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 porovnáváme 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 dimorfismus (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 $X_{11}, \ldots, X_{1n_1}$. Pro druhou populaci vytvoříme reprezentativní vzorek změřením výšky postavy n_2 žen. Získáme tak druhý náhodný výběr $X_{21}, \ldots, X_{2n_2}$. 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. ★

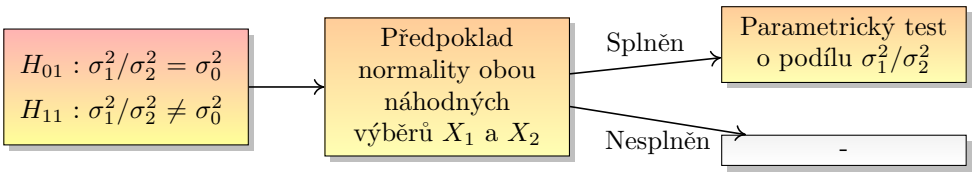
Dalším příkladem využití dvouvýběrových testů je například porovnání téhož znaku (šířka nadočnicového oblouku, délka levé klíční kosti, interorbitální šířka, apod.) u dvou populací (například malajské a čínské populace, německé a bantuské populace, či řecké a indické populace, apod.), nebo u dvou různých 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 $\mathrm{Alt}(p_1)$ a $\mathrm{Alt}(p_2)$.

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 (μ_1, μ_2, σ_1^2, σ_2^2, ρ_1, ρ_2, p_1, p_2) a na hladině významnosti α testujeme potom hypotézy o těchto vybraných parametrech.

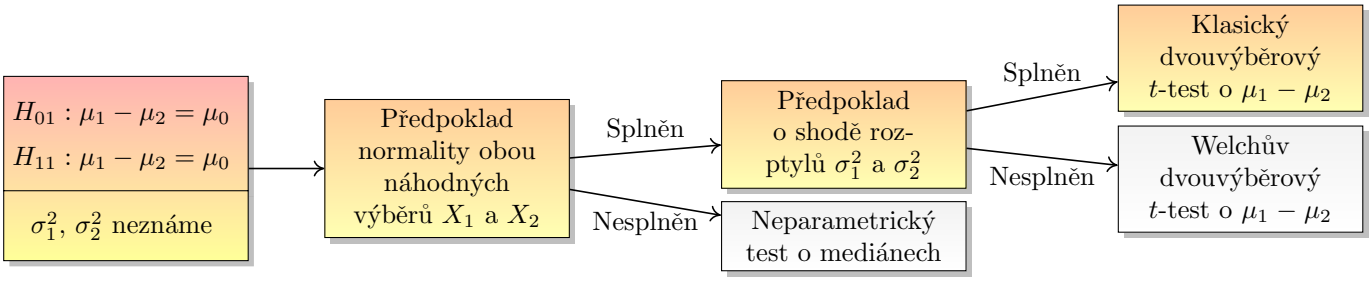
V případě testu o podílu rozptylů σ_1^2/σ_2^2 předpokládáme, že znak sledovaný v dané populaci pochází z normálního rozdělení a tedy také reprezentativní vzorky naměřených hodnot tohoto znaku v první i v druhé populaci pochází z normálního rozdělení. Proto je potřeba před samotným testem o podílu rozptylů σ_2^2/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



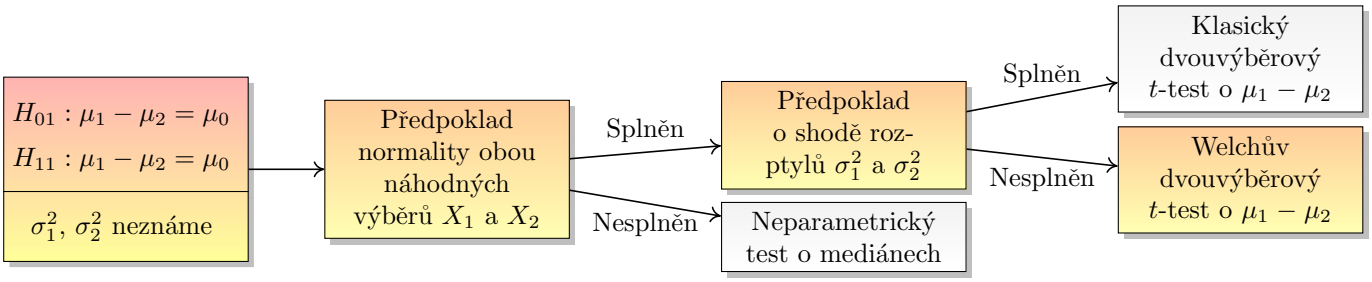
Obrázek 1: Rozhodovací cesta k použití parametrického testu o podílu rozptylů σ_1^2/σ_2^2

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 $\mu_1 - \mu_2$

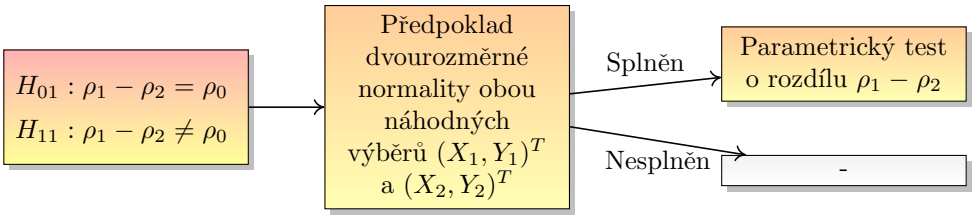
Pokud na základě testu o podílu rozptylů dojdeme k závěru, že rozptyly nejsou shodné, tj. $\sigma_1^2 \neq \sigma_2^2$, použijeme na otestování hypotézy o rozdílu středních hodnot Welchův dvouvýběrový t-test o rozdílu středních hodnot $\mu_1 - \mu_2$ (viz sekce 10.3). Pro lepší orientaci uvádíme rozhodovací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



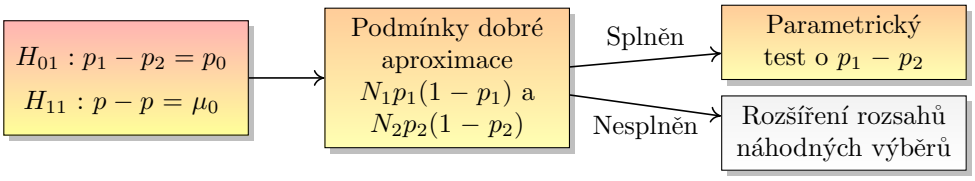
Obrázek 3: Rozhodovací cesta k použití Welchova dvouvýběrového t-testu o rozdílu středních hodnot $\mu_1 - \mu_2$

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 cestu 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 $X_{11}, \ldots, X_{1n_1}$ je náhodný výběr z $N(\mu_1, \sigma^2)$, a $X_{21}, \ldots, X_{2n_2}$ 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.

