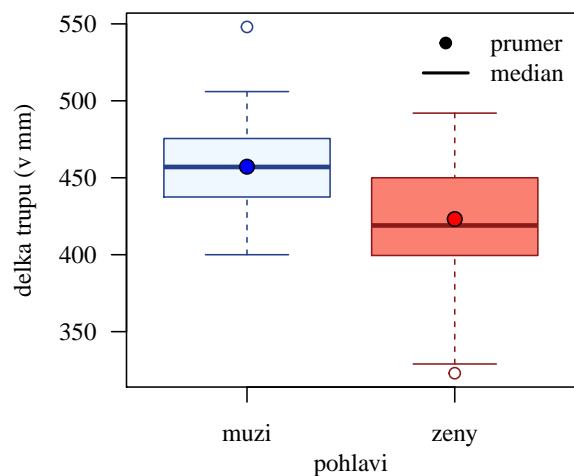
6. Grafická vizualizace výsledků testování

Rozdíl mezi středními hodnotami populací mužů a žen vizualizujeme pomocí krabicového diagramu (viz obrázek 30).

7. Interpretace výsledků:

Na základě všech tří způsobů testování zamítáme hypotézu o shodě středních hodnot. Mezi délkou trupu mužů a žen existuje statisticky významný rozdíl.

Poznámka: K otestování nulové hypotézy o rozdílu středních hodnot můžeme využít funkci `t.test()` s argumenty `conf.level = 0.99` pro volbu hladiny významnosti $\alpha = 0.01$, `alternative = 'two.sided'` pro volbu oboustranné alternativní hypotézy a `var.equal = F` pro volbu Welchova dvouvýběrového t -testu.



Obrázek 30: Krabicový diagram délky trupu mužů a žen

```
255 t.test(tru.Lm, tru.Lf, conf.level = 0.95, alternative = 'two.sided', var.equal = F)
```

```

Welch Two Sample t-test
data:  tru.Lm and tru.Lf
t = 7.352, df = 172.3, p-value = 7.549e-12
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 24.88403 43.14930
sample estimates:
mean of x mean of y
 457.1867  423.1700

```

Součástí výstupu je hodnota testovací statistiky $t = 7.352$, Welchova aproximace počtu stupňů volnosti Studentova rozdělení $\text{num df} = 172.3$, hranice intervalu spolehlivosti 24.88403 a 43.14930 a p -hodnota $p\text{-value} = 7.549e-12$. Jediné, co musíme stanovit zvlášť, jsou dolní a horní hranice kritického oboru. ★

Příklad 10.12. Test o rozdílu středních hodnot

Mějme datový soubor 19-more-samples-correlations-skull.txt a proměnnou nose.B popisující šířku nosu v mm (viz sekce ??). Předpokládejme, že náhodná veličina X , popisující šířku nosu bantuské populace, pochází z normálního rozdělení $N(\mu_1, \sigma_1^2)$, a že náhodná veličina Y , popisující šířku nosu čínské populace, pochází z normálního rozdělení $N(\mu_2, \sigma_2^2)$. Na hladině významnosti $\alpha = 0.05$ testujte nulovou hypotézu o tom, že šířka nosu čínské populace je menší nebo rovna šířce nosu bantuské populace.

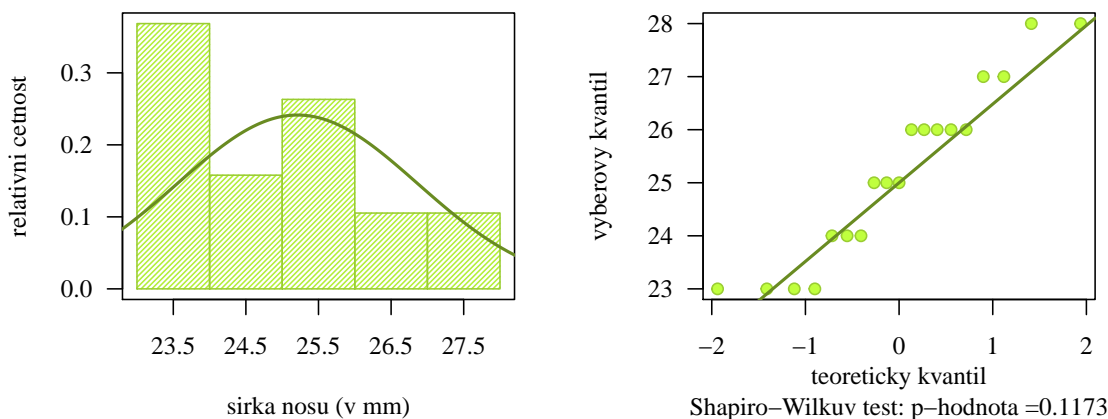
Řešení příkladu 10.12

Příkazem read.delim() načteme datový soubor a příkazem na.omit() z něj odtráhneme NA hodnoty. Pomocí operátoru [] vybereme z datové tabulky naměřené hodnoty šířky nosu (nose.B) čínské populace (pop == 'cin'), resp. bantuské populace (pop == 'ban') a zjistíme rozsahy obou náhodných výběrů.

```
267 data <- read.delim('00-Data//19-more-samples-correlations-skull.txt')
268 data <- na.omit(data)
269 #head(data, n = 3)
270 nose.Bc <- data[data$pop == 'cin', 'nose.B']
271 nose.Bb <- data[data$pop == 'ban', 'nose.B']
272
273 n1 <- length(nose.Bc) # 19
274 n2 <- length(nose.Bb) # 14
```

Datový soubor obsahuje naměřené hodnoty šířky nosu u 19 jedinců čínské populace a 14 jedinců bantuské populace. Řešení příkladu vede na test o rozdílu středních hodnot.

Nejprve otestujeme normalitu obou náhodných výběrů. Na hladině významnosti α testujeme hypotézu H_{01} : *Náhodný výběr naměřených šířek nosu čínské populace pochází z normálního rozdělení.* oproti alternativní hypotéze H_{11} : *Náhodný výběr naměřených šířek nosu čínské populace nepochází z normálního rozdělení.* Vzhledem k rozsahu náhodného výběru použijeme k otestování normality Shapiro-Wilkův test, kvantilový graf a histogram (viz obrázek 31). Datový soubor rozdělíme do pěti ekvidistantních intervalů s šířkou 1 mm prostřednictvím stanovených hranic 23, 24, ..., 28.

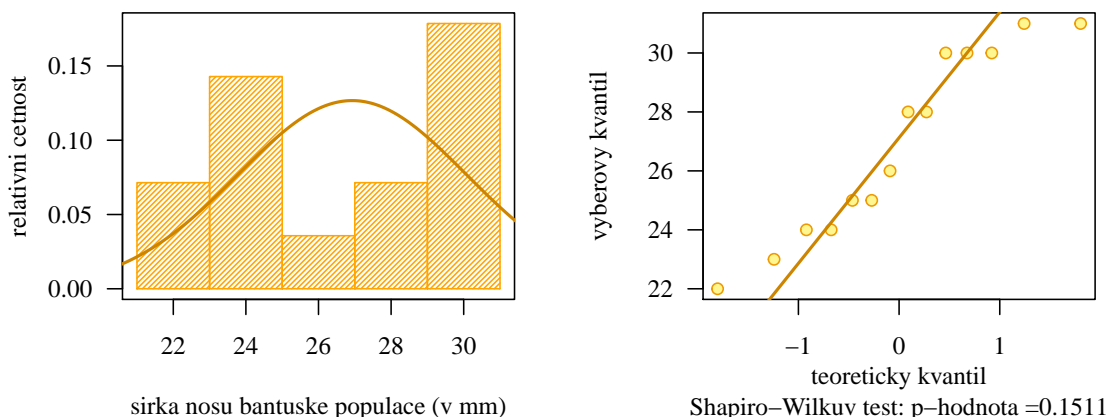


Obrázek 31: Histogram a kvantilový diagram šířky nosu mužů čínské populace

Protože p -hodnota = 0.1173 je větší než 0.05, nulovou hypotézu nezamítáme na hladině významnosti $\alpha = 0.05$. Náhodný výběr naměřených šířek nosu čínské populace pochází z normálního rozdělení. Vykreslený kvantilový graf výsledek testování podporuje, histogram víceméně také.

Analogicky testujeme hypotézu H_{02} : *Náhodný výběr naměřených šířek nosu bantuské populace pochází z normálního rozdělení.* oproti alternativní hypotéze H_{12} : *Náhodný výběr naměřených šířek nosu bantuské populace nepochází z normálního rozdělení.* K otestování předpokladu normality použijeme Shapiro-Wilkův test, kvantilový diagram a

histogram. Datový soubor rozdělíme do pěti ekvidistatních intervalů s šířkou 2 mm prostřednictvím stanovených hranic 21, 23, ..., 31 (viz obrázek 32).



Obrázek 32: Histogram a kvantilový diagram šířky nosu mužů bantuské populace

Protože p -hodnota = 0.1511 je větší než 0.05, nulovou hypotézu nezamítáme na hladině významnosti $\alpha = 0.05$. Náhodný výběr šířek nosu bantuské populace pochází z normálního rozdělení. Ačkoli pohled na histogram nás o normalitě dat příliš nepřesvědčil, nemůžeme případnou normalitu naměřených hodnot navrhnout kvůli malému rozsahu náhodného výběru. Kvantilový graf normalitu naměřených hodnot podporuje. Proto se kloníme k závěru Shapiro-Wilkova testu a data považujeme za normálně rozdělená.

Nyní zbývá rozhodnout o splnění nebo nesplnění předpokladu o shodě rozptylů obou náhodných výběrů. Na hladině významnosti $\alpha = 0.05$ testujeme hypotézu $H_0 : \sigma_1^2/\sigma_2^2 = 1$ (rozptyly σ_1^2 a σ_2^2 jsou shodné) oproti alternativní hypotéze $H_1 : \sigma_1^2/\sigma_2^2 \neq 1$ (rozptyly σ_1^2 a σ_2^2 nejsou shodné). Nulovou hypotézu otestujeme pomocí funkce `var.test()` na základě p -hodnoty (viz příklad 10.2).

```
275 var.test(nose.Bc, nose.Bb, alternative = 'two.sided')$p.val
```

```
[1] 0.01257845
```

276

Jelikož p -hodnota = 0.01258 je menší než 0.05, nulovou hypotézu o shodě rozptylů zamítáme na hladině významnosti $\alpha = 0.05$. Testování tedy prokázalo statisticky významný rozdíl v rozptylech obou náhodných výběrů.

Náhodné výběry pochází z normálních rozdělení s vzájemně odlišnými rozptyly, proto na otestování hypotézy ze zadání použijeme Welchův dvouvýběrový t -test o rozdílu středních hodnot $\mu_1 - \mu_2$. V zadání příkladu máme výslovně definované znění hypotézy H_0 , zbývá nám tedy dodefinovat znění alternativní hypotézy H_1 .

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy
 H_0 : Střední hodnota šířky nosu čínské populace je menší nebo rovna střední hodnotě šířky nosu bantuské populace.
 H_1 : Střední hodnota šířky nosu čínské populace je větší než střední hodnota šířky nosu bantuské populace.
- **matematická formulace** nulové a alternativní hypotézy
 $H_0 : \mu_1 \leq \mu_2 \rightarrow \mu_1 - \mu_2 \leq \mu_0$, kde $\mu_0 = 0$
 $H_1 : \mu_1 > \mu_2 \rightarrow \mu_1 - \mu_2 > \mu_0$, kde $\mu_0 = 0$
 (pravostranná alternativa)

2. Volba hladiny významnosti

- Hladinu významnosti volíme ze zadání $\alpha = 0.05$.

3. Testování kritickým oborem

- Testovací statistika

$$\begin{aligned}
 T_W &= \frac{(M_1 - M_2) - \mu_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \\
 &= \frac{(25.21053 - 26.92857) - 0}{\sqrt{\frac{1.652572^2}{19} + \frac{3.149219^2}{14}}} \\
 &= \frac{-1.71804}{\sqrt{0.8521351}} \\
 &= \frac{-1.71804}{0.9231117} = -1.86114 \doteq -1.8611
 \end{aligned}$$

```

277 alpha <- 0.05
278 mu0 <- 0
279 m1 <- mean(nose.Bc)
280 m2 <- mean(nose.Bb)
281 s1 <- sd(nose.Bc)
282 s2 <- sd(nose.Bb)
283
284 tw <- ((m1 - m2) - mu0) / (sqrt(s1 ^ 2 / n1 + s2 ^ 2 / n2)) # -1.861145

```

- Welchova aproximace počtu stupňů volnosti

$$\begin{aligned}
 df &= \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1-1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2-1}} \\
 &= \frac{\left(\frac{1.652572^2}{19} + \frac{3.149219^2}{14}\right)^2}{\frac{\left(\frac{1.652572^2}{19}\right)^2}{19-1} + \frac{\left(\frac{3.149219^2}{14}\right)^2}{14-1}} \\
 &= \frac{(0.1437365 + 0.7083986)^2}{\frac{0.1437365^2}{18} + \frac{0.7083986^2}{13}} \\
 &= \frac{0.7261342}{0.03974999} \\
 &= 18.26753 \doteq 18.2675
 \end{aligned}$$

```

285 df <- (s1 ^ 2 / n1 + s2 ^ 2 / n2)^2 / ((s1 ^ 2 / n1) ^ 2 / (n1 - 1) + (s2 ^ 2 / n2) ^ 2 / (n2 - 1)) # 18.26753

```

- Kritický obor

$$\begin{aligned}
 W &= \langle t_{df}(1 - \alpha); \infty \rangle \\
 &= \langle t_{18.26753}(1 - 0.05); \infty \rangle \\
 &= \langle 1.732689; \infty \rangle
 \end{aligned}$$

- **Závěr testování**

Protože realizace testovací statistiky $t_W = -1.861145$ nenáleží do kritického oboru, tj. $t_W \notin W$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

4. Testování intervalem spolehlivosti

- **Interval spolehlivosti**

$$\begin{aligned} (d, \infty) &= \left(m_1 - m_2 - \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} t_{df}(1 - \alpha); \infty \right) \\ &= \left((25.21053 - 26.92857) - \sqrt{\frac{1.652572^2}{19} + \frac{3.149219^2}{14}} t_{18.26753}(0.95); \infty \right) \\ &= (-1.71804 - \times 0.9231117 \times 1.732689; \infty) \\ &= (-3.31751; \infty) \doteq (-3.3175; \infty) \end{aligned}$$

```
287 DH <- (m1 - m2) - sqrt(s1^2 / n1 + s2^2 / n2) * qt(1 - alpha, df) # -3.317511
```

- **Závěr testování**

Protože $\mu_0 = 0$ náleží do Waldova 95% empirického levostranného intervalu spolehlivosti, tj. $\mu_0 = 0 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

5. Testování p -hodnotou

- **p -hodnota**

$$\begin{aligned} p\text{-hodnota} &= \Pr(T_W \geq t_W) = 1 - \Pr(T_w < t_W) \\ &= 1 - \Pr(T_W \leq -1.861145) \\ &= 0.960552 \doteq 0.9606 \end{aligned}$$

```
288 p.val <- 1 - pt(tw, df) # 0.960552
```

- **Závěr testování**

Protože p -hodnota = 0.9606 je větší než $\alpha = 0.05$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

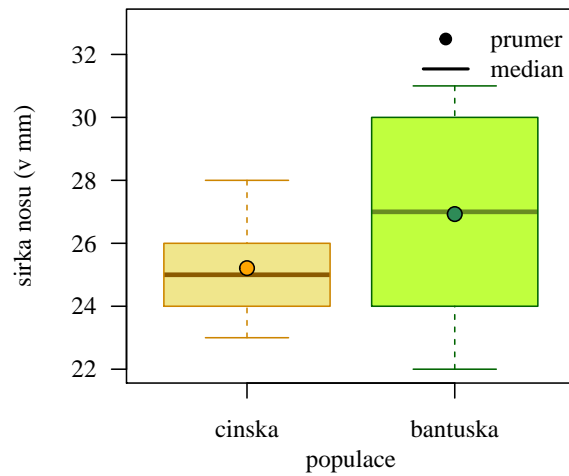
6. Grafická vizualizace výsledků testování

Vhodný grafem porovnávajícím střední hodnoty obou náhodných výběrů je krabicový diagram viz obrázek 33).

7. Interpretace výsledků:

Na základě všech tří způsobů testování nezamítáme hypotézu o tom, že střední hodnota šířky nosu čínské populace je menší nebo rovna střední hodnotě šířky nosu bantuské populace. Nemáme tedy dostatek indicií k zamítnutí tvrzení, že šířka nosu čínské populace je statisticky významně menší nebo rovná šířce nosu bantuské populace.

Poznámka: K otestování nulové hypotézy o rozdílu středních hodnot můžeme využít funkci `t.test()` s argumenty `conf.level = 0.95` pro volbu hladiny významnosti $\alpha = 0.05$, `alternative = 'greater'` pro volbu pravostranné alternativní hypotézy a `var.equal = F` pro volbu Welchova dvouvýběrového t -testu.