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

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

$$F_W = \frac{S_1^2}{S_2^2}, \tag{10.1}$$

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.

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

Kritický obor podle zvolené alternativní hypotézy má tvar

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

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 R a implementované funkce `qf()`.

Interval spolehlivosti má podle zvolené alternativní hypotézy jeden z následujících tvarů

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

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 mínus nekonečnem.

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

$H_{11} : \sigma_1^2/\sigma_2^2 \neq \sigma_0^2$	p-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-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-hodnota $= \Pr(F_W \leq f_W)$

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 R a implementované funkce `pf()`.

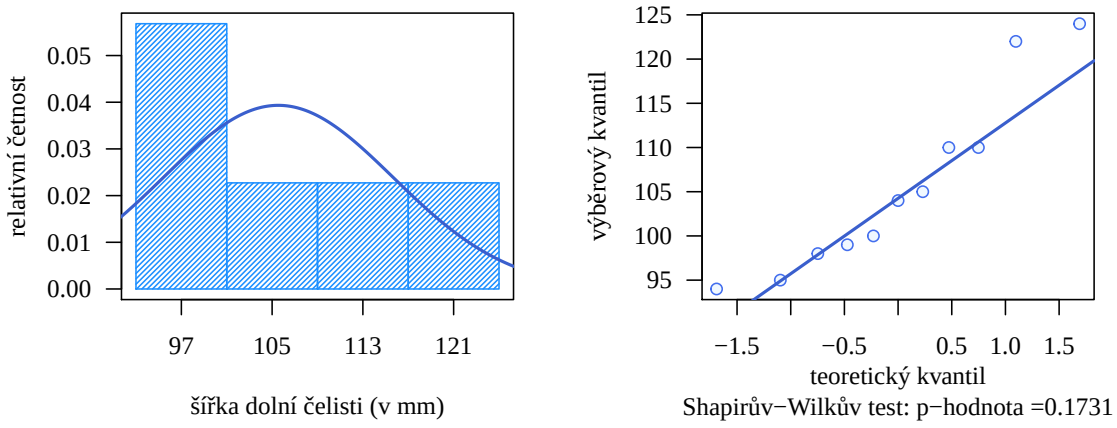
Příklad 10.2. Test o podílu rozptylů (oboustranná alternativa)
Mějme datový soubor 16-anova-head.txt a proměnnou bigo.W popisující šířku dolní čelisti v mm (viz sekce **??**). Předpokládejme, že náhodná veličina X, popisující *šířku dolní čelisti u mužů orientovaných jinak než výlučně heterosexuálně*, 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 Shapirovým-Wilkovým testem. Normalitu ověříme též graficky vykreslením QQ-diagramu (viz sekce **??**) a pomocí histogramu superponovaného křivkou normálního rozdělení (viz sekce **??**). Datový soubor rozdělíme na základě Sturgerova pravidla (viz sekce **??**) do čtyř ekvidistatních intervalů s šířkou 8 mm prostřednictvím stanovených hranic $93, 101, \ldots, 125$ (viz obrázek **??**).

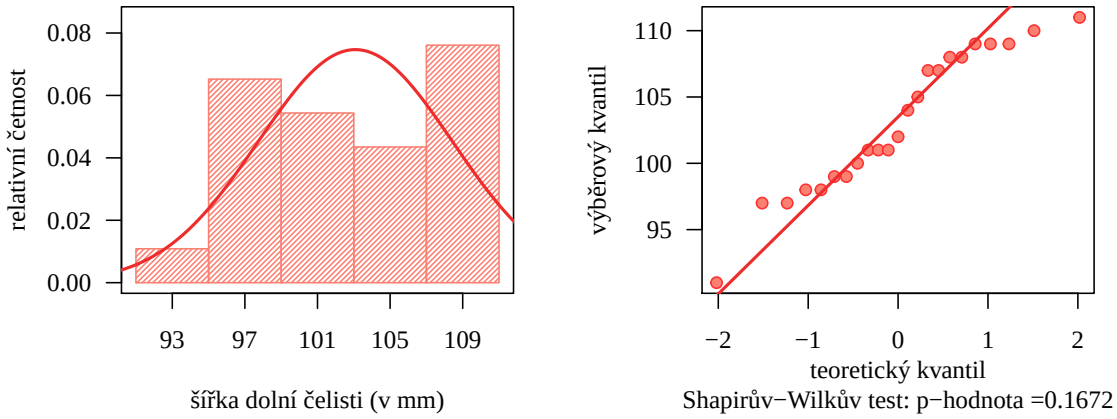


Obrázek 6: Histogram a QQ-diagram šířky dolní čelisti mužů

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

Shapirova-Wilkova 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 Shapirovým-Wilkovým testem. V rámci histogramu rozdělíme soubor do pěti ekvidistatních intervalů s šířkou 4 mm prostřednictvím stanovených hranic $91, 95, \ldots, 111$ (viz obrázek 7.



Obrázek 7: Histogram a QQ-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
     $H_0 : \sigma_1^2 = \sigma_2^2 \quad \rightarrow \quad \sigma_1^2/\sigma_2^2 = \sigma_0^2$, kde $\sigma_0^2 = 1$
     $H_1 : \sigma_1^2 \neq \sigma_2^2 \quad \rightarrow \quad \sigma_1^2/\sigma_2^2 \neq \sigma_0^2$, 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**

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

```
alpha <- 0.05
s1 <- sd(bigo.Wms)
s2 <- sd(bigo.Wfs)
Fw <- s1^2 / s2^2 # 3.60482
```

```
qf(alpha/2, n1 -1, n2 - 1) # 0.2950131
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**

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

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

   - **Závěr testování**
     Protože p-hdonota $= 0.01162$ 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 heterosexuálě 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')
```

```
17
18        F test to compare two variances
19
20 data:  bigo.Wms and bigo.Wfs
21 F = 3.6048, num df = 10, denom df = 22, p-value = 0.01162
22 alternative hypothesis: true ratio of variances is not equal to 1
23 95 percent confidence interval:
24   1.335211 12.219186
25 sample estimates:
26 ratio of variances
27            3.60482
```

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

★

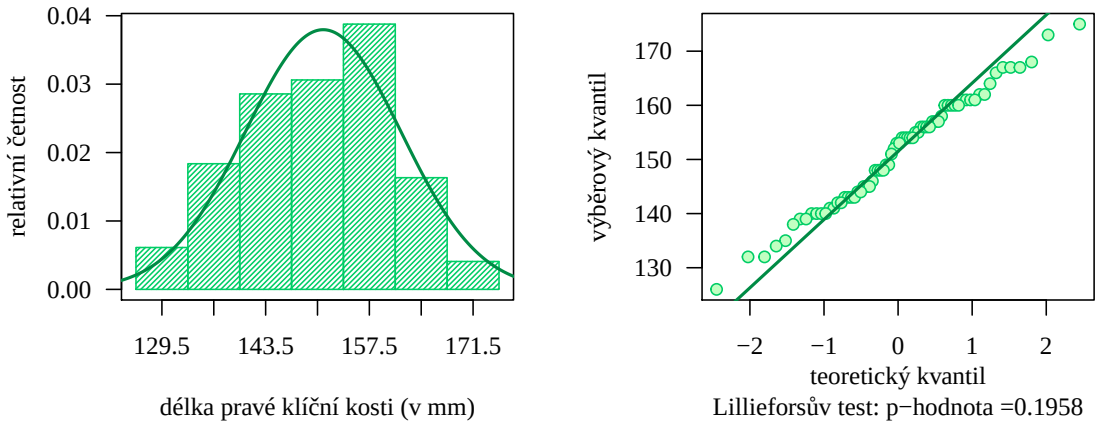
Příklad 10.3. Test o podílu rozptylů (pravostranná alternativa)
Mějme datový soubor 18-more-samples-variances-clavicle.txt a proměnnou cla.L popisující největší délku klíční kosti z pravé strany v mm (viz sekce **??**). Předpokládejme, že náhodná veličina X, popisující *největší délku klíční kosti z 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 $\alpha = 0.10$ zjistěte, zda je rozptyl největší délky klíční kosti z pravé strany u anglické populace statisticky významně větší než u indické populace z Amritsaru.

Řešení příkladu 10.3
Pomocí příkazu read.delim() načteme datový soubor a příkazem na.omit() odtraníme 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.

```
28 data <- read.delim('00-Data//18-more-samples-variances-clavicle.txt')
29 data <- na.omit(data)
30 #head(data, n = 3)
31 cla.Le <- data[data$pop == 'eng', 'cla.L']
32 cla.Li <- data[data$pop == 'ind1', 'cla.L']
33
34 n1 <- length(cla.Le) # 70
35 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.

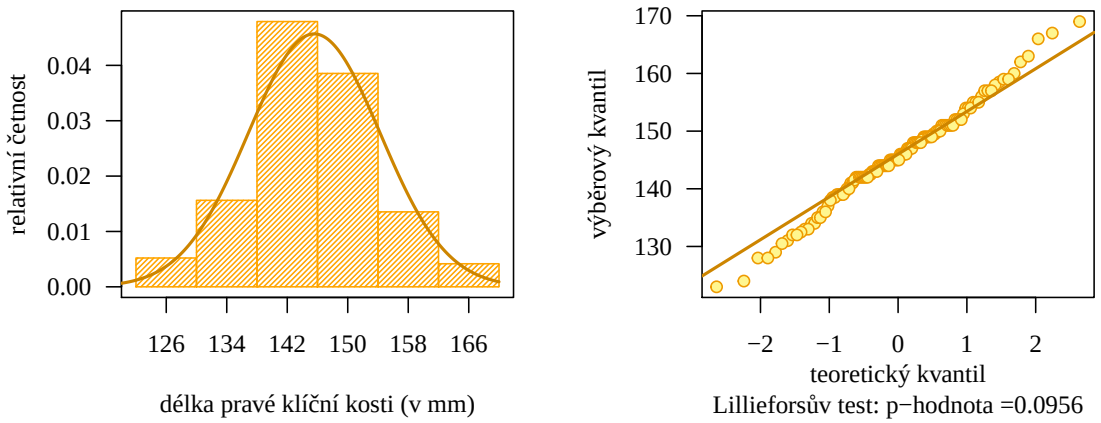
Hladinu významnosti α pro test normality stanovíme standartně, t. $\alpha = 0.05$. Nejprve testujeme hypotézu H_{01}: *Náhodný výběr největších délek klíčních kostí z pravé strany u jedinců anglické populace pochází z normálního rozdělení.* oproti alternativní hypotéze H_{11}: *Náhodný výběr největších 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 QQ-diagramu a histogramu superponovaného křivkou normálního rozdělení (viz obrázek **??**). Datový soubor rozdělíme do devíti ekvidistatních intervalů s šířkou 3 mm stanovením hranic $126, 133, \ldots, 175$.



Obrázek 8: Histogram a QQ-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 ekvidistatních intervalů s šířkou 8 mm prostřednictvím stanovených hranic $122, 130, \ldots, 170$ (viz obrázek E10-Test-ss-cla.LAI).



Obrázek 9: Histogram a QQ-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 Armitsaru. 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, zatímto 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
     $H_0 : \sigma_1^2 \leq \sigma_2^2 \quad \rightarrow \quad \sigma_1^2/\sigma_2^2 \leq \sigma_0^2$, kde $\sigma_0^2 = 1$
     $H_1 : \sigma_1^2 > \sigma_2^2 \quad \rightarrow \quad \sigma_1^2/\sigma_2^2 > \sigma_0^2$, 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**

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

```r
alpha <- 0.10
s1 <- sd(cla.Le)
s2 <- sd(cla.Li)
Fw <- s1^2 / s2^2 # 1.448242
```

- **Kritický obor**

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

```r
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**

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

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

```
[1] 1.107677
```

- **Závěr testování**
  Protože $\sigma_0^2 = 1$ nenáleží do Waldova 90% empirického levostranného 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-hdonota $= 0.0385$ 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.

```
43 p.hodnota <- 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.

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

```
45
46        F test to compare two variances
47
48 data:  cla.Le and cla.Li
49 F = 1.4482, num df = 69, denom df = 119, p-value = 0.03846
50 alternative hypothesis: true ratio of variances is greater than 1
51 90 percent confidence interval:
52  1.107677      Inf
53 sample estimates:
54 ratio of variances
55           1.448242
```

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

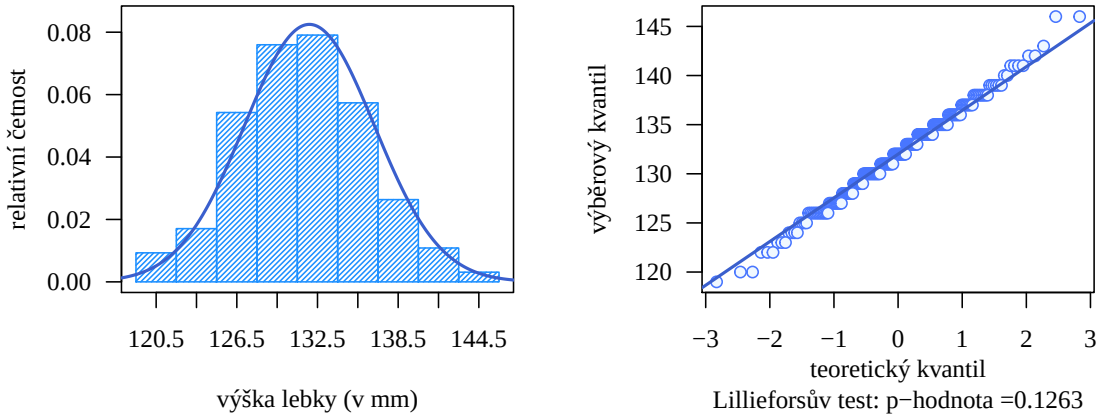
Příklad 10.4. Test o podílu rozptylů (levostranná alternativa)
Mějme datový soubor 11-two-samples-means-skull.txt a proměnnou (skull.H), popisující basion-bregmatickou výšku lebky v mm (viz sekce **??**). Předpokládejme, že náhodná veličina X, popisující *výšku lebky u žen*, pochází z normálního rozdělení $N(\mu_1, \sigma_1^2)$, 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 10.4
Pomocí příkazu read.delim() načteme datový soubor a příkazem na.omit() odtraníme 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.