Obrázek 33: Krabicový diagram šířky nosu mužů čínské a bantuské populace

```
289 t.test(nose.Bb, nose.Bc, conf.level = 0.95, alternative = 'greater', var.equal = F)
```

```

Welch Two Sample t-test
data: nose.Bb and nose.Bc
t = 1.8611, df = 18.268, p-value = 0.03945
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 0.1185797      Inf
sample estimates:
mean of x mean of y
 26.92857  25.21053

```

290
291
292
293
294
295
296
297
298
299
300



10.4 Test o rozdílu korelačních koeficientů $\rho_1 - \rho_2$

Nechť $(X_{11}, Y_{11})^T \dots (X_{1n_1}, Y_{1n_1})^T$ je náhodný výběr z $N_2(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$ a $(X_{21}, Y_{21})^T, \dots (X_{2n_2}, Y_{2n_2})^T$ je na něm nezávislý náhodný výběr z $N_2(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$. Nechť ρ_0 je konstanta. Na hladině významnosti α testujeme jednu z následujících tří hypotéz oproti příslušné alternativní hypotéze.

$$\begin{array}{lll} H_{01} : \rho_1 - \rho_2 = \rho_0 & \text{oproti} & H_{11} : \rho_1 - \rho_2 \neq \rho_0 \quad (\text{oboustranná alt.}) \\ H_{02} : \rho_1 - \rho_2 \leq \rho_0 & \text{oproti} & H_{12} : \rho_1 - \rho_2 > \rho_0 \quad (\text{pravostranná alt.}) \\ H_{03} : \rho_1 - \rho_2 \geq \rho_0 & \text{oproti} & H_{13} : \rho_1 - \rho_2 < \rho_0 \quad (\text{levostranná alt.}) \end{array}$$

Test nazýváme dvouvýběrovým Z -testem o rozdílu korelačních koeficientů $\rho_1 - \rho_2$. Testovací statistika má tvar

$$Z_W = \frac{Z_1 - Z_2 - \xi_0}{\sqrt{\frac{1}{n_1-3} + \frac{1}{n_2-3}}}, \quad (10.5)$$

kde $Z_1 = \frac{1}{2} \ln \frac{1+R_1}{1-R_1}$, $Z_2 = \frac{1}{2} \ln \frac{1+R_2}{1-R_2}$ jsou Fisherovy Z -transformace výběrových korelačních koeficientů R_1 a R_2 a $\xi_0 = \frac{1}{2} \ln \frac{1+\rho_0}{1-\rho_0}$ je Fisherova Z -transformace konstanty ρ_0 z nulové hypotézy, n_1 je rozsah prvního náhodného výběru a n_2 je rozsah druhého náhodného výběru. Testovací statistika Z_W pochází ze standardizovaného normálního rozdělení, tj.

$$Z_W = \frac{Z_1 - Z_2 - \xi_0}{\sqrt{\frac{1}{n_1-3} + \frac{1}{n_2-3}}} \sim N(0, 1).$$

Kritický obor podle zvolené alternativní hypotézy má tvar

$$\begin{array}{ll} H_{11} : \rho_1 - \rho_2 \neq \rho_0 & W = (-\infty; u_{\alpha/2}) \cup (u_{1-\alpha/2}; \infty) \\ H_{12} : \rho_1 - \rho_2 > \rho_0 & W = (u_{1-\alpha}; \infty) \\ H_{13} : \rho_1 - \rho_2 < \rho_0 & W = (-\infty; u_{\alpha}) \end{array}$$

kde $u_{\alpha/2}$, $u_{1-\alpha/2}$, u_{α} , $u_{1-\alpha}$ jsou kvantily standardizovaného normálního rozdělení, jejichž hodnoty získáme pomocí \mathbb{R} a implementované funkce `qnorm()`.

Interval spolehlivosti má podle zvolené alternativní hypotézy jeden z následujících tvarů

$$\begin{array}{ll} H_{11} : \rho_1 - \rho_2 \neq \rho_0 & (d, h) = (\tanh(z_1 - z_2 - s_g u_{1-\alpha/2}); \tanh(z_1 - z_2 - s_g u_{\alpha/2})) \\ H_{12} : \rho_1 - \rho_2 > \rho_0 & (d, 2) = (\tanh(z_1 - z_2 - s_g u_{1-\alpha}); 2) \\ H_{13} : \rho_1 - \rho_2 < \rho_0 & (-2, h) = (-2; \tanh(z_1 - z_2 - s_g u_{\alpha})) \end{array}$$

kde $s_g = \sqrt{\frac{1}{n_1-3} + \frac{1}{n_2-3}}$ a \tanh je hyperbolický tangens, jehož hodnotu získáme pomocí \mathbb{R} a implementované funkce `tanh()`.

Poznámka: Dá se ukázat, že mezi korelačním koeficientem ρ_0 a ξ_0 platí vztah $\rho_0 = \tanh(\xi_0)$, který je inverzí ke vztahu $\xi_0 = \frac{1}{2} \ln \frac{1+\rho_0}{1-\rho_0}$. Všimněme si, že $z_1 - z_2 - s_g u_{1-\alpha/2}$ a $z_1 - z_2 - s_g u_{\alpha/2}$ jsou dolní a horní hranice intervalu spolehlivosti pro Z -transformaci ξ_0 , neboť z_1 i z_2 jsou Z -transformacemi korelačních koeficientů. Hyperbolický tangens potom funguje jako zpětná transformace, která převede hranice intervalu spolehlivosti pro ξ_0 zpátky na hranice intervalu spolehlivosti pro korelační koeficient ρ_0 .

Poznámka: Protože parametry ρ_1 i ρ_2 jsou korelační koeficienty, platí, že $\rho_1 \in \langle -1; 1 \rangle$ a $\rho_2 \in \langle -1; 1 \rangle$. Rozdíl $\rho_1 - \rho_2$ nabývá tedy hodnoty z intervalu $\langle -2; 2 \rangle$. Proto levostranný interval spolehlivosti omezíme shora hodnotou 2, namísto nekonečnem, a pravostranný interval spolehlivosti omezíme zdola hodnotou -2, namísto mínus nekonečnem.

p -hodnota má v závislosti na zvolené alternativní hypotéze jeden z následujících tvarů

$$\begin{array}{ll} H_{11} : \rho_1 - \rho_2 \neq \rho_0 & p\text{-hodnota} = 2 \min\{\Pr(Z_W \leq z_W), \Pr(Z_W > z_W)\} \\ H_{12} : \rho_1 - \rho_2 > \rho_0 & p\text{-hodnota} = \Pr(Z_W > z_W) = 1 - \Pr(Z_W \leq z_W) \\ H_{13} : \rho_1 - \rho_2 < \rho_0 & p\text{-hodnota} = \Pr(Z_W \leq z_W) \end{array}$$

kde Z_W je náhodná veličina, z_W je realizace testovací statistiky Z_W (viz vzorec 10.5), tedy konkrétní číslo, a $\Pr(Z_W \leq z_W)$ je distribuční funkce standardizovaného normálního rozdělení, jejíž hodnotu získáme pomocí \mathcal{R} a implementované funkce `pnorm()`.

Příklad 10.13. Test o rozdílu dvou korelačních koeficientů $\rho_1 - \rho_2$

Mějme datový soubor `19-more-samples-correlations-skull.txt`, proměnnou `nose.H` popisující délku nosu a proměnnou `nose.B` popisující šířku nosu (viz sekce ??). Na hladině významnosti $\alpha = 0.05$ zjistěte, zda se korelační koeficient délky a šířky nosu mužů bantuské populace významně liší od korelačního koeficientu délky a šířky nosu mužů peruánské populace.

Řešení příkladu 10.13

Datový soubor načteme příkazem `read.delim()` a odstraníme z něj chybějící hodnoty příkazem `na.omit()`. Operátorem `[]` vybereme z tabulky nejprve údaje o délce nosu (`nose.H`), resp. šířce nosu `nose.B` u mužů bantuské populace `pop == 'ban'` a následně údaje o délce nosu (`nose.H`), resp. šířce nosu `nose.B` u mužů peruánské populace `pop == 'per'`. Pomocí příkazu `length()` dále zjistíme rozsahy obou náhodných výběrů a pomocí příkazu `range()` zjistíme rozsahy naměřených hodnot každé proměnné. Datový soubor obsahuje údaje o výšce a šířce nosu u 14 mužů bantuské

```
301 data <- read.delim('00-Data//19-more-samples-correlations-skull.txt')
302 data <- na.omit(data)
303 nose.HB <- data[data$pop == 'ban', 'nose.H']
304 nose.BB <- data[data$pop == 'ban', 'nose.B']
305 nose.HP <- data[data$pop == 'per', 'nose.H']
306 nose.BP <- data[data$pop == 'per', 'nose.B']
307
308 n1 <- length(nose.HB) # 14
309 n2 <- length(nose.HP) # 46
310 range(nose.HB) # 41-58
311 range(nose.BB) # 22-31
312 range(nose.HP) # 44-57
313 range(nose.BP) # 19-26
```

populace, přičemž naměřené výšky nosu nabývají hodnot v rozmezí 41–58 mm, naměřené šířky nosu nabývají hodnot v rozmezí 22–31 mm. Soubor dále obsahuje údaje o výšce a šířce nosu u 46 mužů peruánské populace, přičemž naměřené výšky nosu nabývají hodnot v rozmezí 44–57 mm a naměřené šířky nosu nabývají hodnot v rozmezí 19–26 mm.

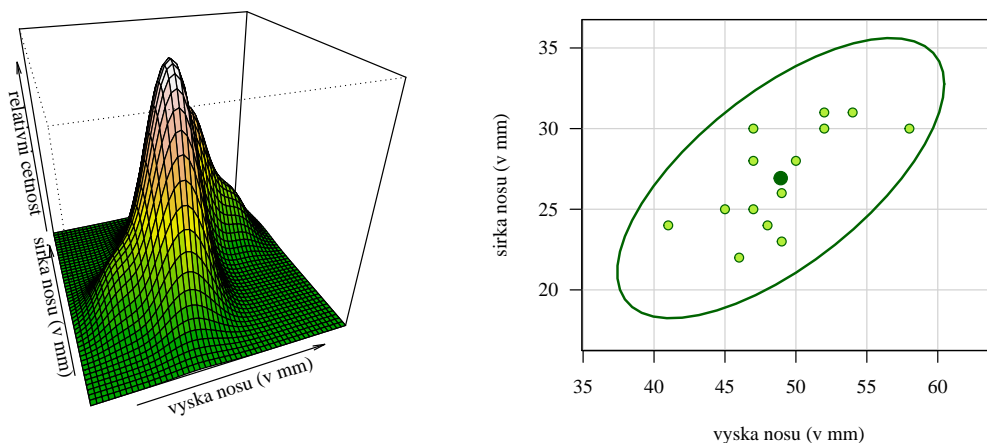
Abychom mohli hypotézu ze zadání otestovat pomocí parametrického testu, musí oba datové soubory splňovat předpoklad dvourozměrné normality. Před samotným testováním hypotézy uvedené v zadání je potřeba ověřit, zda oba výběry tento předpoklad splňují. Závěr o dvourozměrné normalitě obou náhodných výběrů stanovíme na základě Mardiova testu v kombinaci s grafickou vizualizací dat pomocí 3D grafu a tečkového diagramu superponovaného 95 % elipsou spolehlivosti, analogicky, jako je uvedeno v sekci ??). Hladinu významnosti zvolíme $\alpha = 0.05$. Začneme s náhodným výběrem mužů bantuské populace.

```
314 MVN::mvn(cbind(nose.HB, nose.BB), mvnTest = 'mardia')$multivariateNormality
315 # sikkmost: 0.9610029 # spicatost: 0.4619885
```

Protože p -hodnota testu o nevýznamnosti koeficientu šikmosti, tj. 0.9610, je větší než 0.05, hypotézu o nevýznamnosti koeficientu šikmosti nezamítáme na hladině významnosti $\alpha = 0.05$. Dále protože p -hodnota testu o nevýznamnosti koeficientu špičatosti, tj. 0.3198, je větší než 0.05, nezamítáme hypotézu o nevýznamnosti koeficientu špičatosti. Protože náhodný výběr nevykazuje statisticky významné známky zešikmení ani zešpičatění, nezamítáme na hladině významnosti $\alpha = 0.05$ hypotézu o dvourozměrné normalitě náhodného výběru výšek a šířek nosu u mužů bantuské populace. Ke stejnému závěru bychom došli také použitím Henze-Zirklerova testu (p -hodnota = 0.7110 > 0.05) i Roystonova testu (p -hodnota = 0.3198 > 0.05).

Nyní se podíváme na grafickou vizualizaci náhodného výběru (viz graf 34). Z 3D grafu i tečkového diagramu je zřejmá dvourozměrná normalita náhodného výběru. 3D graf nám ukazuje kopcovitý tvar náhodného výběru mužů

bantuské populace. Při pohledu na tečkový diagram vidíme, že všechny body leží uvnitř 95 % elipsy spolehlivosti.



Obrázek 34: 3D graf a tečkový diagram s 95% elipsou spolehlivosti pro výšku nosu a šířku nosu mužů bantuské populace (v mm)

Dále se zaměříme na ověření předpoklad normality výšky a šířky nosu u mužů peruánské populace.

```
316 MVN::mvn(cbind(nose.HP, nose.BP), mvnTest = 'mardia')$multivariateNormality # šikmost:
0.409373 # špičatost: 0.2925019
```

Jelikož p -hodnota testu o nevýznamnosti koeficientu šikmosti, tj. 0.4094, je větší než 0.05, hypotézu o nevýznamnosti koeficientu šikmosti nezamítáme na hladině významnosti $\alpha = 0.05$. Dále jelikož p -hodnota testu o nevýznamnosti koeficientu špičatosti, tj. 0.2925, je větší než 0.05, nezamítáme hypotézu o nevýznamnosti koeficientu špičatosti. Protože náhodný výběr nevykazuje statisticky významné zešikmení ani zešpičatění, nezamítáme hypotézu o dvourozměrné normalitě náhodného výběru výšek a šířek nosu u mužů bantuské populace na hladině významnosti $\alpha = 0.05$. Stejný závěr bychom získali také použitím Henze-Zirklerova testu (p -hodnota = 0.1306 > 0.05). Na základě Roystonova testu bychom hypotézu o dvourozměrné normalitě náhodného výběru hraničně zamítli (p -hodnota = 0.4543 < 0.05). Podíváme se tedy, co nám na dvourozměrnou normalitu náhodného výběru poví grafická vizualizace (viz obrázek 35). 3D graf i tečkový diagram ukazují na dvourozměrné normální rozdělení s jedním odlehlym pozorováním, které se nevešlo do 95 % elipsy spolehlivosti. Vzhledem k tomu, že rozsah náhodného výběru je 46, požadujeme, aby elipsa spolehlivosti pokrývala alespoň 44 bodů. Zbylé dva body mohou ležet mimo elipsu spolehlivosti. V našem případě leží mimo elipsu spolehlivosti pouze jeden bod, proto se kloníme k závěru Mardia a Henze-Zirklerova testu o dvourozměrné normalitě náhodného výběru výšek a šířek nosu u mužů peruánské populace.

Protože oba výběry pochází z dvourozměrného normálního rozdělení, můžeme hypotézu ze zadání otestovat parametrickým testem o rozdílu korelačních koeficientů $\rho_1 - \rho_2$. Testování provedeme v posloupnosti šesti kroků.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

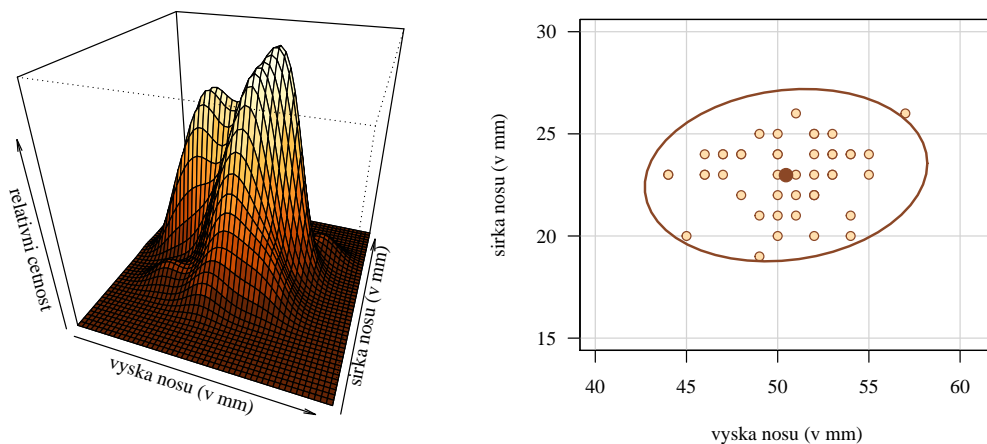
H_0 : Korelační koeficient proměnných výška nosu a šířka nosu u mužů bantuské a korelační koeficient proměnných výška nosu a šířka nosu u mužů peruánské populace jsou shodné.

H_1 : Korelační koeficient proměnných výška nosu a šířka nosu u mužů bantuské a korelační koeficient proměnných výška nosu a šířka nosu u mužů peruánské populace nesou shodné.

- **matematická formulace** nulové a alternativní hypotézy

$H_0 : \rho_1 = \rho_2 \rightarrow \rho_1 - \rho_2 = \rho_0$, kde $\rho_0 = 0$

$H_1 : \rho_1 \neq \rho_2 \rightarrow \rho_1 - \rho_2 \neq \rho_0$, kde $\rho_0 = 0$
(oboustranná alternativa)



Obrázek 35: 3D graf a tečkový diagram s 95% elipsou spolehlivosti pro výšku nosu a šířku nosu mužů peruánské populace (v mm)

2. Volba hladiny významnosti

- Hladinu významnosti volíme v souladu se zadáním $\alpha = 0.05$.