```
56 data <- read.delim('00-Data//11-two-samples-means-skull.txt')
57 data <- na.omit(data)
58 #head(data, n = 3)
59 skull.Hm <- data[data$sex == 'm', 'skull.H']
60 skull.Hf <- data[data$sex == 'f', 'skull.H']
61
62 n1 <- length(skull.Hm) # 215
63 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 QQ-digramu a histogramu superponovaného křivkou normálního rozdělení (viz obrázek **??**). Datový soubor rozdělíme na základě Sturgerova pravidla do devíti ekvidistatních intervalů s šířkou 3 mm prostřednictvím stanovených hranic 119, 121, . . . , 146.

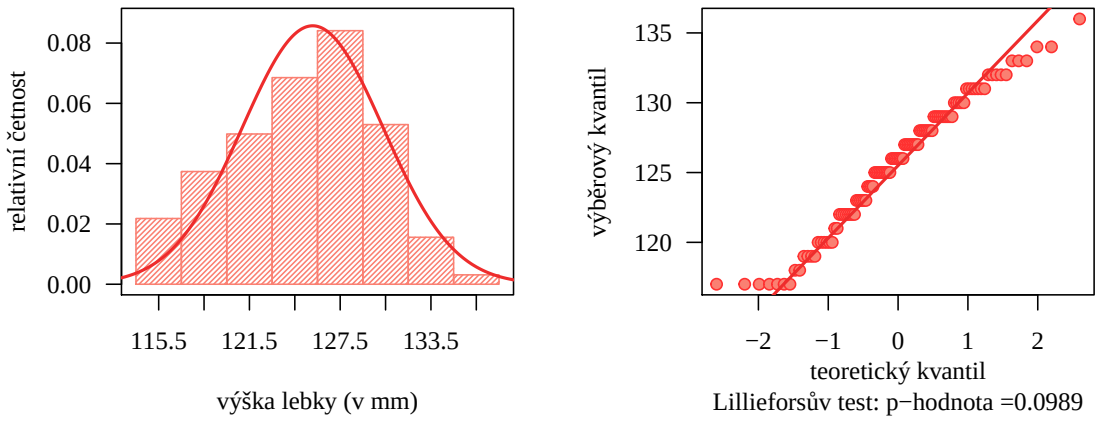


Obrázek 10: Histogram a QQ-diagram výšky lebky mužů

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 ekvidistatních intervalů s šířkou 3 mm prostřednictvím

stanovených hranic 114, 117, . . . , 138 (viz obrázek 11).



Obrázek 11: Histogram a QQ-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
     $H_0 : \sigma_1^2 \geq \sigma_2^2 \quad \rightarrow \quad \sigma_1^2/\sigma_2^2 \geq \sigma_0^2$, kde $\sigma_0^2 = 1$
     $H_1 : \sigma_1^2 < \sigma_2^2 \quad \rightarrow \quad \sigma_1^2/\sigma_2^2 < \sigma_0^2$, 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**

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

```
64 alpha <- 0.01
65 s1 <- sd(skull.Hm)
66 s2 <- sd(skull.Hf)
67 Fw <- s1^2 / s2^2 # 1.079861
```

   - **Kritický obor**

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

```r
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**

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

```r
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 pravostranného intervalu spolehlivosti, tj. $\sigma_0^2 = 1 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

5. **Testování p-hodnotou**

   - **p-hodnota**

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

```r
p.hodnota <- pf(Fw, n1 - 1, n2 - 1) # 0.3165
```

   - **Závěr testování**
     Protože p-hdonota $= 0.3165$ 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.

```
        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
```

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

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

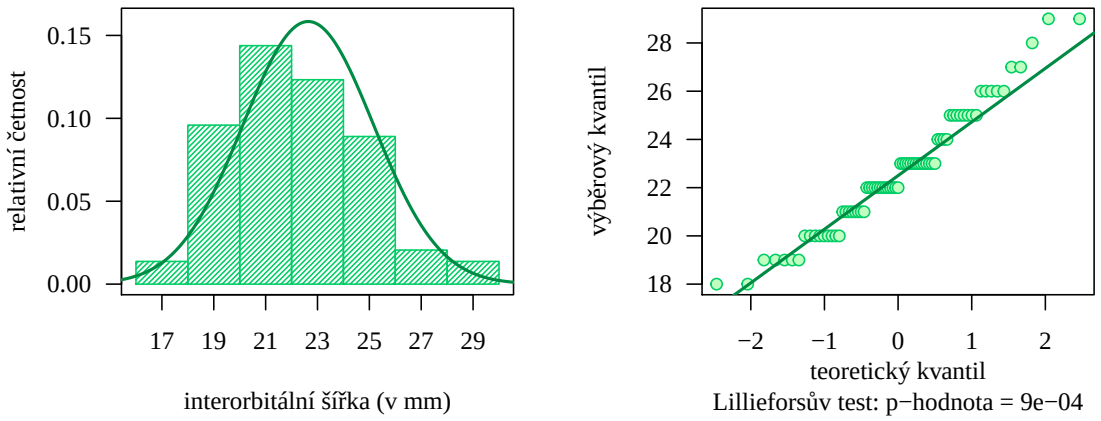
Příklad 10.5. Test o podílu rozptylů (nesplnění předpokladů)
Mějme datový soubor 19-more-samples-correlations-skull.txt a proměnnou intorb.B popisující interorbitální šířku v mm (viz sekce **??**). Na hladině významnosti $\alpha = 0.01$ zjistěte, zda je rozptyl interorbitální šířky mužů malajské populace 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 $\alpha = 0.05$.

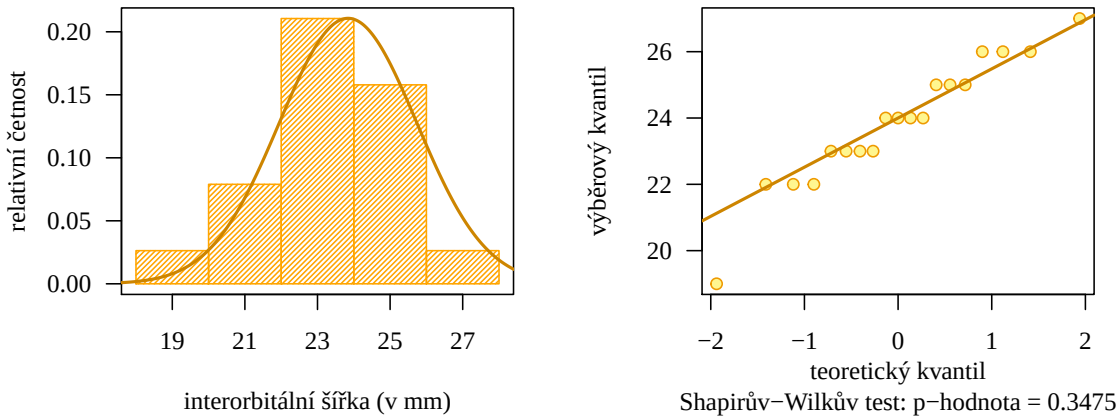
Protože rozsah náhodného výběru interorbitálních šířek mužů malajské populace je větší než 30, použijeme na ověření předpokladu normality Lillieforsův test. Rozdělení datového souboru vizualizujeme pomocí QQ-diagramu a histogramu (viz obrázek 12). Datový soubor rozdělíme do sedmi ekvidistatních intervalů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



Obrázek 12: Histogram a QQ-diagram interorbitální šířky mužů malajské populace

Jelikož p-hodnota $= 9 \times 10^{-4}$ je menší než 0.05, nulovou hypotézu zamítáme na hladině významnosti $\alpha = 0.05$. Takéž histogram ukazuje porušení předpoklaud normality zejména vyšikmením naměřených hodnot doleva oproti předpokládanému rozdělení. V QQ-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 Shapirovým-Wilkovým testem. V rámci histogramu rozdělíme data do pěti ekvidistatních intervalů s šířkou 2 mm prostřednictvím stanovených hranic $18, 20, \ldots, 28$ (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 příchylnost bodů k referenční přímce vyjma krajného levého bodu. Odchýlení tohoto bodů však



Obrázek 13: Histogram a QQ-diagram interorbitální šířky mužů čínské populace

podle Shapirova-Wilkova testu nenarušuje fatálně předpoklad normality náhodného výběru. Náhodný výběr interorbitálních šířek u mužů čínské populace pochází z normálnho 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 $X_{11}, \ldots, X_{1n_1}$ je náhodný výběr z $N(\mu_1, \sigma_1^2)$, a $X_{21}, \ldots, X_{2n_2}$ 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.

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

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

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

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.

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

Kritický obor podle zvolené alternativní hypotézy má tvar

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

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 a implementované funkce `qt()`.

Interval spolehlivosti má podle zvolené alternativní hypotézy jeden z následujících tvarů

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

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

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

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 a implementované funkce `pt()`.

Příklad 10.6. Test o rozdílu středních hodnot (oboustranná alternativa)
Mějme datový soubor 18-more-samples-variances-clavicle.txt obsahující údaje o délce klíční kosti z pravé strany (cla.L) u mužů z indické populace z Amritsaru, z indické populace z Varanasi a z řecké populace z Atén (viz sekce **??**). Na hladině významnosti $\alpha = 0.10$ zjistěte, zda existuje rozdíl mezi délkou klíční kosti z pravé strany u mužů z indické populace z Amritsaru a u mužů z indické populace z Varanasi.

Řešení příkladu ??
Datový soubor 18-more-samples-variances-clavicle.txt načteme příkazem read.delim(). Následně z načtených dat vybereme naměřené hodnoty délky klíční kosti z pravé stranny (cla.L) u mužů z indické populace z Amritsaru (pop == 'ind1') a vložíme je do proměnné cla.LA. Dále z datové tabulky vybereme naměřené hodnoty délky klíční kosti z pravé strany u mužů z indické populace z Varanasi (pop == 'ind2') a vložíme je do proměnné cla.LV. Z obou proměnných cla.LA a cla.LV odstraníme příkazem na.omit() chybějící pozorování, dále příkazem length() zjistíme počet pozorování v obou náhodných výběrech a příkazem range() rozpětí naměřených hodnot.

```
90  data <- read.delim('00-Data//18-more-samples-variances-clavicle.txt')
91  cla.LA <- data[data$pop == 'ind1', 'cla.L']
92  cla.LA <- na.omit(cla.LA)
93  cla.LV <- data[data$pop == 'ind2', 'cla.L']
94  cla.LV <- na.omit(cla.LV)
95  n1 <- length(cla.LA) # 120
96  n2 <- length(cla.LV) # 81
97  range(cla.LA) # 123-169
98  range(cla.LV) # 126-165
```

Datový soubor obsahuje naměřené hodnoty délky klíční kosti z pravé strany u 120 mužů z indické populace z Amritsaru u 81 mužů z indické populace z Varanasi. Naměřené délky klíčních kostí z pravé strany u mužů z Amritsaru nabývají hodnot v rozmezí od 123 mm do 169 mm, naměřené délky klíčních kosti z pravé strany u mužů Varanasi nabývají hodnot v rozmezí od 126 mm do 165 mm.

Řešení příkladu vede na test o rozdílu středních hodnot. Před testováním nulové hypotézy ze zadání musíme ověřit splnění předpokladu normalitu 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 délek klíčních kostí z pravé strany u mužů z indické populace z Amritsaru pochází z normálního rozdělení.* oproti alternativní hypotéze: H_{11} : *Náhodný výběr délek klíčních kostí z pravé strany u mužů z indické populace z Amritsaru nepochází z normálního rozdělení.* Jelikož náhodný výběr sestává ze 120 pozorování, což je více než 30, provedeme test normality Lillieforsovým testem. Normalitu ověříme též graficky vykreslením QQ-diagramu (viz sekce **??**) a pomocí histogramu superponovaného křivkou normálního rozdělení (viz sekce **??**). Datový soubor rozdělíme na základě Sturgerova pravidla (viz sekce **??**) do osmi ekvidistatních intervalů s šířkou 6 mm prostřednictvím stanovených hranic $122, 128, \ldots, 170$ (viz obrázek **??**).
Protože p-hodnota $= 0.2702$ je větší než 0.05, nulovou hypotézu nezamítáme na hladině významnosti $\alpha = 0.05$. Ani v histogramu ani v QQ-diagramu potom nepozorujeme žádné výrazné porušení předpokladu normality. Náhodný výběr délek klíčních kostí z pravé strany u mužů z populace z Amritsaru pochází z normálního rozdělení.

Analogicky testujeme hypotézu H_{02} : *Náhodný výběr délek klíčních kostí z pravé strany u mužů z indické populace z Varanasi pochází z normálního rozdělení.* oproti alternativní hypotéze: H_{12} : *Náhodný výběr délek klíčních kostí z pravé strany u mužů z indické populace z Varanasi nepochází z normálního rozdělení.* Jelikož rozsah náhodného výběru je dostatečně vysoký, použijeme na otestování hypotézy o normalitě Lillieforsův test. Pro účely grafické vizualizace pomocí histogramu rozdělíme náhodný výběr do 7 ekvidistantních třídicích intervalů o šířce 6 mm prostřednictvím stanovených hranic 125, 131, ..., 167 mm (viz obrázek **??**).