3. Testování kritickým oborem

- Fisherovy Z -transformace výběrových korelačních koeficientů

K výpočtu Fisherových Z -transformací potřebujeme znát hodnoty výběrových korelačních koeficientů. Tyto hodnoty získáme pomocí příkazu `cor()` s nastaveným argumentem `method == 'pearson'`. Výběrový korelační koeficient výšky a šířky nosu mužů bantuské populace $r_1 = 0.6719$, výběrový korelační koeficient výšky a šířky nosu peruánské populace vyšel $r_2 = 0.1371$. Nyní již můžeme vypočítat Fisherovy Z -transformace těchto korelačních koeficientů i Fisherovu Z -transformaci konstanty $\rho_0 = 0$ z nulové hypotézy.

$$\begin{aligned}
 Z_1 &= \frac{1}{2} \ln \frac{1 + R_1}{1 - R_1} \\
 &= \frac{1}{2} \ln \frac{1 + 0.6719065}{1 - 0.6719065} = \frac{1}{2} \ln 5.095816 \\
 &= 0.5 \times 1.62842 = 0.81421 \doteq 0.8142
 \end{aligned}$$

$$\begin{aligned}
 Z_2 &= \frac{1}{2} \ln \frac{1 + R_2}{1 - R_2} \\
 &= \frac{1}{2} \ln \frac{1 + 0.1370691}{1 - 0.1370691} = \frac{1}{2} \ln 1.317683 \\
 &= 0.5 \times 0.2758746 = 0.1379373
 \end{aligned}$$

$$\begin{aligned}
 \xi_0 &= \frac{1}{2} \ln \frac{1 + \rho_0}{1 - \rho_0} \\
 &= \frac{1}{2} \ln \frac{1 + 0}{1 - 0} = \frac{1}{2} \ln 1 \\
 &= 0.5 \times 0 = 0
 \end{aligned}$$


```

317 alpha <- 0.05
318 ksi0 <- 0
319
320 r1 <- cor(nose.HB, nose.BB) # 0.6719065
321 r2 <- cor(nose.HP, nose.BP) # 0.1370691
322
323 z1 <- 1 / 2 * log((1 + r1) / (1 - r1)) # 0.8142106
324 z2 <- 1 / 2 * log((1 + r2) / (1 - r2)) # 0.1379373

```

- Testovací statistika

$$\begin{aligned}
Z_W &= \frac{Z_1 - Z_2 - \xi_0}{\sqrt{\frac{1}{n_1-3} + \frac{1}{n_2-3}}} \\
&= \frac{0.8142106 - 0.1379373 - 0}{\sqrt{\frac{1}{14-3} + \frac{1}{46-3}}} \\
&= \frac{0.6762733}{\sqrt{0.09090909 + 0.02325581}} \\
&= \frac{0.6762733}{\sqrt{0.1141649}} \\
&= \frac{0.6762733}{0.337883} \\
&= 2.001501
\end{aligned}$$

```

325 zw <- (z1 - z2 - ksi0) / sqrt(1 / (n1 - 3) + 1 / (n2 - 3)) # 2.001502

```

- Kritický obor

$$\begin{aligned}
W &= (-\infty; u_{\alpha/2}) \cup (u_{1-\alpha/2}; \infty) \\
&= (-\infty; u_{0.05/2}) \cup (u_{1-0.05/2}; \infty) \\
&= (-\infty; u_{0.025}) \cup (u_{0.975}; \infty) \\
&= (-\infty; -1.959964) \cup (1.959964; \infty)
\end{aligned}$$

```

326 qnorm(alpha / 2) # -1.959964
327 qnorm(1 - alpha / 2) # 1.959964

```

- Závěr testování

Protože realizace testovací statistiky $z_W = 2.001501$ náleží do kritického oboru, tj. $z_W \in W$, H_0 zamítáme na hladině významnosti $\alpha = 0.05$.

4. Testování intervalem spolehlivosti

- Interval spolehlivosti

Pro výpočet intervalu spolehlivosti si nejdříve vypočítáme hodnotu s_g .

$$\begin{aligned}
s_g &= \sqrt{\frac{1}{n_1-3} + \frac{1}{n_2-3}} \\
&= \sqrt{\frac{1}{14-3} + \frac{1}{46-3}} = \sqrt{0.09090909 + 0.02325581} \\
&= \sqrt{0.1141649} = 0.337883 \doteq 0.3379
\end{aligned}$$

$$\begin{aligned}
(d, h) &= (\tanh(z_1 - z_2 - s_g u_{1-\alpha/2}); \tanh(z_1 - z_2 - s_g u_{\alpha/2})) \\
&= (\tanh(0.8142106 - 0.1379373 - 0.337883 \times u_{1-0.05/2}); \tanh(0.8142106 - 0.1379373 - 0.337883 \times u_{0.05/2})) \\
&= (\tanh(0.6762733 - 0.337883 \times 1.959964); \tanh(0.6762733 - 0.337883 \times (-1.959964))) \\
&= (\tanh(0.01403478); \tanh(1.338512)) \\
&= (0.01403386; 0.8713144)
\end{aligned}$$

```

328 sg <- sqrt(1 / (n1 - 3) + 1 / (n2 - 3)) # 0.337883
329 dh <- tanh(z1 - z2 - sg * qnorm(1 - alpha / 2)) # 0.01403392
330 hh <- tanh(z1 - z2 - sg * qnorm(alpha / 2)) # 0.8713143

```

- **Závěr testování**

Protože $\rho_0 = 0$ nenáleží do Waldova 95% empirického oboustranného intervalu spolehlivosti, tj. $\rho_0 = 0 \notin IS$, H_0 zamítáme na hladině významnosti $\alpha = 0.05$.

5. Testování p -hodnotou

- **p -hodnota**

$$\begin{aligned}
p\text{-hodnota} &= 2 \min\{\Pr(Z_w \leq z_w), \Pr(Z_W > z_w)\} \\
&= 2 \min\{\Pr(Z_w \leq z_w), 1 - \Pr(Z_W \leq z_w)\} \\
&= 2 \min\{\Pr(Z_w \leq 2.001502), 1 - \Pr(Z_W \leq 2.001502)\} \\
&= 2 \min\{0.9773308, 0.02266918\} \\
&= 2 \times 0.02266918 \\
&= 0.04533837
\end{aligned}$$

```

331 2 * min(pnorm(zw), 1 - pnorm(zw)) # 0.04533837

```

- **Závěr testování**

Protože p -hodnota = 0.04533837 je menší než $\alpha = 0.05$, H_0 zamítáme na hladině významnosti $\alpha = 0.05$.

6. Grafická vizualizace výsledků testování

Vhodný grafem porovnávajícím míru závislosti v obou náhodných výběrech je tečkový diagram superponovaný lineárními regresními přímkami prokládajícími zobrazené body (viz obrázek 36). Nejprve příkazem `plot()` vykreslíme body výšky a šířky nosu mužů bantuské populace. Koeficienty lineární regresní přímky získáme pomocí funkce `lm()`, jejímž jediným argumentem bude vztah `nose.BB ~ nose.HB`, tj. vztah vyjadřující závislost mezi proměnnou `nose.BB` na ose y a proměnnou `nose.HB` na ose x . Koeficienty regresní přímky, které jsou vloženy v poloze `coefficients`, získáme z výstupu funkce `lm` pomocí odkazu `$coef`. Dále vytvoříme posloupnost tisíce bodů x v rozsahu hodnot proměnné `nose.HB` a vypočítáme hodnoty regresní přímky v bodech posloupnosti x , které vložíme do proměnné y . Nakonec vykreslíme lineární regresní přímku v bodech x, y příkazem `lines()`.

Do grafu dále přikreslíme body výšky a šířky nosu peruánské populace příkazem `points()`, které opět superponujeme lineární regresní přímkou získanou funkcí `lm()`, jejímž vstupním argumentem bude vztah `nose.BP ~ nose.HP` budou proměnné `nose.HP`, tj. vztah vyjadřující závislost mezi proměnnou `nose.BP` na ose y a proměnnou `nose.HP` na ose x . Koeficienty regresní přímky opět získáme z výstupu funkce `lm` pomocí odkazu `$coef`. Dále vytvoříme posloupnost tisíce bodů x v rozsahu hodnot proměnné `nose.HP`, vypočítáme hodnoty regresní přímky v bodech posloupnosti x a vložíme je do proměnné y . Vykreslíme lineární regresní přímku v bodech x, y příkazem `lines()`. Nakonec do grafu doplníme pod osu x popisek obsahující hodnoty výběrových korelačních koeficientů (příkaz `mtext()`) a legendu (příkaz `legend()`).

```

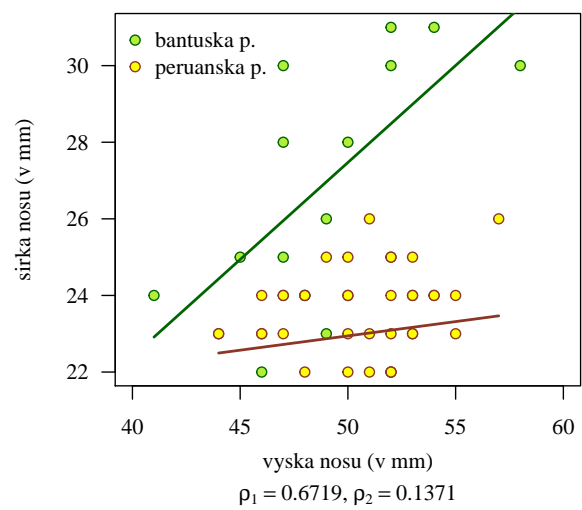
332 par(mar = c(5, 4, 1, 1), family = 'Times')
333 plot(nose.HB, nose.BB, xlab = '', ylab = 'sirka nosu (v mm)',
334       xlim = c(40, 60), las = 1,
335       pch = 21, col = 'darkgreen', bg = 'olivedrab2')
336 k <- lm(nose.BB ~ nose.HB)$coef
337 x <- seq(min(nose.HB), max(nose.HB), length = 1000)
338 y <- k[1] + x * k[2]
339 lines(x, y, col = 'darkgreen', lwd = 2)
340
341 points(nose.HP, nose.BP, pch = 21, col = 'tomato4', bg = 'yellow')
342 k <- lm(nose.BP ~ nose.HP)$coef
343 x <- seq(min(nose.HP), max(nose.HP), length = 1000)
344 y <- k[1] + x * k[2]
345 lines(x, y, col = 'tomato4', lwd = 2)
346
347 r1 <- round(r1, digit = 4)
348 r2 <- round(r2, digit = 4)
349 mtext('vyska nosu (v mm)', side = 1, line = 2.3)
350 mtext(bquote(paste(rho[1] == .(r1), ', ', rho[2] == .(r2))), side = 1, line = 3.7)
351 legend('topleft', pch = 21, pt.bg = c('olivedrab2', 'yellow'),
352       col = c('darkgreen', 'tomato4'), legend = c('bantuska p.', 'peruanska p.'),
353       bty = 'n')

```

7. Interpretace výsledků:

Na základě všech tří způsobů testování zamítáme hypotézu o shodě korelačních koeficientů pro výšku nosu a šířku nosu bantuské a peruánské populace. Mezi korelačním koeficientem výšky a šířky nosu u mužů bantuské populace a u mužů peruánské populace existuje statisticky významný rozdíl. Mezi výškou a šířkou nosu u mužů bantuské populace existuje význačný stupeň přímé lineární závislosti ($\rho_1 = 0.6719$). Mezi výškou a šířkou nosu u mužů peruánské populace existuje nízký stupeň přímé lineární závislosti ($\rho_2 = 0.1379$) (viz stupnice míry závislosti pro Pearsonův korelační koeficient, kapitola ??).





Obrázek 36: Tečkový diagram závislosti výšky nosu a šířky nosu u mužů bantuské a peruánské populace

Příklad 10.14. Test o rozdílu dvou korelačních koeficientů $\rho_1 - \rho_2$

Mějme datový soubor 16-anova-head.txt, proměnnou head.W popisující šířku hlavy a proměnnou bizyg.W popisující šířku tváře (viz sekce ??). Na hladině významnosti $\alpha = 0.05$ testujte hypotézu o shodě korelačního koeficientu šířky hlavy a šířky tváře mužů a korelačního koeficientu šířky hlavy a šířky tváře žen.

Řešení příkladu 10.14

Datový soubor načteme příkazem read.delim() a odstraníme z něj chybějící hodnoty příkazem na.omit(). Operátorem [] vybereme z tabulky údaje o šířce hlavy (head.W) a šířce tváře (bizyg.W) u mužů (sex == 'm'), resp. žen (sex == 'f'). Dále zjistíme rozsahy obou náhodných výběrů příkazem length() a rozsahy naměřených hodnot každé proměnné příkazem range(). Datový soubor obsahuje naměřené hodnoty šířky hlavy a šířky tváře u 63 mužů,

```
354 data <- read.delim('00-Data//16-anova-head.txt')
355 data <- na.omit(data)
356 head.WM <- data[data$sex == 'm', 'head.W']
357 bizyg.WM <- data[data$sex == 'm', 'bizyg.W']
358 head.WF <- data[data$sex == 'f', 'head.W']
359 bizyg.WF <- data[data$sex == 'f', 'bizyg.W']
360
361 n1 <- length(head.WM) # 63
362 n2 <- length(head.WF) # 85
363 range(head.WM) # 142-170
364 range(bizyg.WM) # 113-155
365 range(head.WF) # 135-162
366 range(bizyg.WF) # 120-151
```

přičemž naměřené šířky hlavy nabývají hodnot v rozmezí 142–170 mm, naměřené šířky tváře nabývají hodnot v rozmezí 113–155 mm. Soubor dále obsahuje údaje o šířce hlavy a šířce tváře 85 žen, přičemž naměřené šířky hlavy nabývají hodnot v rozmezí 135–162 mm a naměřené šířky tváře nabývají hodnot v rozmezí 120–151 mm.

Aby bylo možné otestovat hypotézu ze zadání pomocí parametrického testu o rozdílu dvou korelačních koeficientů, musí oba datové soubory pocházet z dvourozměrného normálního rozdělení. Před samotným testem hypotézu ze zadání je tedy potřeba ověřit, zda oba výběry tento předpoklad splňují. Závěr o dvourozměrné normalitě obou náhodných výběrů stanovíme na základě Roystonova testu ($\alpha = 0.05$) v kombinaci s grafickou vizualizací dat

```
367 MVN::mvt(cbind(head.WM, bizyg.WM), mvtTest = 'royston')$multivariateNormality #
0.003572831
```

Protože p -hodnota Roystonova testu, tj. 0.003573 je menší než 0.05, hypotézu o dvourozměrné normalitě náhodného výběru šířky lebky a šířky tváře u mužů zamítáme na hladině významnosti $\alpha = 0.05$. Ke stejnému závěru dojdeme také použitím Henze-Zirklerova testu (p -hodnota = 0.002498 < 0.05) i Mardiova testu (p -hodnota pro koeficient šikmosti = 0.006636 < 0.05, p -hodnota pro koeficient špičatosti = 0.2778 > 0.05). Nyní se podíváme na grafickou vizualizaci náhodného výběru (viz obrázek 37).

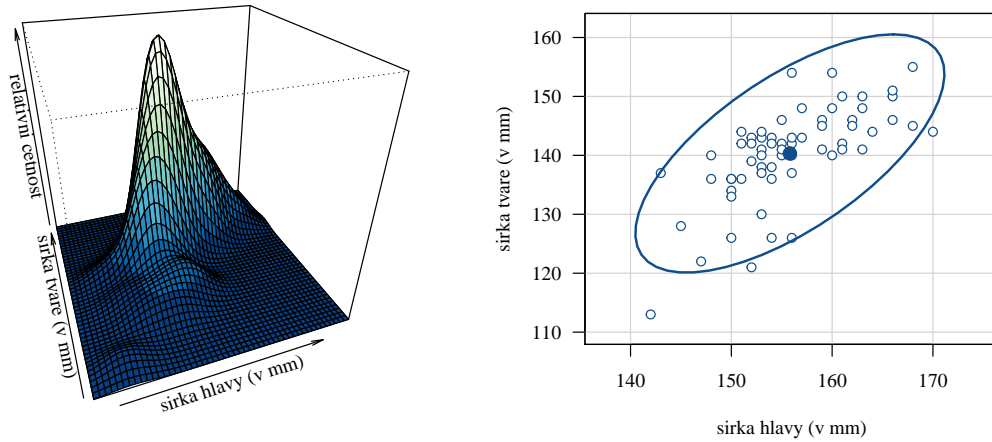
```
Loading required package: viridisLite
```

368

3D graf nám nabízí pohled na dvourozměrné normální rozdělení, tvořené pospolitým kopcem s jednou výrazně odlehlou hodnotou a několika méně odlehlými hodnotami. Aby data pocházela z dvourozměrného normálního rozdělení, je potřeba, aby alespoň 60 bodů leželo uvnitř 95% elipsy spolehlivosti. Zbylé 3 body se mohou realizovat mimo elipsu spolehlivosti. Z grafu vidíme, že mimo elipsu spolehlivosti leží právě 3 body, což ukazuje na dvourozměrnou normalitu náhodného výběru. V tomto případě tedy nastává rozkol mezi grafickou vizualizací a závěry testování dvourozměrné normality. Protože však všechny tři testy ukazují na porušení předpokladu normality, předpokládáme, že datový soubor šířek hlavy a šířek tváře mužů nepochází z normálního rozdělení.

Nyní prozkoumáme předpoklad normality šířky hlavy a šířky tváře u žen.

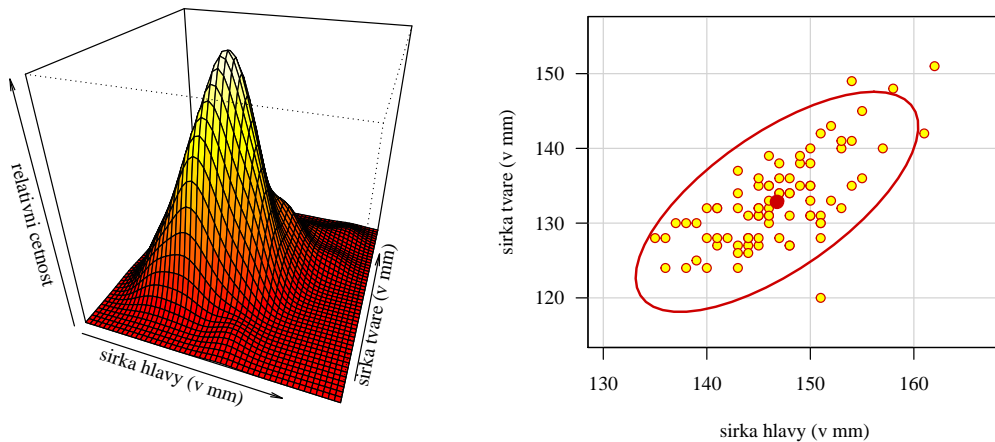
Jelikož p -hodnota Roystonova testu, tj. 0.02646746, je menší než 0.05, hypotézu o dvourozměrné normalitě náhodného výběru šířky hlavy a šířky tváře žen zamítáme na hladině významnosti $\alpha = 0.05$. Ke stejnému závěru



Obrázek 37: 3D graf a tečkový diagram s 95% elipsou spolehlivosti pro šířku hlavy a šířku tváře mužů (v mm)

```
369 MVN::mvm(cbind(head.WF, bizyg.WF), mvmTest = 'royston')$multivariateNormality #
0.02646746
```

bychom došli také Mardiovým testem (p -hodnota pro koeficient šikmosti = 0.00289277 > 0.05, p -hodnota pro koeficient špičatosti = 0.10385298 > 0.05). Na základě Henze-Zirklerova testu bychom hypotézu o dvourozměrné normalitě náhodného výběru nezamítli (p -hodnota = 0.2727 > 0.05). Zde tedy nastává rozkol mezi výsledky testů normality. Před stanovením závěru o rozdělení náhodného výběru se podívejme na vizualizaci datového souboru (viz obrázek 38).



Obrázek 38: 3D graf a tečkový diagram s 95% elipsou spolehlivosti pro šířku hlavy a šířku tváře žen (v mm)

3D graf nám nabízí pohled na dvourozměrné normální rozdělení tvořené pospolitým kopcem s alespoň jednou výrazně odlehlou hodnotou a několika méně odlehlými hodnotami. Aby data pocházela z dvourozměrného normálního rozdělení, je potřeba, aby alespoň 81 bodů leželo uvnitř 95% elipsy spolehlivosti. Zbylé 4 body mohou ležet mimo elipsu spolehlivosti. Z tečkového grafu je zřejmé, že mimo elipsu spolehlivosti však leží 5 bodů. V tomto případě tedy grafická vizualizace podporuje výsledek Roystonova a Mardiova testu. Na základě výsledků testování a grafické vizualizace tedy zamítáme hypotézu o dvourozměrném normálním rozdělení náhodného výběru šířek hlavy a šířek tváře žen.

Protože předpoklad dvourozměrné normality obou náhodných výběrů je porušen, nemůžeme k testování hypotézy ze zadání použít parametrický test o rozdílu korelačních koeficientů $\rho_1 - \rho_2$. ★

Příklad 10.15. Test o rozdílu dvou korelačních koeficientů $\rho_1 - \rho_2$

Mějme datový soubor 05-one-sample-correlation-skull-mf.txt, proměnnou skull.pH popisující největší výšku mozkovny a proměnnou face.H popisující morfologickou výšku tváře (viz sekce ??). Na hladině významnosti $\alpha = 0.10$ zjistíte, zda korelační koeficient největší výšky mozkovny a morfologické výšky tváře mužů je statisticky významně větší než korelační koeficient největší výšky mozkovny a morfologické výšky tváře žen.

Řešení příkladu 10.15

Datový soubor načteme příkazem read.delim() a odstraníme z něj chybějící hodnoty příkazem na.omit(). Operátorem [] vybereme z tabulky nejprve údaje o největší výšce mozkovny (skull.pH), resp. morfologické výšce tváře (face.H) u mužů (sex == 'm') a následně údaje o největší výšce mozkovny (skull.pH), resp. morfologické výšce tváře (face.H) u žen (sex == 'f'). Dále zjistíme rozsahy obou náhodných výběrů (length()) rozsahy naměřených hodnot (range()) každé proměnné. Datový soubor obsahuje údaje o největší výšce mozkovny a morfologické výšce tváře u 164 mužů,

```
370 data <- read.delim('00-Data//05-one-sample-correlation-skull-mf.txt')
371 data <- na.omit(data)
372 head(data)
373 skull.pHM <- data[data$sex == 'm', 'skull.pH']
374 face.HM <- data[data$sex == 'm', 'face.H']
375 skull.pHF <- data[data$sex == 'f', 'skull.pH']
376 face.HF <- data[data$sex == 'f', 'face.H']
377
378 n1 <- length(skull.pHM) # 164
379 n2 <- length(skull.pHF) # 78
380 range(skull.pHM) # 127-149
381 range(face.HM) # 100-136
382 range(skull.pHF) # 120-142
383 range(face.HF) # 95-119
```

přičemž naměřené výšky mozkovny nabývají hodnot v rozmezí 127–149 mm, naměřené morfologické výšky tváře nabývají hodnot v rozmezí 100–136 mm. Soubor dále obsahuje údaje o největší výšce mozkovny a morfologické výšce tváře u 78 žen, přičemž naměřené výšky mozkovny nabývají hodnot v rozmezí 120–142 mm a naměřené hodnoty morfologické výšky tváře nabývají hodnot v rozmezí 95–119 mm.

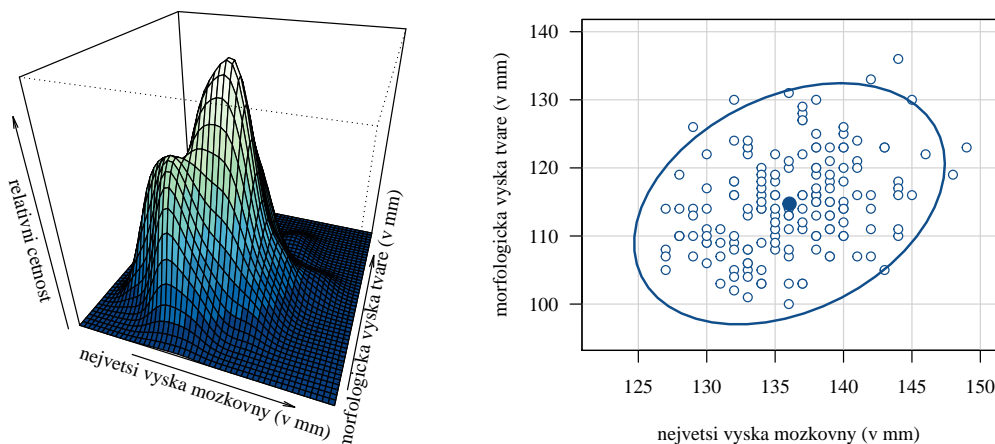
Řešení příkladu vede na test o rozdílu dvou korelačních koeficientů. Aby bylo možné provést parametrický test o rozdílu dvou korelačních koeficientů, musí oba datové soubory splňovat předpoklad dvourozměrné normality. Před samotným testováním je tedy potřeba ověřit, zda oba výběry tento předpoklad splňují. Závěr o dvourozměrné normalitě obou náhodných výběrů stanovíme na základě Henze-Zirklerova testu ($\alpha = 0.05$) v kombinaci s grafickou vizualizací dat.

```
384 MVN::mvn(cbind(skull.pHM, face.HM), mvnTest = 'hz')$multivariateNormality # 0.719099
```

Protože p -hodnota Henze-Zirklerova testu, tj. 0.7191 je větší než 0.05, hypotézu o dvourozměrné normalitě náhodného výběru největší výšky mozkovny a morfologické výšky tváře u mužů nezamítáme na hladině významnosti $\alpha = 0.05$. Ke stejnému závěru bychom došli také použitím Mardiova testu (p -hodnota pro koeficient šikmosti = 0.3853 > 0.05, p -hodnota pro koeficient špičatosti = 0.4620 > 0.05) i Roystonova testu (p -hodnota = 19998 > 0.05).

Nyní se podíváme na grafickou vizualizaci náhodného výběru (viz obrázek 39). 3D graf nám nabízí pohled na dvourozměrné normální rozdělení, tvořené dvěma vzájemně se prolínajícími kopci s několika odlehlými hodnotami. Zda je množství odlehlých hodnot v normě, zjistíme pohledem na tečkový diagram. Aby data pocházela z dvourozměrného normálního rozdělení stačí, aby alespoň 95 % bodů, tj. 156 bodů, leželo uvnitř elipsy spolehlivosti. Zbylých 8 bodů může ležet mimo elipsu spolehlivosti. Z grafu vidíme, že mimo elipsu spolehlivosti leží právě 8 bodů, což je v pořádku. O náhodném výběru největší výšky mozkovny a morfologické výšky tváře u mužů tedy na základě testování i grafické vizualizace předpokládáme, že pochází z dvourozměrného normálního rozdělení. Nyní prozkoumáme předpoklad normality největší výšky mozkovny a morfologické výšky tváře u žen.

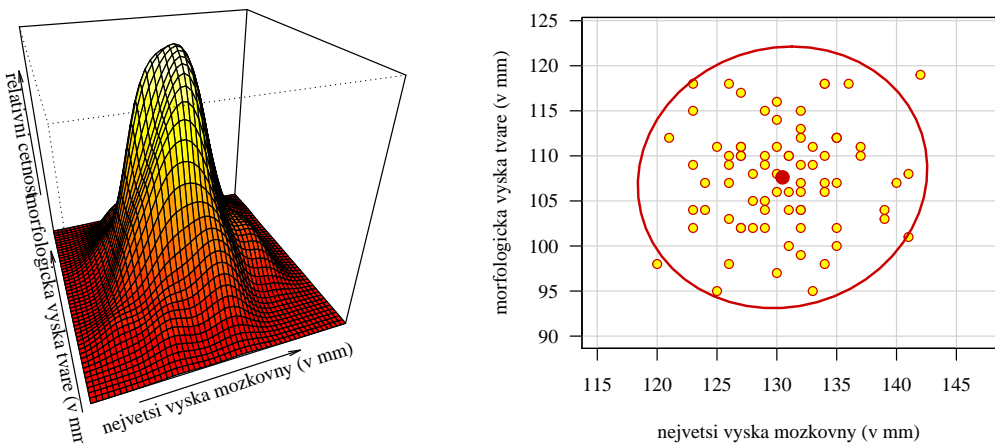
Jelikož p -hodnota Henze-Zirklerova testu, tj. 0.9997 je větší než 0.05, hypotézu o dvourozměrné normalitě náhodného výběru největší výšky mozkovny a morfologické výšky tváře žen nezamítáme na hladině významnosti



Obrázek 39: 3D graf a tečkový diagram s 95% elipsou spolehlivosti pro největší výšku mozkovny a morfológickou výšku tváře mužů (v mm)

```
385 MVN::mvm(cbind(skull.pHF, face.HF), mvnTest = 'hz')$multivariateNormality # 0.9997521
```

$\alpha = 0.05$. Ke stejnému závěru bychom došli také Mardiovým testem (p -hodnota pro koeficient šikmosti = $0.9674 > 0.05$, p -hodnota pro koeficient špičatosti = $0.6223 > 0.05$) i Roystonovým testem (p -hodnota = $0.4054 > 0.05$).



Obrázek 40: 3D graf a tečkový diagram s 95% elipsou spolehlivosti pro největší výšku mozkovny a morfológickou výšku tváře žen (v mm)

3D graf nabízí pohled na pospolité dvourozměrné normální rozdělení tvořené jedním souvislým kopcem (viz obrázek 40). Aby data pocházela z dvourozměrného normálního rozdělení, je potřeba, aby 95% elipsa spolehlivosti pokrývala alespoň 74 bodů. Zbylé čtyři body se mohou realizovat mimo elipsu spolehlivosti. Pohledem na tečkový diagram zjišťujeme, že mimo elipsu spolehlivosti leží pouze tři body, což podporuje náš závěr o dvourozměrné normalitě náhodného výběru. O náhodném výběru největší výšky mozkovny a morfológické výšky tváře u žen tedy na základě testování i grafické vizualizace předpokládáme, že pochází z dvourozměrného normálního rozdělení.

Protože oba výběry pochází z dvourozměrného normálního rozdělení, můžeme k testování použít parametrický test o rozdílu korelačních koeficientů $\rho_1 - \rho_2$. Zadaným úkolem je zjistit, zda korelační koeficient největší výšky mozkovny a morfológické výšky tváře mužů je statisticky významně větší než korelační koeficient největší výšky mozkovny a morfológické výšky tváře žen. Toto tvrzení je zněním alternativní hypotézy a nulové hypotéza tvoří doplněk k alternativní hypotéze.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : Korelační koeficient největší výšky mozkovny a morfologické výšky tváře mužů je menší nebo roven korelačnímu koeficientu největší výšky mozkovny a morfologické výšky tváře žen.

H_1 : Korelační koeficient největší výšky mozkovny a morfologické výšky tváře mužů je větší než korelační koeficient největší výšky mozkovny a morfologické výšky tváře žen.

- **matematická formulace** nulové a alternativní hypotézy

$H_0 : \rho_1 \leq \rho_2 \rightarrow \rho_1 - \rho_2 \leq \rho_0$, kde $\rho_0 = 0$

$H_1 : \rho_1 > \rho_2 \rightarrow \rho_1 - \rho_2 > \rho_0$, kde $\rho_0 = 0$

(pravostranná alternativa)

2. Volba hladiny významnosti

- Hladinu významnosti volíme podle zadání $\alpha = 0.10$.

3. Testování kritickým oborem

- Fisherovy Z -transformace výběrových korelačních koeficientů

K výpočtu Fisherových Z -transformací je potřeba nejprve stanovit hodnoty výběrových korelačních koeficientů, a to pomocí příkazu `cor()`. Výběrový korelační koeficient výšky mozkovny a morfologické výšky tváře mužů $r_1 = 0.33064$, výběrový korelační koeficient výšky a šířky nosu peruánské populace vyšel $r_2 = 0.06417$. Nyní vypočítáme Fisherovy Z -transformace obou výběrových korelačních koeficientů a konstanty $\rho_0 = 0$ z nulové hypotézy.

$$\begin{aligned} Z_1 &= \frac{1}{2} \ln \frac{1 + R_1}{1 - R_1} \\ &= \frac{1}{2} \ln \frac{1 + 0.3306431}{1 - 0.3306431} = \frac{1}{2} \ln 1.987943 \\ &= 0.5 \times 1.62842 = 0.81421 \doteq 0.8142 \end{aligned}$$

$$\begin{aligned} Z_2 &= \frac{1}{2} \ln \frac{1 + R_2}{1 - R_2} \\ &= \frac{1}{2} \ln \frac{1 + 0.06417166}{1 - 0.06417166} = \frac{1}{2} \ln 1.137144 \\ &= 0.5 \times 0.1285199 = 0.06425995 \end{aligned}$$

$$\begin{aligned} \xi_0 &= \frac{1}{2} \ln \frac{1 + \rho_0}{1 - \rho_0} \\ &= \frac{1}{2} \ln \frac{1 + 0}{1 - 0} = \frac{1}{2} \ln 1 \\ &= 0.5 \times 0 = 0 \end{aligned}$$

```
386 alpha <- 0.10
387 ksi0 <- 0
388 r1 <- cor(skull.pHM, face.HM) # 0.3306431
389 r2 <- cor(skull.pHF, face.HF) # 0.06417166
390 z1 <- 1 / 2 * log((1 + r1) / (1 - r1)) # 0.3435501
391 z2 <- 1 / 2 * log((1 + r2) / (1 - r2)) # 0.06425996
```

- Testovací statistika

$$\begin{aligned}
 Z_W &= \frac{Z_1 - Z_2 - \xi_0}{\sqrt{\frac{1}{n_1-3} + \frac{1}{n_2-3}}} \\
 &= \frac{0.3435501 - 0.06425996 - 0}{\sqrt{\frac{1}{164-3} + \frac{1}{78-3}}} \\
 &= \frac{0.2792901}{0.2792901} \\
 &= \frac{0.2792901}{\sqrt{0.00621118 + 0.01333333}} \\
 &= \frac{0.2792901}{\sqrt{0.01954451}} \\
 &= \frac{0.2792901}{0.1398017} \\
 &= 1.997759
 \end{aligned}$$

```
392 zW <- (z1 - z2 - ksi0) / sqrt(1 / (n1 - 3) + 1 / (n2 - 3)) # 1.997759
```

- Kritický obor

$$\begin{aligned}
 W &= \langle u_{1-\alpha}; \infty \rangle \\
 &= \langle u_{1-0.10}; \infty \rangle \\
 &= \langle u_{0.90}; \infty \rangle \\
 &= \langle 1.281552; \infty \rangle
 \end{aligned}$$

```
393 qnorm(1 - alpha) # 1.281552
```

- Závěr testování

Protože realizace testovací statistiky $z_W = 1.9978$ náleží do kritického oboru, tj. $z_W \in W$, H_0 zamítáme na hladině významnosti $\alpha = 0.10$.

4. Testování intervalem spolehlivosti

- Interval spolehlivosti

Pro výpočet intervalu spolehlivosti si nejdříve vypočítáme hodnotu s_g .

$$\begin{aligned}
 s_g &= \sqrt{\frac{1}{n_1-3} + \frac{1}{n_2-3}} \\
 &= \sqrt{\frac{1}{164-3} + \frac{1}{78-3}} = \sqrt{0.00621118 + 0.01333333} \\
 &= \sqrt{0.01954451} = 0.1398017 \doteq 0.1398
 \end{aligned}$$

$$\begin{aligned}
 (d, 2) &= (\tanh(z_1 - z_2 - s_g u_{1-\alpha}); 2) \\
 &= (\tanh(0.3435501 - 0.06425996 - 0.1398017 \times u_{1-0.10}); 2) \\
 &= (\tanh(0.2792901 - 0.1398017 \times 1.281552); 2) \\
 &= (\tanh(0.100127); 2) \\
 &= (0.09979368; 2)
 \end{aligned}$$

```

394 sg <- sqrt(1 / (n1 - 3) + 1 / (n2 - 3)) # 0.1398017
395 dh <- tanh(z1 - z2 - sg * qnorm(1 - alpha)) # 0.0997938

```

- **Závěr testování**

Protože $\rho_0 = 0$ nenáleží do Waldova 90% empirického jednostranného intervalu spolehlivosti, tj. $\rho_0 = 0 \notin IS$, H_0 zamítáme na hladině významnosti $\alpha = 0.10$.

5. Testování p -hodnotou

- p -hodnota

$$\begin{aligned}
 p\text{-hodnota} &= \Pr(Z_W > z_w) \\
 &= 1 - \Pr(Z_W \leq z_w) \\
 &= 1 - \Pr(Z_W \leq 1.997759) \\
 &= 0.02287137
 \end{aligned}$$

```

396 1 - pnorm(zw) # 0.02287137

```

- **Závěr testování**

Protože p -hodnota = 0.02287 je menší než $\alpha = 0.10$, H_0 zamítáme na hladině významnosti $\alpha = 0.10$.