Protože p-hodnota Lillieforsova testu $p = 0.0274$ je menší než 0.05, hypotézu H_{02} zamítáme na hladině významnosti $\alpha = 0.05$. Zamítnutí normality je však hraniční, protože p-hodnota neklesla pod hladinu významnosti $\alpha = 0.01$. Proto

pro kontrolu otestujeme normalitu také pomocí Andersonova-Darlingova testu. Protože p-hodnota Andersonova-Darlingova testu $p = 0.0628$ je větší než 0.05, hypotézu H_{02} nezamítáme na hladině významnosti $\alpha = 0.05$. Andersonův-Darlingův test tedy hraničně normalitu náhodného výběru nezamítá. Bezpochyby se rozdělení náhodného výběru délky klíční kosti z pravé strany u mužů z indické populace z Varanasi pohybuje na pomezí normality. Grafické zhodnocení normality je v tomto případě velmi subjektivní. Z histogramu je patrná normalita náhodného výběru, nicméně je zde viditelné také kladné vyšikmení hodnot s prodlouženým pravým koncem. QQ-diagram potom zobrazuje stejnou vlastnost prostřednictvím bodů v pravém horním rohu, které se odchylují od referenční přímky. Situace je tedy opravdu hraniční a záleží pouze na nás, jak se o normalitě náhodného výběru rozhodneme. Zde se zachováme tak, jak bychom se pravděpodobně zachovali také v praxi. Obecně je naším zájmem, pokud to předpoklad normality dovolí, použít na otestování libovolné hypotézy parametrický test, neboť je silnější (s větší pravděpodobností zamítá H_0, pokud H_0 neplatí). Proto se v tomto hraničním případě přikloníme k výsledku Andersonova-Darlingova testu, že náhodný výběr délek klíčních kostí z pravé strany u mužů z indické populace z Varanasi pochází z normálního rozdělení.
Jelikož tedy 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.

```
99 var.test(cla.LA, cla.LV, alternative = 'two.sided')$p.val # 0.5652848
```

```
[1] 0.5652848                                                                    100
```

Jelikož p-hodnota $= 0.5653$ 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 hodnoty délky klíční z pravé strany u mužů z Amritsaru a z Varanasi jsou shodné.*
     H_1 : *Střední hodnoty délky klíční z pravé strany u mužů z Amritsaru a z Varanasi nejsou shodné.*
   - **matematická formulace** nulové a alternativní hypotézy
     $H_0 : \mu_1 = \mu_2 \quad \rightarrow \quad \mu_1 - \mu_2 = \mu_0$, kde $\mu_0 = 0$
     $H_1 : \mu_1 \neq \mu_2 \quad \rightarrow \quad \mu_1 - \mu_2 \neq \mu_0$, kde $\mu_0 = 0$
     (oboustranná alternativa)

2. **Volba hladiny významnosti**

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

3. **Testování kritickým oborem**

- **Aritmetický průměr výběrových rozptylů**

$$
\begin{aligned}
S_* &= \sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}} \\
&= \sqrt{\frac{(120-1)\times 8.733432^2 + (81-1)\times 8.220209^2}{120+81-2}} \\
&= \sqrt{\frac{119\times 76.27283 + 80\times 67.57184}{199}} \\
&= \sqrt{\frac{14482.21}{199}} \qquad\qquad = \sqrt{72.77492} = 8.530822 \doteq 8.53082
\end{aligned}
$$

```
s1 <- sd(cla.LA) # 8.733432
s2 <- sd(cla.LV) # 8.220209
sh <- sqrt(((n1 - 1) * s1^2 + (n2 - 1) * s2^2) / (n1 + n2 - 2)) # 8.530823
```

- **Testovací statistika**

$$
\begin{aligned}
T_W &= \frac{(M_1 - M_2) - \mu_0}{S_*\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \\
&= \frac{(145.5667 - 141.4938) - 0}{8.530823\sqrt{\frac{1}{120}+\frac{1}{81}}} \\
&= \frac{4.0729}{8.530823\times 0.143802} \\
&= \frac{4.0729}{1.226749} = 3.320076
\end{aligned}
$$

```
alpha <- 0.10
mu0 <- 0
m1 <- mean(cla.LA) # 145.5667
m2 <- mean(cla.LV) # 141.4938
tw <- ((m1 - m2) - mu0) / (sh * sqrt(1 / n1 + 1 / n2)) # 3.320026
```

Kritický obor

$$
\begin{aligned}
W &= (-\infty\,;\, t_{n_1+n_2-2}(\alpha/2)\rangle \cup \langle t_{n_1+n_2-2}(1-\alpha/2)\,;\,\infty) \\
&= (-\infty\,;\, t_{120+81-2}(0.10/2)\rangle \cup \langle t_{120+81-2}(1-0.10/2)\,;\,\infty) \\
&= (-\infty\,;\, t_{199}(0.05)\rangle \cup \langle t_{199}(0.95)\,;\,\infty) \\
&= (-\infty\,;\, -1.652547\rangle \cup \langle 1.652547\,;\,\infty)
\end{aligned}
$$

```
qt(alpha / 2, n1 + n2 - 2) # -1.652547
qt(1 - alpha / 2, n1 + n2 - 2) # 1.652547
```

- **Závěr testování**
  Protože realizace testovací statistiky $t_W = 3.320026$ náleží do kritického oboru, tj. $t_W \in W$, H_0 zamítáme na hladině významnosti $\alpha = 0.10$.

4. **Testování intervalem spolehlivosti**

- **Interval spolehlivosti**

$$
\begin{aligned}
(d,h) &= \left(m_1 - m_2 - s_* \sqrt{\frac{1}{n1} + \frac{1}{n_2}} t_{n_1+n_2-2}(1-\alpha/2)\,;\, m_1 - m_2 - s_* \sqrt{\frac{1}{n1} + \frac{1}{n_2}} t_{n_1+n_2-2}(\alpha/2)\right) \\
&= ((145.5667 - 141.4938) - 8.530823\sqrt{\frac{1}{120} + \frac{1}{81}} t_{120+81-2}(0.95)\,;\, (145.5667 - 141.4938) - 8.530823\sqrt{\frac{1}{120} + \frac{1}{81}} \\
&= (4.0729 - 8.530823\sqrt{0.02067901} t_{199}(0.95)\,;\, (4.0729 - 8.530823\sqrt{0.02067901} t_{199}(0.05)) \\
&= (4.0729 - 8.530823 \times 0.143802 \times 1.652547\,;\, 4.0729 - 8.530823 \times 0.143802 \times (-1.652547)) \\
&= (2.045639\,;\, 6.100161) \doteq (2.04564\,;\, 6.100161)
\end{aligned}
$$

```
dh <- (m1 - m2) - sh * sqrt(1 / n1 + 1 / n2) * qt(1 - alpha / 2, n1 + n2 - 2) # 2.045579
hh <- (m1 - m2) - sh * sqrt(1 / n1 + 1 / n2) * qt(alpha / 2, n1 + n2 - 2) # 6.1001
```

- **Závěr testování**
  Protože $\mu_0 = 0$ nenáleží do Waldova 90% empirického oboustranného intervalu spolehlivosti, tj. $\mu_0 = 0 \notin IS$, H_0 zamítáme na hladině významnosti $\alpha = 0.10$.