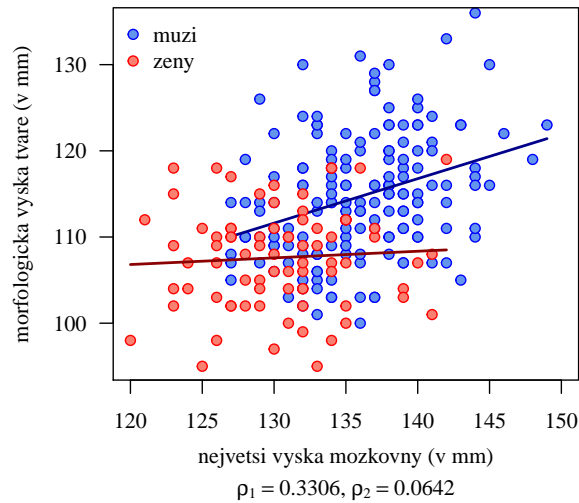
6. Grafická vizualizace výsledků testování

Rozdíl v míře závislosti mezi oběma náhodnými výběry vizualizujeme pomocí tečkového diagramu s lineárními regresními přímkami prokládajícími zobrazené body (viz obrázek 44). Koefficienty lineární regresní přímky pro největší výšku mozkovny a morfologickou výšku tváře mužů získáme pomocí funkce `lm()`, jehož argumentem bude vztah `face.HM ~ skull.pHM` vyjadřující závislost mezi proměnnou `face.HM` na ose y a proměnnou `skull.pHM` na ose x . Do grafu dále dokreslíme body největší výšky mozkovny a morfologické výšky tváře žen, které superponujeme lineární regresní přímkou získanou funkcí `lm()` s argumentem `face.HF ~ skull.pHF`. Nakonec do grafu doplníme popisek obsahující hodnoty výběrových korelačních koeficientů a legendu.

```

397 par(mar = c(5, 4, 1, 1), family = 'Times')
398 plot(skull.pHM, face.HM, xlab = '', ylab = 'morfologicka vyska tvare (v mm)',
399      xlim = c(120, 150), ylim = c(95, 135), las = 1,
400      pch = 21, col = 'blue', bg = 'cornflowerblue')
401 k <- lm(face.HM ~ skull.pHM)$coef
402 x <- seq(min(skull.pHM), max(skull.pHM), length = 1000)
403 y <- k[1] + x * k[2]
404 lines(x, y, col = 'darkblue', lwd = 2)
405
406 points(skull.pHF, face.HF, pch = 21, col = 'red', bg = 'salmon')
407 k <- lm(face.HF ~ skull.pHF)$coef
408 x <- seq(min(skull.pHF), max(skull.pHF), length = 1000)
409 y <- k[1] + x * k[2]
410 lines(x, y, col = 'darkred', lwd = 2)
411
412 r1 <- round(r1, digit = 4)
413 r2 <- round(r2, digit = 4)
414 mtext('nejvetsi vyska mozkovny (v mm)', side = 1, line = 2.3)
415 mtext(bquote(paste(rho[1] == .(r1), ', ', rho[2] == .(r2))), side = 1, line = 3.7)
416 legend('topleft', pch = 21, pt.bg = c('cornflowerblue', 'salmon'),
417       col = c('blue', 'red'), legend = c('muzi', 'zeny'),
418       bty = 'n')

```



Obrázek 41: Krabicový diagram největší výšky mozkovny a morfologické výšky tváře žen (v mm)

7. Interpretace výsledků:

Na základě všech tří způsobů testování zamítáme hypotézu H_0 . Korelační koeficient největší výšky mozkovny a morfologické výšky tváře u mužů je statisticky významně větší než korelační koeficient největší výšky mozkovny a morfologické výšky tváře u žen. Mezi největší výškou mozkovny a morfologickou výškou tváře u mužů existuje mírný stupeň přímé lineární závislosti ($\rho_1 = 0.3306$). Mezi největší výškou mozkovny a morfologickou výškou tváře u žen existuje velmi nízký stupeň přímé lineární závislosti ($\rho_1 = 0.0642$).

★

Příklad 10.16. Test o rozdílu dvou korelačních koeficientů $\rho_1 - \rho_2$

Mějme datový soubor 13-two-samples-correlations-trunk.txt, proměnnou lowex.L popisující délku dolní končetiny v mm a proměnnou tru.L popisující délku trupu v mm (viz sekce ??). Na hladině významnosti $\alpha = 0.01$ zjistíte, zda je korelační koeficient délky dolní končetiny a délky trupu u mužů menší než korelační koeficient délky dolní končetiny a délky trupu žen.

Řešení příkladu 10.16

Datový soubor načteme příkazem read.delim() a odstraníme z něj chybějící hodnoty. Operátorem [] vybereme z tabulky nejprve údaje o délce dolní končetiny (lowex.L) a délce trupu (tru.L) u mužů (sex == 'm'), resp. u žen (sex == 'f'). Dále zjistíme rozsahy obou náhodných výběrů a rozsahy naměřených hodnot každé proměnné. Datový

```
419 data <- read.delim('00-Data//13-two-samples-correlations-trunk.txt')
420 data <- na.omit(data)
421 lowex.LM <- data[data$sex == 'm', 'lowex.L']
422 tru.LM <- data[data$sex == 'm', 'tru.L']
423 lowex.LF <- data[data$sex == 'f', 'lowex.L']
424 tru.LF <- data[data$sex == 'f', 'tru.L']
425
426 n1 <- length(lowex.LM) # 75
427 n2 <- length(lowex.LF) # 100
428 range(lowex.LM) # 915-1114
429 range(tru.LM) # 400-548
430 range(lowex.LF) # 836-1076
431 range(tru.LF) # 323-492
```

soubor obsahuje údaje o délce dolní končetiny a délce trupu u 75 mužů, přičemž naměřené délky dolní končetiny nabývají hodnot v rozmezí 915–1114 mm, naměřené délky trupu nabývají hodnot v rozmezí 400–548 mm. Soubor dále obsahuje údaje o největší délce dolní končetiny a délce trupu u 100 žen, přičemž naměřené délky dolní končetiny nabývají hodnot v rozmezí 836–1076 mm a naměřené délky trupu nabývají hodnot v rozmezí 323–492 mm.

Řešení příkladu vede na test o rozdílu dvou korelačních koeficientů. Před provedením parametrického testu o rozdílu korelačních koeficientů je potřeba ověřit předpoklad dvourozměrné normality obou náhodných výběrů. Závěr o dvourozměrné normalitě obou výběrů stanovíme na základě Henze-Zirklerova testu ($\alpha = 0.05$) v kombinaci s grafickou vizualizací dat.

```
432 MVN::mvt(cbind(lowex.LM, tru.LM), mvnTest = 'hz')$multivariateNormality # 0.543489
```

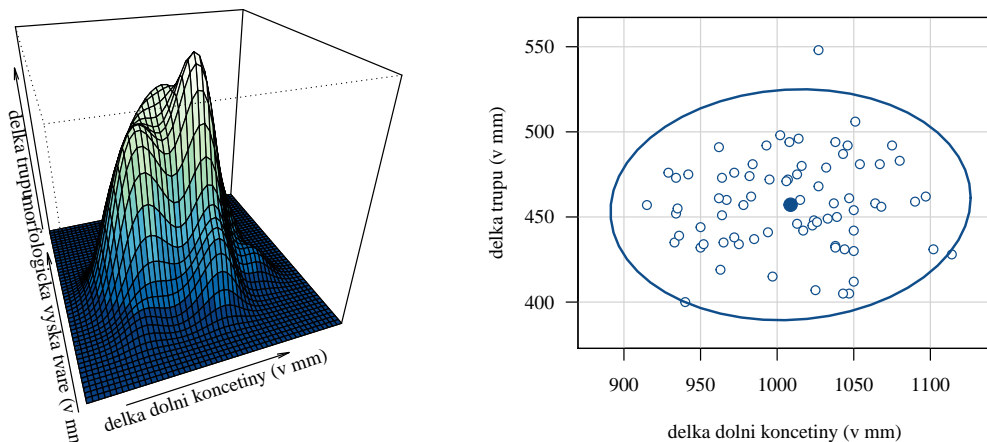
Protože p -hodnota Henze-Zirklerova testu, tj. 0.5435 je větší než 0.05, hypotézu o dvourozměrné normalitě náhodného výběru délky dolní končetiny a délky trupu mužů nezamítáme na hladině významnosti $\alpha = 0.05$. Ke stejnému závěru bychom došli také použitím Mardiova testu (p -hodnota pro koeficient šikmosti = 0.5599 > 0.05, p -hodnota pro koeficient špičatosti = 0.4298 > 0.05) i Roystonova testu (p -hodnota = 0.1876 > 0.05).

Na 3D grafu vidíme dvourozměrné normální rozdělení tvořené jedním kopcem (viz obrázek 42). Aby data pocházela z dvourozměrného normálního rozdělení stačí, aby alespoň 71 bodů leželo uvnitř 95% elipsy spolehlivosti. Zbylé 4 body mohou ležet mimo elipsu spolehlivosti. Z tečkového diagramu vidíme, že mimo elipsu spolehlivosti leží pouze dva body a jeden bod leží na hranici elipsy spolehlivosti. O náhodném výběru největší délky dolní končetiny a délky trupu mužů tedy předpokládáme, že pochází z dvourozměrného normálního rozdělení.

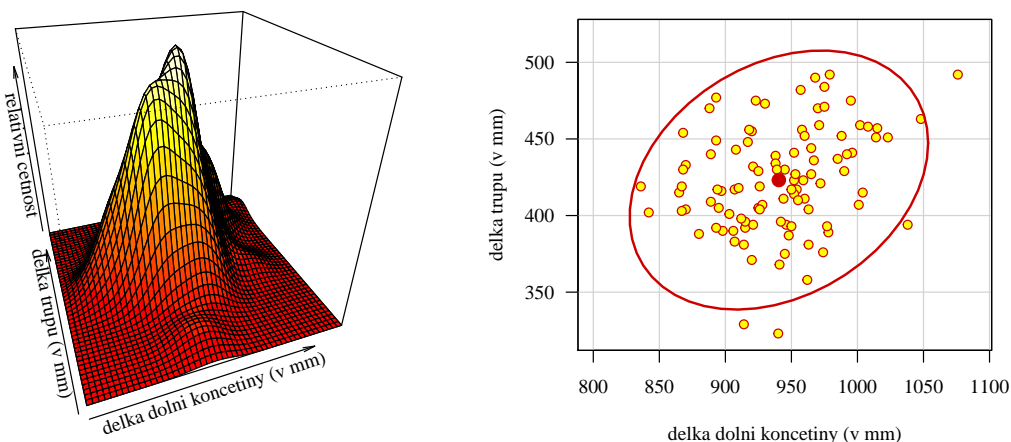
Nyní prozkoumáme předpoklad normality délky dolní končetiny a délky trupu u žen.

```
433 MVN::mvt(cbind(lowex.LF, tru.LF), mvnTest = 'hz')$multivariateNormality # 0.587225
```

Jelikož p -hodnota Henze-Zirklerova testu, tj. 0.5872 je větší než 0.05, hypotézu o dvourozměrné normalitě náhodného výběru největší výšky mozkovny a morfologické výšky tváře žen nezamítáme na hladině významnosti $\alpha = 0.05$. Ke stejnému závěru bychom došli také Mardiovým testem (p -hodnota pro koeficient šikmosti = 0.1769 > 0.05, p -hodnota pro koeficient špičatosti = 0.8360 > 0.05) i Roystonovým testem (p -hodnota = 0.6339 > 0.05).



Obrázek 42: 3D graf a tečkový diagram s 95% elipsou spolehlivosti pro délku dolní končetiny a délku trupu u mužů (v mm)



Obrázek 43: 3D graf a tečkový diagram s 95% elipsou spolehlivosti pro délku dolní končetiny a délku trupu u žen (v mm)

3D graf nabízí pohled na dvourozměrné normální rozdělení tvořené jedním souvislým kopcem a několik odlehlými hodnotami (viz obrázek 43). Aby data pocházela z dvourozměrného normálního rozdělení, je potřeba, aby odlehlých hodnot nebylo více než pět. Pohledem na tečkový diagram zjišťujeme, že mimo elipsu spolehlivosti leží pouze čtyři body. O náhodném výběru délky dolní končetiny a délky trupu žen tedy předpokládáme, že pochází z dvourozměrného normálního rozdělení.

Protože oba výběry pochází z dvourozměrného normálního rozdělení, můžeme k testování použít parametrický test o rozdílu korelačních koeficientů $\rho_1 - \rho_2$. Zadaným úkolem je zjistit, zda korelační koeficient délky dolní končetiny a délky trupu mužů je statisticky významně menší než korelační koeficient délky dolní končetiny a délky trupu žen. Toto tvrzení je zněním alternativní hypotézy a nulové hypotéza tvoří doplněk k alternativní hypotéze.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : Korelační koeficient délky dolní končetiny a délky trupu mužů je větší nebo rovek korelačnímu koeficientu délky dolní končetiny a délky trupu žen.

H_1 : Korelační koeficient délky dolní končetiny a délky trupu mužů je menší než korelační koeficient délky dolní končetiny a délky trupu žen.

- **matematická formulace** nulové a alternativní hypotézy

$$H_0 : \rho_1 \geq \rho_2 \quad \rightarrow \quad \rho_1 - \rho_2 \geq \rho_0, \text{ kde } \rho_0 = 0$$

$$H_1 : \rho_1 < \rho_2 \quad \rightarrow \quad \rho_1 - \rho_2 < \rho_0, \text{ kde } \rho_0 = 0$$

(levostranná alternativa)

2. Volba hladiny významnosti

- Hladinu významnosti volíme podle zadání $\alpha = 0.05$.

3. Testování kritickým oborem

- Fisherovy Z -transformace výběrových korelačních koeficientů

K výpočtu Fisherových Z -transformací je potřeba nejprve stanovit hodnoty výběrových korelačních koeficientů. Výběrový korelační koeficient délky dolní končetiny a délky trupu mužů $r_1 = 0.05976$, výběrový korelační koeficient délky dolní končetiny a délky trupu žen vyšel $r_2 = 0.285256$. Nyní vypočítáme Fisherovy Z -transformace obou výběrových korelačních koeficientů a konstanty $\rho_0 = 0$ z nulové hypotézy.

$$\begin{aligned} Z_1 &= \frac{1}{2} \ln \frac{1 + R_1}{1 - R_1} \\ &= \frac{1}{2} \ln \frac{1 + 0.05976}{1 - 0.05976} = \frac{1}{2} \ln 1.127116 \\ &= 0.5 \times 0.1196626 = 0.0598313 \doteq 0.05983 \end{aligned}$$

$$\begin{aligned} Z_2 &= \frac{1}{2} \ln \frac{1 + R_2}{1 - R_2} \\ &= \frac{1}{2} \ln \frac{1 + 0.285256}{1 - 0.285256} = \frac{1}{2} \ln 1.798205 \\ &= 0.5 \times 0.5867888 = 0.2933944 \end{aligned}$$

$$\begin{aligned} \xi_0 &= \frac{1}{2} \ln \frac{1 + \rho_0}{1 - \rho_0} \\ &= \frac{1}{2} \ln \frac{1 + 0}{1 - 0} = \frac{1}{2} \ln 1 \\ &= 0.5 \times 0 = 0 \end{aligned}$$

```
434 alpha <- 0.01
435 ksi0 <- 0
436 r1 <- cor(lowex.LM, tru.LM) # 0.05975781
437 r2 <- cor(lowex.LF, tru.LF) # 0.285256
438 z1 <- 1 / 2 * log((1 + r1) / (1 - r1)) # 0.0598291
439 z2 <- 1 / 2 * log((1 + r2) / (1 - r2)) # 0.2933943
```

- **Testovací statistika**

$$\begin{aligned}
Z_W &= \frac{Z_1 - Z_2 - \xi_0}{\sqrt{\frac{1}{n_1-3} + \frac{1}{n_2-3}}} \\
&= \frac{0.0598291 - 0.2933944 - 0}{\sqrt{\frac{1}{75-3} + \frac{1}{100-3}}} \\
&= \frac{-0.2335653}{\sqrt{0.01388889 + 0.01030928}} \\
&= \frac{-0.2335653}{\sqrt{0.02419817}} \\
&= \frac{-0.2335653}{0.1555576} \\
&= -1.501471
\end{aligned}$$

```
440 zw <- (z1 - z2 - ksi0) / sqrt(1 / (n1 - 3) + 1 / (n2 - 3)) # -1.501471
```

- Kritický obor

$$\begin{aligned}
W &= (-\infty; u_\alpha) \\
&= (-\infty; u_{0.01}) \\
&= (-\infty; -1.644854)
\end{aligned}$$

```
441 qnorm(alpha) # -2.326348
```

- Závěr testování

Protože realizace testovací statistiky $z_W = -1.5015$ nenáleží do kritického oboru, tj. $z_W \notin W$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

4. Testování intervalem spolehlivosti

- Interval spolehlivosti

Pro výpočet intervalu spolehlivosti si nejdříve vypočítáme hodnotu s_g .

$$\begin{aligned}
s_g &= \sqrt{\frac{1}{n_1-3} + \frac{1}{n_2-3}} \\
&= \sqrt{\frac{1}{75-3} + \frac{1}{100-3}} = \sqrt{0.01388889 + 0.01030928} \\
&= \sqrt{0.02419817} = 0.1555576 \doteq 0.1556
\end{aligned}$$

$$\begin{aligned}
(-2, h) &= (-2; \tanh(z_1 - z_2 - s_g u_\alpha)) \\
(-2, h) &= (-2; \tanh(0.0598291 - 0.2933944 - 0.1555576 u_{0.01})) \\
&= (-2; \tanh(-0.2335653 - 0.1555576 \times (-2.326348))) \\
&= (-2; \tanh(0.1283158)) \\
&= (-2; 0.1276162)
\end{aligned}$$


```
442 sg <- sqrt(1 / (n1 - 3) + 1 / (n2 - 3)) # 0.1555576
443 hh <- tanh(z1 - z2 - sg * qnorm(alpha)) # 0.1276162
```

- **Závěr testování**

Protože $\rho_0 = 0$ náleží do Waldova 99% empirického pravostanného intervalu spolehlivosti, tj. $\rho_0 = 0 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

5. Testování p -hodnotou

- p -hodnota

$$\begin{aligned} p\text{-hodnota} &= \Pr(Z_W \leq z_w) \\ &= \Pr(Z_W \leq -1.501471) \\ &= 0.06661688 \end{aligned}$$

```
444 pnorm(zw) # 0.06661688
```

- **Závěr testování**

Protože p -hodnota = 0.06662 je větší než $\alpha = 0.01$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

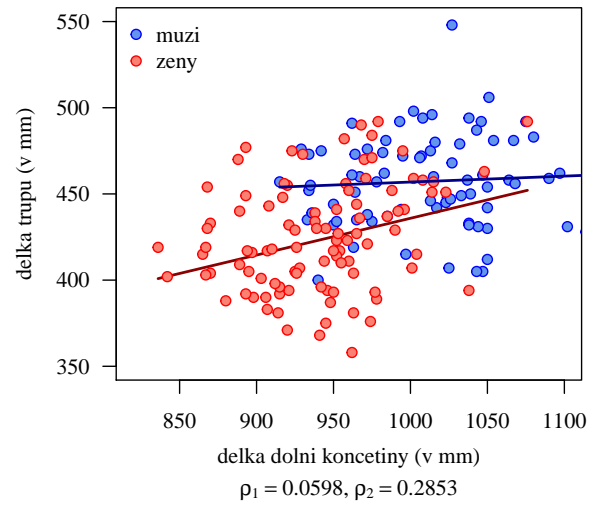
6. Grafická vizualizace výsledků testování

Míru závislosti obou náhodných výběrů vizualizujeme pomocí tečkového diagramu s lineárními křivkami prokládajícími naměřené body (viz obrázek 44). Koefficienty lineární přímky pro délku dolní končetiny a délku trupu mužů, resp. žen, získáme pomocí funkce `lm()`, analogicky jako v předchozích příkladech.

```
445 par(mar = c(5, 4, 1, 1), family = 'Times')
446 plot(lowex.LM, tru.LM, xlab = '', ylab = 'delka trupu (v mm)',
447       xlim = c(820, 1100), ylim = c(350, 550), las = 1,
448       pch = 21, col = 'blue', bg = 'cornflowerblue')
449 k <- lm(tru.LF ~ lowex.LF)$coef
450 x <- seq(min(lowex.LF), max(lowex.LF), length = 1000)
451 y <- k[1] + x * k[2]
452 lines(x, y, col = 'darkred', lwd = 2)
453
454 points(lowex.LF, tru.LF, pch = 21, col = 'red', bg = 'salmon')
455 k <- lm(tru.LM ~ lowex.LM)$coef
456 x <- seq(min(lowex.LM), max(lowex.LM), length = 1000)
457 y <- k[1] + x * k[2]
458 lines(x, y, col = 'darkblue', lwd = 2)
459
460 r1 <- round(r1, digit = 4)
461 r2 <- round(r2, digit = 4)
462 mtext('delka dolni koncetiny (v mm)', side = 1, line = 2.3)
463 mtext(bquote(paste(rho[1] == .(r1), ', ', rho[2] == .(r2))), side = 1, line = 3.7)
464 legend('topleft', pch = 21, pt.bg = c('cornflowerblue', 'salmon'),
465       col = c('blue', 'red'), legend = c('muzi', 'zeny'),
466       bty = 'n')
```

7. Interpretace výsledků:

Na základě všech tří způsobů testování nezamítáme hypotézu H_0 . Korelační koeficient délky dolní končetiny a délky trupu mužů není statisticky významně menší než korelační koeficient délky dolní končetiny a délky trupu žen. Mezi délkou dolní končetiny a délkou trupu mužů existuje velmi nízký stupeň přímé lineární závislosti



Obrázek 44: Krabicový diagram s 95% elipsou spolehlivosti pro délku dolní končetiny a délku trupu u žen (v mm)

($\rho_1 = 0.0598$). Mezi délkou dolní končetiny a délkou trupu žen existuje nízký stupeň přímé lineární závislosti ($\rho_1 = 0.2853$).



10.5 Test o rozdílu parametrů $p_1 - p_2$ dvou alternativních rozdělení

Nechť X_{11}, \dots, X_{1N_1} je náhodný výběr z alternativního rozdělení $\text{Alt}(p_1)$ a X_{21}, \dots, X_{2N_2} je na něm nezávislý náhodný výběr s alternativního rozdělení $\text{Alt}(p_2)$. Nechť p_0 je konstanta. Na hladině významnosti α testujeme jednu z následujících tří hypotéz oproti příslušné alternativní hypotéze.

$$\begin{array}{lll} H_{01} : p_1 - p_2 = p_0 & \text{oproti} & H_{11} : p_1 - p_2 \neq p_0 \quad (\text{oboustranná alt.}) \\ H_{02} : p_1 - p_2 \leq p_0 & \text{oproti} & H_{12} : p_1 - p_2 > p_0 \quad (\text{pravostranná alt.}) \\ H_{03} : p_1 - p_2 \geq p_0 & \text{oproti} & H_{13} : p_1 - p_2 < p_0 \quad (\text{levostranná alt.}) \end{array}$$

Test nazýváme dvouvýběrovým Z -testem o rozdílu pravděpodobností $p_1 - p_2$. Testovací statistika má tvar

$$Z_W = \frac{M_1 - M_2 - p_0}{\sqrt{\frac{M_1(1-M_1)}{N_1} + \frac{M_2(1-M_2)}{N_2}}}, \quad (10.6)$$

kde $M_1 = \frac{1}{N_1} \sum_{i=1}^{N_1} X_{1i}$, kde X_{1i} nabývá hodnot 0 nebo 1, je výběrový průměr prvního náhodného výběru, $M_2 = \frac{1}{N_2} \sum_{i=1}^{N_2} X_{2i}$, kde X_{2i} nabývá hodnot 0 nebo 1, je výběrový průměr druhého náhodného výběru, N_1 je rozsah prvního náhodného výběru, N_2 je rozsah druhého náhodného výběru a p_0 je konstanta z nulové hypotézy. Testovací statistika Z_W pochází asymptoticky ze standardizovaného normálního rozdělení, tj.

$$Z_W = \frac{M_1 - M_2 - p_0}{\sqrt{\frac{M_1(1-M_1)}{N_1} + \frac{M_2(1-M_2)}{N_2}}} \stackrel{A}{\sim} N(0, 1).$$

To znamená, že pro malé rozsahy náhodných výběrů N_1 a N_2 nemusí mít rozdělení statistiky Z_W charakter standardizovaného normálního rozdělení. To je ale problém, neboť testování kritickým oborem, intervalem spolehlivosti i p -hodnotou je založeno na předpokladu, že Z_W má charakter standardizovaného normálního rozdělení. Proto pro malé rozsahy N_1 či N_2 náhodných výběrů nemůžeme statistiku Z_W k testování nulové hypotézy použít, neboť závěry takového testování by mohly být mylné. S rostoucími rozsahy náhodných výběrů N_1 a N_2 se však rozdělení statistiky Z_W čím dál více blíží ke standardizovanému normálnímu rozdělení, získává tak všechny jeho vlastnosti a závěry testování se stávají spolehlivě správnými.

Zda máme dostatečný počet pozorování k provedení testu o rozdílu pravděpodobností $p_1 - p_2$ prověříme podmínkami dobré aproximace, které mají obecně tvar

$$N_1 p_1 (1 - p_1) > 9 \quad \text{a} \quad N_2 p_2 (1 - p_2) > 9. \quad (10.7)$$

V případě, že skutečné hodnoty parametrů p_1 a p_2 neznáme, nahrazujeme je jejich nestrannými odhady, tj. výběrovými průměry M_1 a M_2 . Podmínky dobré aproximace potom vypadají následujícím způsobem

$$N_1 M_1 (1 - M_1) > 9 \quad \text{a} \quad N_2 M_2 (1 - M_2) > 9. \quad (10.8)$$

Pokud je podmínka 10.7 (resp. 10.8) dobré aproximace pro oba parametry p_1 i p_2 splněna, můžeme test o rozdílu pravděpodobností $p_1 - p_2$ použít, aniž bychom se vystavili riziku zavádějících výsledků testování. Pokud by však podmínka dobré aproximace pro jeden z parametrů nebo pro oba parametry p_1 a p_2 nebyla splněna, nemůžeme test použít, dokud nerozšíříme jeden nebo oba datové soubory o další pozorování a nezvýšíme tak dostatečně rozsahy N_1 i N_2 obou náhodných výběrů.

Poznámka: Problému s nedostatkem pozorování při použití testu o rozdílu pravděpodobností $p_1 - p_2$ se můžeme vyvarovat, pokud si ve fázi plánování experimentu spočítáme minimální potřebné rozsahy obou náhodných výběrů. Ve fázi sběru dat si potom již snadno ohlídáme, aby oba datové soubory obsahovaly potřebný počet pozorování, ideálně s nějakou rezervou. K výpočtu minimálního rozsahu náhodného výběru potřebujeme znát pouze předpokládanou hodnotu pravděpodobností p_1 a p_2 .

Kritický obor podle zvolené alternativní hypotézy má tvar

$$\begin{array}{ll} H_{11} : p_1 - p_2 \neq p_0 & W = (-\infty; u_{\alpha/2}) \cup (u_{1-\alpha/2}; \infty) \\ H_{12} : p_1 - p_2 > p_0 & W = (u_{1-\alpha}; \infty) \\ H_{13} : p_1 - p_2 < p_0 & W = (-\infty; u_{\alpha}) \end{array}$$

kde $u_{\alpha/2}$, $u_{1-\alpha/2}$, u_{α} , $u_{1-\alpha}$ jsou kvantily standardizovaného normálního rozdělení, jejichž hodnoty získáme pomocí \mathbb{R} a implementované funkce `qnorm()`.

Interval spolehlivosti má podle zvolené alternativní hypotézy jeden z následujících tvarů

$$\begin{array}{ll} H_{11} : p_1 - p_2 \neq p_0 & (d, h) = \left(m_1 - m_2 - \sqrt{\frac{m_1(1-m_1)}{N_1} + \frac{m_2(1-m_2)}{N_2}} u_{1-\alpha/2}; m_1 - m_2 - \sqrt{\frac{m_1(1-m_1)}{N_1} + \frac{m_2(1-m_2)}{N_2}} u_{\alpha/2} \right) \\ H_{12} : p_1 - p_2 > p_0 & (d, 1) = \left(m_1 - m_2 - \sqrt{\frac{m_1(1-m_1)}{N_1} + \frac{m_2(1-m_2)}{N_2}} u_{1-\alpha}; 1 \right) \\ H_{13} : p_1 - p_2 < p_0 & (-1, h) = \left(-1; m_1 - m_2 - \sqrt{\frac{m_1(1-m_1)}{N_1} + \frac{m_2(1-m_2)}{N_2}} u_{\alpha} \right) \end{array}$$

Poznámka: Protože parametry p_1 i p_2 značí pravděpodobnosti, platí, že $p_1 \in (0; 1)$ a $p_2 \in (0, 1)$. Rozdíl $p_1 - p_2$ nabývá tedy hodnoty z intervalu -1 až 1 , tj. $p_1 - p_2 \in (-1, 1)$. Proto levostranný interval spolehlivosti omezíme shora hodnotou 1 , namísto nekonečnem, a pravostranný interval spolehlivosti omezíme zdola hodnotou -1 , namísto mínus nekonečnem.

p -hodnota má v závislosti na zvolené alternativní hypotéze jeden z následujících tvarů

$$\begin{array}{ll} H_{11} : p_1 - p_2 \neq p_0 & p\text{-hodnota} = 2 \min\{\Pr(Z_W \leq z_W), \Pr(Z_W > z_W)\} \\ H_{12} : p_1 - p_2 > p_0 & p\text{-hodnota} = \Pr(Z_W > z_W) = 1 - \Pr(Z_W \leq z_W) \\ H_{13} : p_1 - p_2 < p_0 & p\text{-hodnota} = \Pr(Z_W \leq z_W) \end{array}$$

kde Z_W je náhodná veličina, z_W je realizace testovací statistiky Z_W (viz vzorec 10.6), tedy konkrétní číslo, a $\Pr(Z_W \leq z_W)$ je distribuční funkce standardizovaného normálního rozdělení, jejíž hodnotu získáme pomocí \mathbb{R} a implementované funkce `pnorm()`.

Příklad 10.17. Test o rozdílu parametrů p_1-p_2

Mějme k dispozici údaje o frekvenci výskytu epigenetického znaku *sutura metopica* (binomické proměnná) na lebkách jedinců z Anatolské populace (viz tabulka 1; (Eroglu, 2008)).

Tabulka 1: Výskyt epigenetického znaku *sutura metopica* u jedinců Anatolské populace

	<i>sutura metopica</i>	Σ
muži	26	334
ženy	15	153

Na hladině významnosti $\alpha = 0.01$ zjistěte, zda je frekvence výskytu epigenetického znaku *sutura metopica* u mužů Anatolské populace nižší než u žen Anatolské populace.

Řešení příkladu 10.17

Nejprve se zorientujme v zadání příkladu. Ze zadání víme, že proměnná popisující frekvenci výskytu epigenetického znaku *sutura metopica* je binárního typu, tj. nabývá pouze hodnoty 1 (zástupný symbol pro úspěch, což je v našem případě výskyt epigenetického znaku *sutura metopica*) a hodnoty 0 (neúspěch, což je v našem případě absence epigenetického znaku *sutura metopica*). Náhodná veličina X popisující frekvenci výskytu epigenetického znaku *sutura metopica* u mužů tedy pochází z alternativního rozdělení, tj. $X \sim \text{Alt}(p_1)$, kde p_1 je pravděpodobnost nastání úspěchu, tedy pravděpodobnost výskytu epigenetického znaku *sutura metopica* u mužů. Analogicky náhodná veličina Y popisující frekvenci výskytu epigenetického znaku *sutura metopica* u žen také pochází z alternativního rozdělení, tj. $X \sim \text{Alt}(p_2)$, kde p_2 je pravděpodobnost výskytu epigenetického znaku *sutura metopica* u žen.

Nejprve vytvoříme datový soubor pomocí příkazu `data.frame()`.

```
467 (data <- data.frame(sutura.metopica = c(26, 15), suma = c(334, 153),  
468 row.names = c('muzi', 'zeny')))
```

```
 sutura.metopica suma  
muzi            26  334  
zeny            15  153
```

469
470
471

Z tabulky vidíme, že z celkového počtu 334 lebek mužů byl zaznamenán výskyt epigenetického znaku *sutura metopica*) na 26 lebkách. Dále z celkového počtu 153 lebek žen byl zaznamenán výskyt epigenetického znaku *sutura metopica*) na 15 lebkách. Rozsah náhodného výběru lebek mužů $N_1 = 334$, rozsah náhodného výběru lebek žen $N_2 = 153$. Protože oba výběry pochází z alternativních rozdělení, použijeme pro ověření hypotézy ze zadání test o rozdílu parametrů $p_1 - p_2$. K použití tohoto testu je nejprve potřeba ověřit podmínku dobré aproximace zvlášť pro každý náhodný výběr. Jelikož přesné hodnoty parametrů p_1 a p_2 neznáme, musíme je v podmínce dobré aproximace nahradit jejich bodovými odhady, tedy hodnotami výběrových průměrů M_1 a M_2 .

$$\begin{aligned} m_1 &= \frac{1}{N_1} \sum_{i=1}^{N_1} x_i = \frac{1}{334} (1 + 1 + 1 + 1 + 1 + 1 + 1 + 0 + \dots + 0) \\ &= \frac{26}{334} = 0.07784431. \end{aligned}$$

$$\begin{aligned} m_2 &= \frac{1}{N_2} \sum_{i=1}^{N_2} x_i = \frac{1}{153} (1 + \dots + 1 + 0 + \dots + 0) \\ &= \frac{15}{153} = 0.09803922. \end{aligned}$$

```

472 N1 <- 334
473 N2 <- 153
474 x1 <- 26
475 x2 <- 15
476
477 m1 <- x1 / N1 # 0.07784431
478 m2 <- x2 / N2 # 0.09803922

```

Podmínky dobré aproximace potom dopočítáme následujícím způsobem

$$\begin{aligned}
N_1 M_1 (1 - M_1) &= 334 \times 0.07784431 \times (1 - 0.07784431) \\
&= 334 \times 0.07784431 \times 0.9221557 \\
&= 23.97605
\end{aligned}$$

```

479 N1 * m1 * (1 - m1) # 23.97605

```

Jelikož číslo $23.9761 > 5$, je podmínka dobré aproximace pro muže splněna.

$$\begin{aligned}
N_2 M_2 (1 - M_2) &= 153 \times 0.09803922 \times (1 - 0.09803922) \\
&= 153 \times 0.09803922 \times 0.9019608 \\
&= 13.52941
\end{aligned}$$

```

480 N2 * m2 * (1 - m2) # 13.52941

```

Jelikož číslo $13.5294 > 5$, je podmínka dobré aproximace pro ženy splněna.

Obě podmínky dobré aproximace jsou splněny, můžeme tedy přistoupit k testování otázky ze zadání. Zde poznamenejme, že v zadání příkladu není zmínka o znění nulové hypotézy, pouze o záměru zjistit, zda je frekvence výskytu znaku *sutura metopica* u mužů nižší než u žen. Toto tedy bude znění alternativní hypotézy, zatímco nulovou hypotézu musíme vhodně dodefinovat. Testování nulové hypotézy potom provedeme v posloupnosti šesti kroků.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : Frekvence výskytu epigenetického znaku *sutura metopica* u mužů je vyšší nebo rovna frekvenci znaku *sutura metopica* u žen.

H_1 : Frekvence výskytu epigenetického znaku *sutura metopica* u mužů je nižší než u žen.

- **matematická formulace** nulové a alternativní hypotézy

$H_0 : p_1 \geq p_2 \rightarrow p_1 - p_2 \geq p_0$, kde $p_0 = 0$

$H_1 : p_1 < p_2 \rightarrow p_1 - p_2 < p_0$, kde $p_0 = 0$

(levostranná alternativa)

2. Volba hladiny významnosti

- Hladina významnosti $\alpha = 0.01$ (viz zadání příkladu).