5. **Testování p-hodnotou**

   - **p-hodnota**

$$
\begin{aligned}
p\text{-hodnota} &= 2\min\{\Pr(T_W \leq t_W)\,,\, \Pr(T_W > t_W)\} \\
&= 2\min\{\Pr(T_W \leq 3.320026)\,,\, 1 - \Pr(T_W \leq 3.320026)\} \\
&= 2\min\{0.9994647, 0.0005352937\} \\
&= 2 \times (0.0005352937) \\
&= 0.001070587 \doteq 0.00107059
\end{aligned}
$$

   Výpočet p-hodnoty provedeme také pomoc softwaru R , pomocí příkazu pt(), který vrátí hodnotu distribuční funkce Studentova rozdělení o $n_1 + n_2 - 2$ stupních volnosti v hodnotě testovací statistiky t_W. Vstupními argumenty funkce pt() budou tedy neprve hodnota testovací statistiky t_W a následně argument df = n1 + n2 - 2 definující počet stupňů volnosti Studentova rozdělení.

```
p.hodnota <- 2 * min(pt(tw, df = n1 + n2 - 2), 1 - pt(tw, df = n1 + n2 - 2)) #
    0.001070587
```

   - **Závěr testování**
     Protože p-hodnota $= 0.00107059$ je menší než $\alpha = 0.10$, H_0 zamítáme na hladině významnosti $\alpha = 0.10$.

6. **Grafická vizualizace výsledků testování**
   Vhodný grafem porovnávajícím střední hodnoty obou náhodných výběrů je krabicový diagram (obrázek **??**).

7. **Interpretace výsledků:** Na základě všech tří způsobů testování zamítáme hypotézu o shodě středních hodnot délky klíční kosti z pravé strany u mužů z indické populace z Amritsaru a u mužů z populace z Varanasi. Mezi délkou klíční kosti z pravé strany u mužů z indické populace z Amritsaru a u mužů z populace z Varanasi existuje statisticky významný rozdíl.

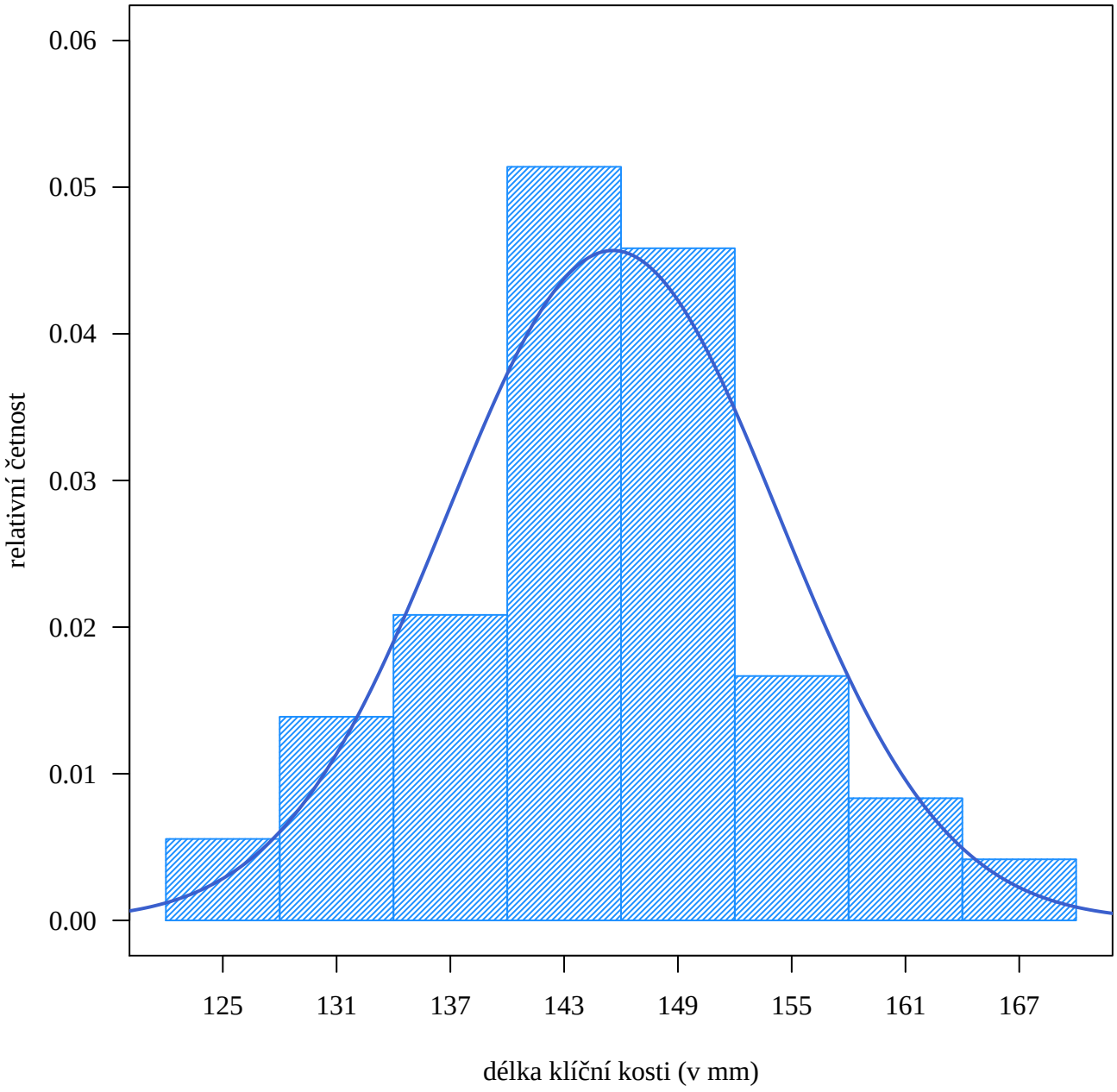
Poznámka: K otestování nulové hypotézy o rozdílu středních hodnot můžeme využít funkci t.test(). Vstupními parametry budou nejprve dva vektory reprezentující náhodné výběry, tj. cla.LA a cla.LV, dále hodnota hladiny významnosti α zadaná prostřednictvím koeficientu spolehlivosti $1 - \alpha$ nastavením hodnoty argumentu conf.level = 0.90, 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.