3. Testování kritickým oborem

- Testovací statistika

$$\begin{aligned}
 Z_W &= \frac{(M_1 - M_2) - p_0}{\sqrt{\frac{M_1(1-M_1)}{N_1} + \frac{M_2(1-M_2)}{N_2}}} \\
 &= \frac{(0.07784431 - 0.09803922) - 0}{\sqrt{\frac{0.07784431(1-0.07784431)}{334} + \frac{0.09803922(1-0.09803922)}{153}}} \\
 &= \frac{-0.02019491}{\sqrt{0.0002149239 + 0.0005779577}} \\
 &= \frac{-0.02019491}{0.02815815} = -0.7171959 \doteq -0.7172
 \end{aligned}$$

```

481 alpha <- 0.01
482 p0 <- 0
483 zw <- ((m1 - m2) - p0) / (sqrt(m1 * (1 - m1) / N1 + m2 * (1 - m2) / N2)) # -0.7171956

```

- Kritický obor

$$\begin{aligned}
 W &= (-\infty; u_\alpha) \\
 &= (-\infty; u_{0.01}) \\
 &= (-\infty; -2.3263)
 \end{aligned}$$

```

484 qnorm(alpha) # -2.326348

```

- Závěr testování

Protože realizace testovací statistiky $z_W = -0.7172$ nenáleží do kritického oboru, tj. $z_W \notin W$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

4. Testování intervalem spolehlivosti

- Interval spolehlivosti

$$\begin{aligned}
 (-1, h) &= \left(-1; m_1 - m_2 - \sqrt{\frac{m_1(1-m_1)}{N_1} + \frac{m_2(1-m_2)}{N_2}} u_\alpha \right) \\
 &= \left(-\infty; (0.07784431 - 0.09803922) - \sqrt{\frac{0.07784431(1-0.07784431)}{334} + \frac{0.09803922(1-0.09803922)}{153}} u_{0.01} \right) \\
 &= (-1; -0.02019491 - \sqrt{0.0002149239 + 0.0005779577} \times (-2.326348)) \\
 &= (-1; -0.02019491 - 0.02815815 \times (-2.326348)) \\
 &= (-1; 0.04531075) \doteq (-1; 0.04531)
 \end{aligned}$$

```

485 HH <- (m1 - m2) - sqrt(m1 * (1 - m1) / N1 + m2 * (1 - m2) / N2) * qnorm(alpha) #
0.04531075

```

- Závěr testování

Protože $p_0 = 0$ náleží do Waldova 99% empirického pravostranného intervalu spolehlivosti, tj. $p_0 = 0 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

5. Testování p -hodnotou

- p -hodnota

$$\begin{aligned} p\text{-hodnota} &= \Pr(Z_W \leq z_W) \\ &= \Pr(Z_W \leq -0.7171956) \\ &= 0.2366267 \doteq 0.2366 \end{aligned}$$

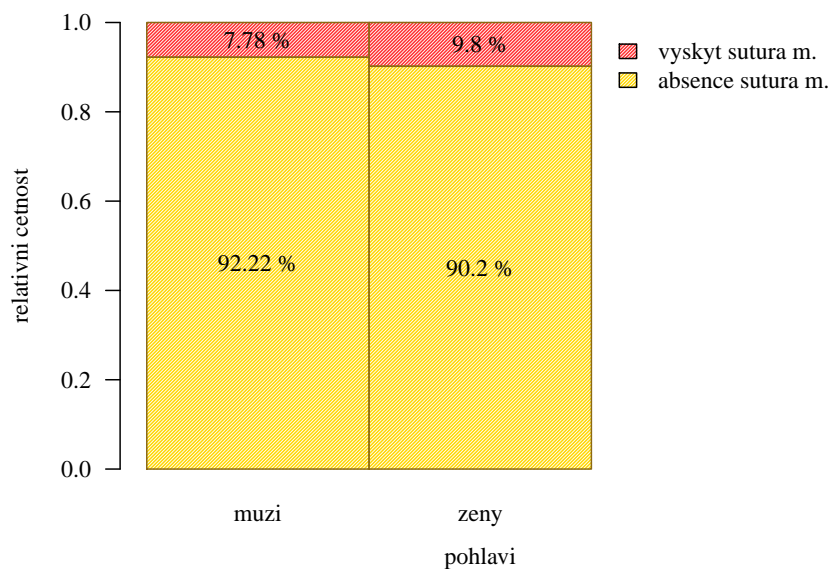
```
486 p.val <- pnorm(zw) # 0.2366267
```

- **Závěr testování**

Protože p -hodnota = 0.2366 je větší než $\alpha = 0.01$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

6. Grafická vizualizace výsledků testování

Výsledek testování vizualizujeme pomocí sloupcového diagramu relativních četností (viz obrázek 45), který vygenerujeme příkazem `rel.barplot.two()`. Funkce `rel.barplot.two()` je implementována v RSkriptu *Sbirka-AS-I-2018-funkce.R*. Před použitím samotné funkce je tedy potřeba načíst RSkript příkazem `source()`.



Obrázek 45: Sloupcový diagram relativních četností výskytu epigenetického znaku *sutura metopica* mužů a žen Anatolské populace

7. **Interpretace výsledků:** Na základě všech tří způsobů testování nezamítáme hypotézu H_0 . Frekvence výskytu epigenetického znaku *sutura metopica* u mužů není statisticky významně nižší než u žen.

★

Příklad 10.18. Test o rozdílu parametrů $p_1 - p_2$

Mějme datový soubor `20-more-samples-probabilities-pubis.txt` obsahující údaje o frekvenci výskytu tří stupňů změn kostního reliéfu (nepřítomné změny, malé změny, střední až výrazné změny) na vnitřní straně stydké kosti (*os pubis*) v blízkosti stydké spony sedací (*symphysis publica*) u žen z tří kosterních souborů (evropského původu, afrického původu a Inuitů). Pro více informací o datovém souboru viz sekci ???. Na hladině významnosti $\alpha = 0.05$ testujte hypotézu, že frekvence výskytu středních změn na stydké kosti u žen afrického původu je větší nebo rovna frekvenci výskytu středních změn na stydké kosti u inuitských žen.

Řešení příkladu 10.18

Datový soubor načteme příkazem `read.delim()` a odstraníme z něj NA hodnoty příkazem `na.omit()`.

```
487 data <- read.delim('00-Data/20-more-samples-probabilities-pubis.txt')
488 data <- na.omit(data)
489 head(data)
```

	origin	absence	trace.to.small	moderate.to.large	number.of.cases	
1	European	30	20	10	60	490
2	African	56	37	17	110	491
3	Inuits	16	6	13	35	492

Načtená tabulka obsahuje tři řádky, z nichž nás zajímají pouze řádky týkající se africké a inuitské populace (tj. řádek 2 a 3), a pět sloupců, z nichž nás zajímají pouze sloupce obsahující údaje o výskytu středních změn na vnitřní straně stydké kosti (`trace.to.small`) a celkový počet skeletů (`number.of.cases`). Pouze tyto údaje si vložíme do nové tabulky `data.AI`, a to pomocí operátoru `[]` a množinového operátoru `%in%`.

```
494 data.AI <- data[data$origin %in% c('African', 'Inuits'), c('trace.to.small', 'number.of.
      cases')]
495 row.names(data.AI) <- c('African', 'Inuits')
496 data.AI
```

	trace.to.small	number.of.cases	
African	37	110	497
Inuits	6	35	498

Mějme nyní náhodnou veličinu X popisující frekvenci výskytu středních změn na vnitřní straně stydké kosti v populaci Afrických žen. Tato veličina nabývá hodnoty 1 (úspěch; výskyt středních změn na vnitřní straně stydké kosti) nebo hodnoty 0 (neúspěch; výskyt jiného typu změn (výrazná nebo žádná změna) na vnitřní straně stydké kosti). Předpokládáme tedy, že náhodná veličina X pochází z alternativního rozdělení, tj. $X \sim \text{Alt}(p_1)$, kde p_1 je výskytu středních změn na vnitřní straně stydké kosti v populaci afrických žen. Analogicky mějme náhodnou veličinu Y popisující frekvenci výskytu středních změn na vnitřní straně stydké kosti v populaci Inuitek. Tato náhodná veličina pochází rovněž z alternativního rozdělení, tj. $Y \sim \text{Alt}(p_2)$, kde p_2 je pravděpodobnost výskytu středních změn na vnitřní straně stydké kosti v populaci Inuitek.

Z tabulky `data.AI` vidíme, že z celkového počtu 110 skeletů afrických žen byl zaznamenán výskyt středních změn na vnitřní straně stydké kosti u 37 skeletů. Z celkového počtu 35 skeletů Inuitek byl zaznamenán výskyt středních změn na vnitřní straně stydké kosti u 6 skeletů. Rozsah náhodného výběru skeletů afrických žen $N_1 = 110$, rozsah náhodného výběru skeletů Inuitek $N_2 = 35$. Protože oba výběry pochází z alternativních rozdělení, použijeme pro ověření hypotézy ze zadání test o rozdílu parametrů $p_1 - p_2$. Nejprve však ověříme podmínku dobré aproximace zvláště pro každý náhodný výběr. Přesné hodnoty parametrů p_1 a p_2 neznáme, nahradíme je proto v podmínce dobré aproximace jejich bodovými odhady, tedy hodnotami výběrových průměrů M_1 a M_2 .

$$m_1 = \frac{1}{N_1} \sum_{i=1}^{N_1} x_i = \frac{37}{110} = 0.3363636.$$

$$m_2 = \frac{1}{N_2} \sum_{i=1}^{N_2} x_i = \frac{6}{35} = 0.1714286.$$

Podmínky dobré aproximace potom dopočítáme následujícím způsobem

$$\begin{aligned} N_1 M_1 (1 - M_1) &= 110 \times 0.3363636 \times (1 - 0.3363636) \\ &= 110 \times 0.3363636 \times 0.6636364 \\ &= 24.55454 \doteq 24.5545 \end{aligned}$$

500 $N_1 * m_1 * (1 - m_1) \# 24.55455$

Jelikož číslo $24.5545 > 5$, je podmínka dobré aproximace pro populaci afrických žen splněna.

$$\begin{aligned} N_2 M_2 (1 - M_2) &= 35 \times 0.1714286 \times (1 - 0.1714286) \\ &= 35 \times 0.1714286 \times 0.8285714 \\ &= 4.971429 \doteq 4.9714 \end{aligned}$$

501 $N_2 * m_2 * (1 - m_2) \# 4.971429$

Jelikož číslo $4.9714 < 5$, není podmínka dobré aproximace pro populaci Inuitských žen splněna.

Jelikož podmínka dobré aproximace pro populaci Inuitských žen není splněna, nemůžeme otestovat nulovou hypotézu zadanou v zadání příkladu pomocí test o rozdílu parametrů $p_1 - p_2$. Připomeňme, že test o rozdílu parametrů $p_1 - p_2$ je asymptotickým testem, tj. testem, jehož výsledky jsou spolehlivé pouze pro dostatečné rozsahy obou náhodných výběrů, které nám zajišťuje právě splnění obou podmínek dobré aproximace.



Poznámka: Při analýze reálných dat bychom měli několik možností, jak se se situací podobnou situaci nastalé v příkladu 10.18 vypořádat. První možností je zamyslet se, zda by bylo možné doplnit datový soubor o další údaje. Doplnění údajů není v příkladě 10.18 preferováno, neboť data byla nasbírána v rámci studie (*Stewart, 1970*) a nasbírání dalších hodnot v podobné kvalitě na podobných skeletech by nebylo snadné. Druhou možností by bylo zvolit na otestování hypotézy ze zadání jiný test, který není tak striktní na splnění podmínek dobré aproximace. O takových testech si povíme více v kapitole ???. V praxi bychom však pravděpodobně volili postup, v rámci kterého bychom do striktně analytického postupu vložili špetku lidskosti, zamhouřili obě oči nad mírným porušením podmínky dobré aproximace a po řádném okomentování a zdůvodnění bychom test o rozdílu parametrů $p_1 - p_2$ na analýzu nulové hypotézy ze zadání příkladu 10.18 použili. Je však nezbytné pracovat s tímto postupem velmi opatrně, a při závažnějším porušení být jen jedné podmínky dobré aproximace volit raději jiné, vhodnější metody, nebo zauvažovat o rozšíření datového souboru.

V případě příkladu 10.18 jsme zůstali z důvodu větší názornosti striktní, test nulové hypotézy ze zadání jsme neprovedli a příklad jsme hned po zjištění porušení podmínky dobré aproximace ukončili.

Závěrem poznamenejme, že této nepříjemné situaci bychom se vyvarovali, kdybychom v rámci fáze plánování experimentu na základě podmínek dobré aproximace stanovili minimální rozsahy obou náhodných výběrů, a to analogicky, jako tomu bylo v příkladech ?? a ?? v kapitole ??.

Příklad 10.19. Test o rozdílu parametrů p_1-p_2

Mějme datový soubor 14-two-samples-probabilities-sexratio.txt. Nechť binární proměnná sex obsahuje údaje o pohlaví novorozenců a binární proměnná o.sib.N obsahuje údaje o počtu starších sourozenců novorozence (viz sekce ??). Na hladině významnosti $\alpha = 0.01$ testujte hypotézu, že pravděpodobnost narození chlapce je stejná u novorozenců s žádným sourozencem a u novorozenců s jedním sourozencem.

Řešení příkladu 10.19

Datový soubor načteme příkazem read.delim() a odstraníme z něj NA hodnoty příkazem na.omit().

```
502 data <- read.delim('00-Data/14-two-samples-probabilities-sexratio.txt')
503 data <- na.omit(data)
504 head(data)
```

	sex	o.sib.N
1	m	0
2	m	0
3	f	0
4	m	0
5	m	0
6	m	1

505
506
507
508
509
510
511

Načtená tabulka obsahuje dva sloupce. Sloupec sex, popisující pohlaví novorozence, disponuje dvěma variantami (m = mužské pohlaví a f = ženské pohlaví). Sloupec o.sib.N, popisující počet starších sourozenců, disponuje též dvěma variantami (0 = žádný starší sourozenec, 1 = jeden starší sourozenec). Datový soubor nyní rozdělíme pomocí operátoru [] na dva subsoubory. První bude obsahovat pouze údaje týkající se novorozenců s žádným starším sourozencem (data0) a druhý bude obsahovat údaje týkající se pouze novorozenců s jedním starším sourozencem (data1). Mějme nyní náhodnou veličinu X popisující frekvenci narození novorozenců mužského pohlaví s žádným

```
512 data0 <- data[data$o.sib.N == 0, ]
513 data1 <- data[data$o.sib.N == 1, ]
```

starším sourozencem. Tato veličina nabývá hodnoty 1 (úspěch; narození chlapce) nebo hodnoty 0 (neúspěch; narození děvčete). Předpokládáme tedy, že náhodná veličina X pochází z alternativního rozdělení, tj. $X \sim \text{Alt}(p_1)$, kde p_1 je pravděpodobnost nastání úspěchu, neboli pravděpodobnost narození chlapce u populace novorozenců s žádným starším sourozencem. Analogicky mějme náhodnou veličinu Y popisující frekvenci narození novorozenců mužského pohlaví s žádným starším sourozencem. Tato náhodná veličina pochází rovněž z alternativního rozdělení, tj. $Y \sim \text{Alt}(p_2)$, kde p_2 je pravděpodobnost narození chlapce v populaci novorozenců s jedním starším sourozencem.

Nyní zjistíme počet narozených chlapců v populaci novorozenců s žádným a s jedním starším sourozencem a rozsahy obou náhodných výběrů. Z tabulky vidíme, že z celkového počtu 595 novorozenců s žádným starším sourozencem

```
514 x1 <- sum(data0$sex == 'm')
515 x2 <- sum(data1$sex == 'm')
516 N1 <- length(data0$sex)
517 N2 <- length(data1$sex)
518 tab <- data.frame(x = c(x1, x2), N = c(N1, N2))
```

bylo 310 novorozenců mužského pohlaví a že z celkového počtu 518 novorozenců s jedním starším sourozencem bylo 277 novorozenců mužského pohlaví. Rozsah náhodného novorozenců s žádným sourozencem $N_1 = 595$, rozsah náhodného novorozenců s jedním starším sourozencem $N_2 = 518$. Protože oba výběry pochází z alternativních rozdělení, použijeme pro ověření hypotézy ze zadání test o rozdílu parametrů p_1-p_2 . Nejprve však ověříme podmínku dobré aproximace zvláště pro každý náhodný výběr. Protože přesné hodnoty parametrů p_1 a p_2 neznáme, nahradíme je v podmínce dobré aproximace jejich bodovými odhady, tedy hodnotami výběrových průměrů M_1 a M_2 .

$$m_1 = \frac{1}{N_1} \sum_{i=1}^{N_1} x_i = \frac{310}{595} = 0.5210084.$$

$$m_2 = \frac{1}{N_2} \sum_{i=1}^{N_2} x_i = \frac{277}{518} = 0.534749.$$

```
519 m1 <- x1 / N1 # 0.5210084
520 m2 <- x2 / N2 # 0.534749
```

Podmínky dobré aproximace potom dopočítáme následujícím způsobem

$$\begin{aligned} N_1 M_1 (1 - M_1) &= 595 \times 0.5210084 \times (1 - 0.5210084) \\ &= 595 \times 0.5210084 \times 0.4789916 \\ &= 148.4874 \end{aligned}$$

```
521 N1 * m1 * (1 - m1) # 148.4874
```

Jelikož číslo $148.4874 > 5$, je podmínka dobré aproximace pro populaci novorozenců s žádným starším sourozencem splněna.

$$\begin{aligned} N_2 M_2 (1 - M_2) &= 518 \times 0.534749 \times (1 - 0.534749) \\ &= 518 \times 0.534749 \times 0.465251 \\ &= 128.8745 \end{aligned}$$

```
522 N2 * m2 * (1 - m2) # 128.8745
```

Jelikož číslo $128.8745 > 5$, je podmínka dobré aproximace populaci novorozenců s jedním starším sourozencem splněna.

Obě podmínky dobré aproximace jsou splněny, můžeme tedy přistoupit k testování otázky ze zadání. V zadání máme uvedeno přesné znění nulové hypotézy, zbývá stanovit alternativní hypotézu tak, aby byla doplňkem k hypotéze nulové.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : Pravděpodobnost narození chlapce je shodná pro populaci novorozenců s žádným starším sourozencem a pro populaci s jedním starším sourozencem.

H_1 : Pravděpodobnost narození chlapce není shodná pro populaci novorozenců s žádným starším sourozencem a pro populaci s jedním starším sourozencem.

- **matematická formulace** nulové a alternativní hypotézy

$H_0 : p_1 = p_2 \rightarrow p_1 - p_2 = p_0$, kde $p_0 = 0$

$H_1 : p_1 \neq p_2 \rightarrow p_1 - p_2 \neq p_0$, kde $p_0 = 0$
(oboustranná alternativa)

2. Volba hladiny významnosti

- Hladina významnosti $\alpha = 0.01$ (viz zadání příkladu).

3. Testování kritickým oborem

- Testovací statistika

$$\begin{aligned}Z_W &= \frac{(M_1 - M_2) - p_0}{\sqrt{\frac{M_1(1-M_1)}{N_1} + \frac{M_2(1-M_2)}{N_2}}} \\&= \frac{(0.5210084 - 0.534749) - 0}{\sqrt{\frac{0.5210084(1-0.5210084)}{595} + \frac{0.534749(1-0.534749)}{518}}} \\&= \frac{-0.0137406}{\sqrt{0.0004194263 + 0.0004802944}} \\&= \frac{-0.0137406}{0.02999534} = -0.4580912 \doteq -0.45809\end{aligned}$$

```
523 alpha <- 0.01
524 p0 <- 0
525 zw <- ((m1 - m2) - p0) / (sqrt(m1 * (1 - m1) / N1 + m2 * (1 - m2) / N2)) # -0.4580921
```

- Kritický obor

$$\begin{aligned}W &= (-\infty; u_{\alpha/2}) \cup (u_{1-\alpha/2}; \infty) \\&= (-\infty; u_{0.01/2}) \cup (u_{1-0.01/2}; \infty) \\&= (-\infty; u_{0.005}) \cup (u_{0.975}; \infty) \\&= (-\infty; -2.5758) \cup (2.575829; \infty)\end{aligned}$$

```
526 qnorm(alpha / 2) # -2.575829
527 qnorm(1 - alpha / 2) # 2.575829
```

- Závěr testování

Protože realizace testovací statistiky $z_W = -0.45809$ nenáleží do kritického oboru, tj. $z_W \notin W$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

4. Testování intervalem spolehlivosti

- Interval spolehlivosti

$$\begin{aligned}(d, h) &= \left(m_1 - m_2 - \sqrt{\frac{m_1(1-m_1)}{N_1} + \frac{m_2(1-m_2)}{N_2}} u_{1-\alpha/2}; m_1 - m_2 - \sqrt{\frac{m_1(1-m_1)}{N_1} + \frac{m_2(1-m_2)}{N_2}} u_{\alpha/2} \right) \\&= \left(0.5210084 - 0.534749 - \sqrt{\frac{0.5210084(1-0.5210084)}{595} + \frac{0.534749(1-0.534749)}{518}} u_{1-0.01/2}; 0.5210084 - 0.534749 - \sqrt{\frac{0.5210084(1-0.5210084)}{595} + \frac{0.534749(1-0.534749)}{518}} u_{0.005} \right) \\&= (-0.0137406 - \sqrt{0.0004194263 + 0.0004802944} u_{0.975}; -0.0137406 - \sqrt{0.0004194263 + 0.0004802944} u_{0.005}) \\&= (-0.0137406 - 0.02999534 \times 2.575829; -0.0137406 - 0.02999534 \times (-2.575829)) \\&= (-0.09100347; 0.06352227)\end{aligned}$$

- Závěr testování

Protože $p_0 = 0$ náleží do Waldova 99% empirického oboustranného intervalu spolehlivosti, tj. $p_0 = 0 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

```

528 dh <- (m1 - m2) - sqrt(m1 * (1 - m1) / N1 + m2 * (1 - m2) / N2) * qnorm(1 - alpha / 2) #
      -0.09100352
529 hh <- (m1 - m2) + sqrt(m1 * (1 - m1) / N1 + m2 * (1 - m2) / N2) * qnorm(alpha / 2) #
      0.06352226

```

5. Testování p -hodnotou

- p -hodnota

$$\begin{aligned}
 p\text{-hodnota} &= 2 \min\{\Pr(Z_W \leq z_W), \Pr(Z_W > z_W)\} \\
 &= 2 \min\{\Pr(Z_W \leq z_W), 1 - \Pr(Z_W \leq z_W)\} \\
 &= 2 \min\{\Pr(Z_W \leq -0.4580921), 1 - \Pr(Z_W \leq -0.4580921)\} \\
 &= 2 \min\{0.3234431, 0.6765569\} \\
 &= 2 \times 0.32344310 \\
 &= 0.6468862 \doteq 0.6469
 \end{aligned}$$

```

530 p.val <- 2 * min(pnorm(zw), 1 - pnorm(zw)) # 0.6468863

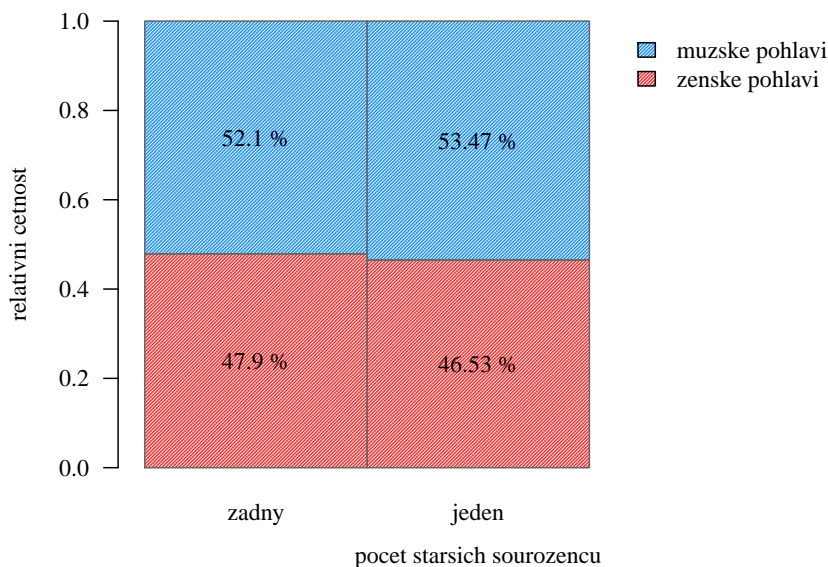
```

- Závěr testování

Protože p -hodnota = 0.6469 je větší než $\alpha = 0.01$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

6. Grafická vizualizace výsledků testování

, Výsledek testování doložíme sloupcovým diagramem relativních četností, jenž vykreslíme pomocí funkce `rel.barplot.two()` implementované v RSkriptu `Sbirka-AS-I-2018-funkce.R` (viz obrázek 46).



Obrázek 46: Sloupcový diagram relativních četností narození chlapce v populaci novorozenců s žádným starším sourozencem a s jedním starším sourozencem

7. **Interpretace výsledků:** Na základě všech tří způsobů testování nezamítáme hypotézu H_0 . Mezi pravděpodobností narození chlapce v populaci novorozenců s žádným starším sourozencem a s jedním starším sourozencem neexistuje statisticky významný rozdíl.



Příklad 10.20. Test o rozdílu parametrů p_1-p_2

Mějme datový soubor `26-two-samples-probabilities-palmar.txt` obsahující údaje o vysokém, středním, nízkém a jiném zakončení dlaňových linií na pravé a levé straně 50 žen z mechske populace a 87 žen z populace Rajbanshi (viz sekce ?? a tabulky 2 a 3). Na hladině významnosti $\alpha = 0.05$ zjistíte, zda je frekvence výskytu nízkého zakončení tří dlaňových linií na levé straně u žen z mechske populace vyšší než frekvence výskytu nízkého zakončení tří dlaňových linií na levé straně u žen z populace Rajbanshi.

Tabulka 2: Frekvence zakončení dlaňových linií u žen z mechske populace

f-mech	l	r	n
Hi	5	6	11
Mi	16	15	31
Lo	19	22	41
others	10	7	17

Tabulka 3: Frekvence zakončení dlaňových linií u žen z populace Rajbanshi

f-raj	l	r	n
Hi	27	37	64
Mi	17	26	43
Lo	21	10	31
others	22	14	36

Řešení příkladu 10.20

Z obou tabulek si vytáhneme pouze relevantní údaje, a sice počet výskytů nízkého zakončení tří dlaňových linií na levé straně u žen mechske populace ($x_1 = 19$) a počet výskytů nízkého zakončení tří dlaňových linií na levé straně u žen z populace Rajbanshi ($x_2 = 21$). Ze zadání příkladu dále víme, že rozsah náhodného výběru žen z mechske populace $N_1 = 50$ a rozsah náhodného výběru žen z populace Rajbanshi $N_2 = 87$.

```
531 x1 <- 19
532 x2 <- 21
533 N1 <- 50
534 N2 <- 87
535 tab <- data.frame(x = c(x1, x2), N = c(N1, N2))
```

Naším úkolem je porovnat frekvenci výskytu nízkého zakončení tří dlaňových linií na levé straně u dvou indických populací. U mechske populace máme k dispozici počet úspěchů, tj. počet výskytů nízkého zakončení tří dlaňových linií na levé straně. Náhodná veličina X popisující výskyt nízkého zakončení tří dlaňových linií na levé straně u žen z mechske populace pochází z alternativního rozdělení, tj. $X \sim \text{Alt}(p_1)$, kde p_1 je pravděpodobnost nastání úspěchu, tedy pravděpodobnost výskytu nízkého zakončení tří dlaňových linií na levé straně. Podobně u populace žen z Rajbanshi máme k dispozici počet výskytů nízkého zakončení tří dlaňových linií na levé straně. Náhodná veličina Y popisující výskyt nízkého zakončení tří dlaňových linií na levé straně u žen z populace Rajbanshi pochází z alternativního rozdělení, tj. $X \sim \text{Alt}(p_2)$, kde p_2 je pravděpodobnost výskytu nízkého zakončení tří dlaňových linií na levé straně. Řešení příkladu vede na situaci, kdy pravděpodobnost p_1 porovnááme s pravděpodobností p_2 , tedy na dvouvýběrový test o rozdílu $p_1 - p_2$. Před použitím testu je potřeba ověřit pro každý náhodný výběr zvlášť podmínku dobré aproximace. Protože přesné hodnoty parametrů p_1 a p_2 neznáme, nahradíme je v podmínce dobré aproximace jejich bodovými odhady, tedy hodnotami výběrových průměrů M_1 a M_2 .

$$m_1 = \frac{1}{N_1} \sum_{i=1}^{N_1} x_i = \frac{19}{50} = 0.38.$$

$$m_2 = \frac{1}{N_2} \sum_{i=1}^{N_2} x_i = \frac{21}{87} = 0.2413793.$$

```
536 m1 <- x1 / N1 # 0.38
537 m2 <- x2 / N2 # 0.2413793
```

Podmínky dobré aproximace potom dopočítáme následujícím způsobem

$$\begin{aligned} N_1 M_1 (1 - M_1) &= 50 \times 0.38 \times (1 - 0.38) \\ &= 50 \times 0.38 \times 0.62 \\ &= 11.78 \end{aligned}$$

```
538 N1 * m1 * (1 - m1) # 11.78
```

Jelikož číslo $11.78 > 5$, je podmínka dobré aproximace pro ženy z mechské populace splněna.

$$\begin{aligned} N_2 M_2 (1 - M_2) &= 87 \times 0.2413793 \times (1 - 0.2413793) \\ &= 87 \times 0.2413793 \times 0.7586207 \\ &= 15.93103 \end{aligned}$$

```
539 N2 * m2 * (1 - m2) # 15.93103
```

Jelikož číslo $15.93103 > 5$, je podmínka dobré aproximace pro ženy z populace Rajbanshi splněna.

Obě podmínky dobré aproximace jsou splněny, můžeme se tedy zaměřit na otázku v zadání. Naším úkolem je zjistit, zda je frekvence výskytu nízkého zakončení tří dlaňových linií na levé straně u žen z mechské populace vyšší než frekvence výskytu nízkého zakončení tří dlaňových linií na levé straně u žen z populace Rajbanshi. Protože v zadání není zmínka o nulové hypotéze, je tvrzení z předchozí věty zněním alternativní hypotézy. Zbývá stanovit nulovou hypotézu tak, aby byla doplňkem k alternativní hypotéze.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : frekvence výskytu nízkého zakončení tří dlaňových linií na levé straně u žen z mechské populace je menší nebo rovna frekvenci výskytu nízkého zakončení tří dlaňových linií na levé straně u žen z populace Rajbanshi.

H_1 : Frekvence výskytu nízkého zakončení tří dlaňových linií na levé straně u žen z mechské populace je vyšší než frekvence výskytu nízkého zakončení tří dlaňových linií na levé straně u žen z populace Rajbanshi.

- **matematická formulace** nulové a alternativní hypotézy

$H_0 : p_1 \leq p_2 \rightarrow p_1 - p_2 \leq p_0$, kde $p_0 = 0$

$H_1 : p_1 > p_2 \rightarrow p_1 - p_2 > p_0$, kde $p_0 = 0$

(pravostranná alternativa)

2. Volba hladiny významnosti

- Hladinu významnosti volíme v souladu se zadáním $\alpha = 0.05$.

3. Testování kritickým oborem

- Testovací statistika

$$\begin{aligned}
 Z_W &= \frac{(M_1 - M_2) - p_0}{\sqrt{\frac{M_1(1-M_1)}{N_1} + \frac{M_2(1-M_2)}{N_2}}} \\
 &= \frac{(0.38 - 0.2413793) - 0}{\sqrt{\frac{0.38(1-0.38)}{50} + \frac{0.2413793(1-0.2413793)}{87}}} \\
 &= \frac{0.1386207}{\sqrt{0.004712 + 0.002104774}} \\
 &= \frac{0.1386207}{\sqrt{0.006816774}} \\
 &= \frac{0.1386207}{0.08256376} \\
 &= 1.678953 \doteq 1.6790
 \end{aligned}$$

```

540 alpha <- 0.05
541 p0 <- 0
542 zw <- ((m1 - m2) - p0) / (sqrt(m1 * (1 - m1) / N1 + m2 * (1 - m2) / N2)) # 1.678953

```

- Kritický obor

$$\begin{aligned}
 W &= \langle u_{1-\alpha}; \infty \rangle \\
 &= \langle u_{1-0.05}; \infty \rangle \\
 &= \langle u_{0.95}; \infty \rangle \\
 &= \langle 1.6449; \infty \rangle
 \end{aligned}$$

```

543 qnorm(1 - alpha) # 1.644854

```

- Závěr testování

Protože realizace testovací statistiky $z_W = 1.6790$ náleží do kritického oboru, tj. $z_W \in W$, H_0 zamítáme na hladině významnosti $\alpha = 0.05$.

4. Testování intervalem spolehlivosti

- Interval spolehlivosti

$$\begin{aligned}
 (d, h) &= \left(m_1 - m_2 - \sqrt{\frac{m_1(1-m_1)}{N_1} + \frac{m_2(1-m_2)}{N_2}} u_{1-\alpha/2}; 1 \right) \\
 &= \left(0.38 - 0.2413793 - \sqrt{\frac{0.38(1-0.38)}{50} + \frac{0.2413793(1-0.2413793)}{87}} u_{1-0.05}; 1 \right) \\
 &= (0.1386207 - \sqrt{0.004712 + 0.002104774} \times 1.644854; 1) \\
 &= (0.1386207 - 0.08256376 \times 1.644854; 1) \\
 &= (0.0028153; 1)
 \end{aligned}$$

```
544 dh <- (m1 - m2) - sqrt(m1 * (1 - m1) / N1 + m2 * (1 - m2) / N2) * qnorm(1 - alpha) #
      0.002815394
```

- **Závěr testování**

Protože $p_0 = 0$ nenáleží do Waldova 95% empirického jednostranného intervalu spolehlivosti, tj. $p_0 = 0 \notin IS$, H_0 zamítáme na hladině významnosti $\alpha = 0.05$.

5. Testování p -hodnotou

- p -hodnota

$$p\text{-hodnota} = 2 \Pr(Z_W > z_W) = 1 - \Pr(Z_W \leq z_W) = \Pr(Z_W \leq 1.678953) = 0.04658058 \doteq 0.04658$$

```
545 p.val <- 1 - pnorm(zw) # 0.04658058
```

- **Závěr testování**

Protože p -hodnota = 0.04658 je menší než 0.05, H_0 zamítáme na hladině významnosti $\alpha = 0.05$.

6. Grafická vizualizace výsledků testování

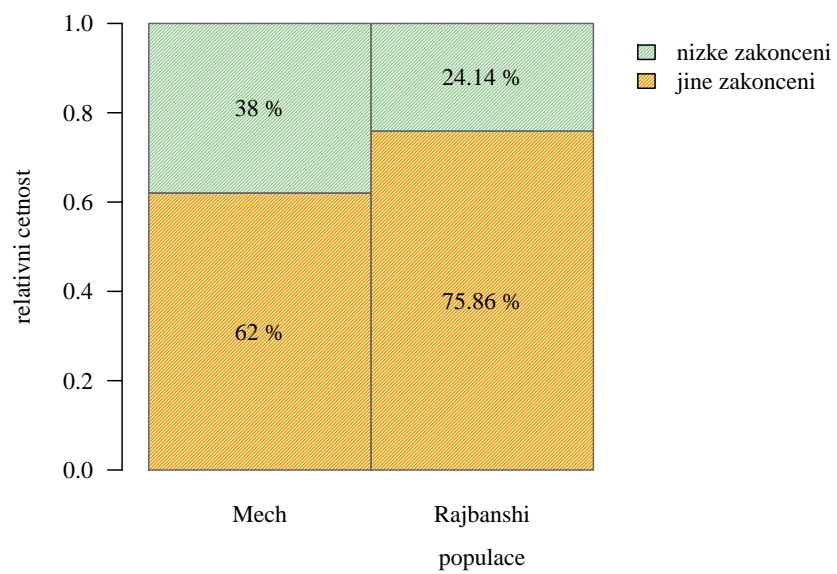
Rozdíly mezi oběma populacemi vizualizujeme sloupcovým diagramem relativních četností (viz obrázek 47).

```
546 source('Sbirka-AS-I-2018-funkce.R')
547 data <- cbind(c(19, N1 - 19), c(21, N2 - 21))
548 par(mar = c(5, 4, 1, 1), family = 'Times')
549 rel.barplot.two(data, col = c('orange2', 'darkseagreen3'),
550                    border = 'grey40',
551                    names = c('Mech', 'Rajbanshi'),
552                    legend = c('nizke zakonceni', 'jine zakonceni'),
553                    density = 60, col.number = 'black', main = '',
554                    xlab = 'populace', ylab = 'relativni cetnost')
```

7. Interpretace výsledků:

Na základě všech tří způsobů testování zamítáme hypotézu H_0 . Frekvence výskytu nízkého zakončení dlaňových linií na levé straně u žen z mečské populace je statisticky významně vyšší než frekvence výskytu nízkého zakončení dlaňových linií na levé straně u žen z populace Rajbanshi.





Obrázek 47: Sloupcový diagram relativních četností zakončení dlaňových linií z levé strany u žen z mečské populace a žen z populace v Rajbanshi