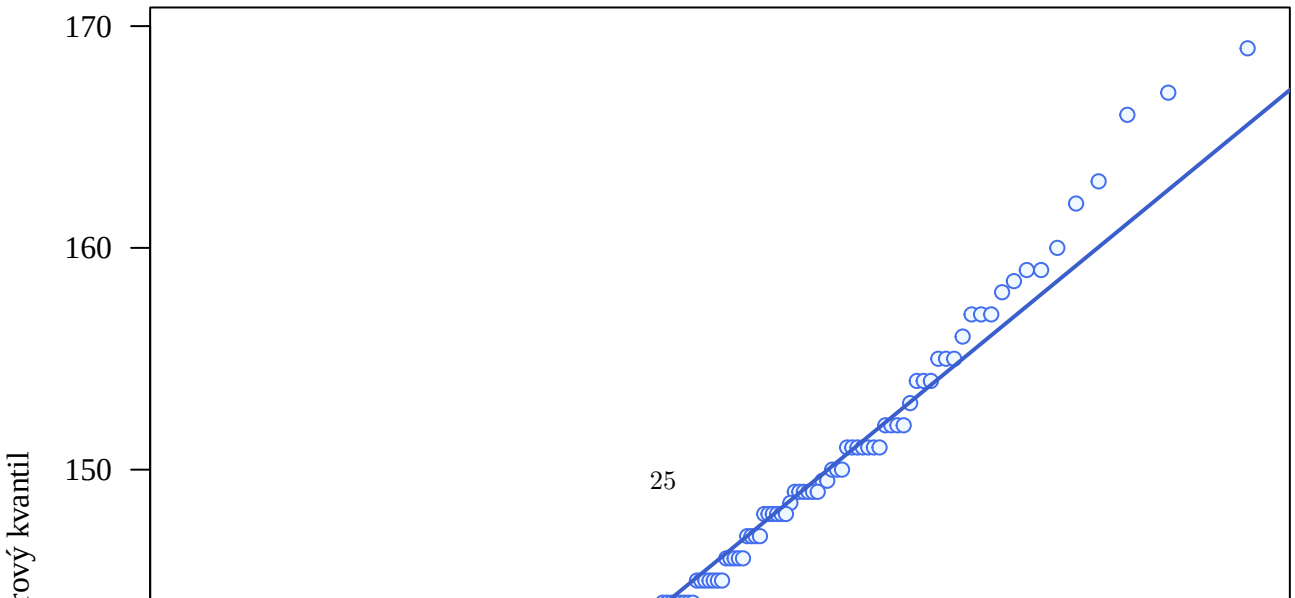
```
114 t.test(cla.LA, cla.LV, conf.level = 0.90, alternative = 'two.sided', var.equal = T)
```

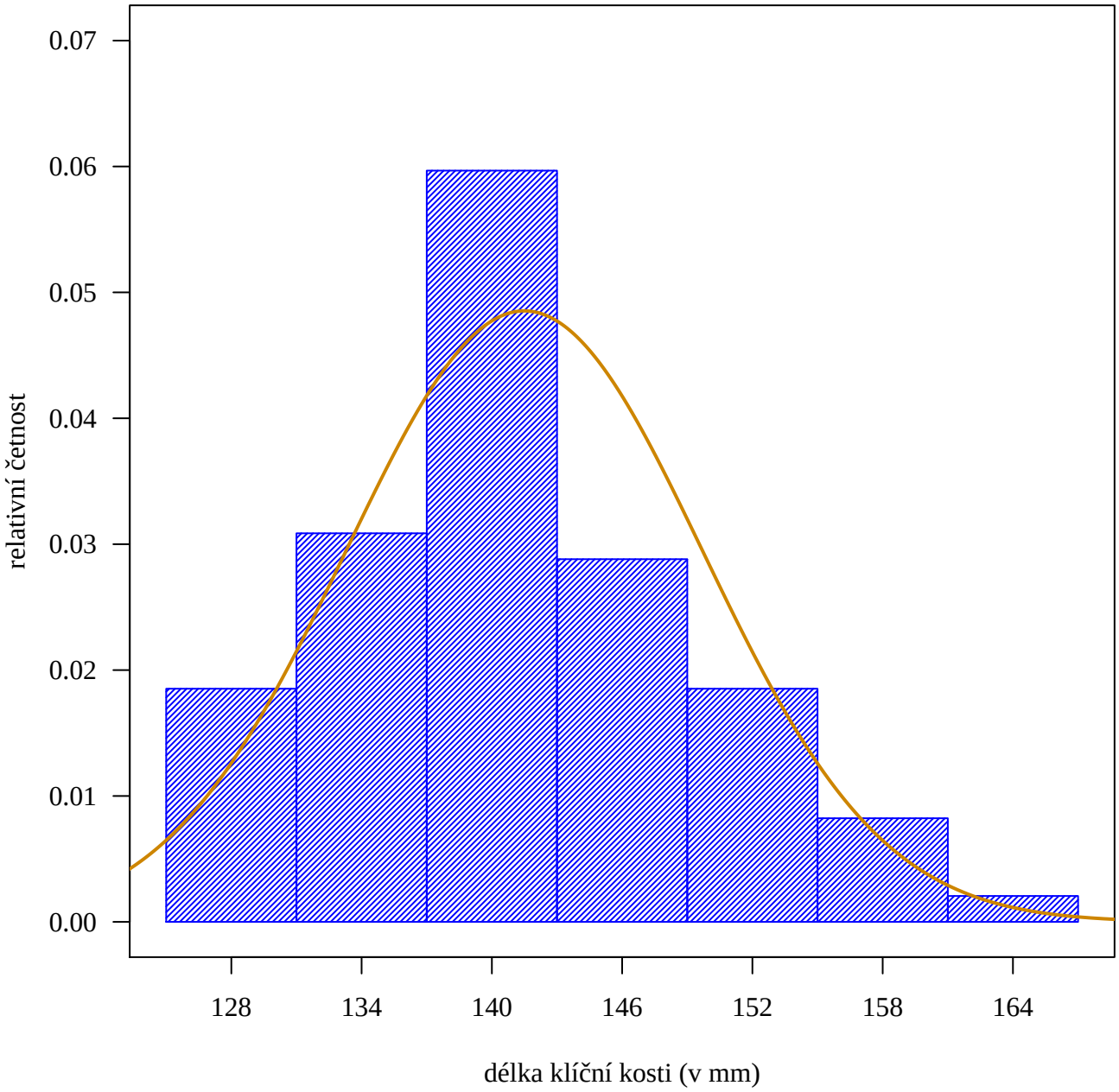
```
115
116        Two Sample t-test
117
118 data:  cla.LA and cla.LV
119 t = 3.32, df = 199, p-value = 0.001071
120 alternative hypothesis: true difference in means is not equal to 0
121 90 percent confidence interval:
122  2.045579 6.100100
123 sample estimates:
124 mean of x mean of y
125  145.5667  141.4938
```

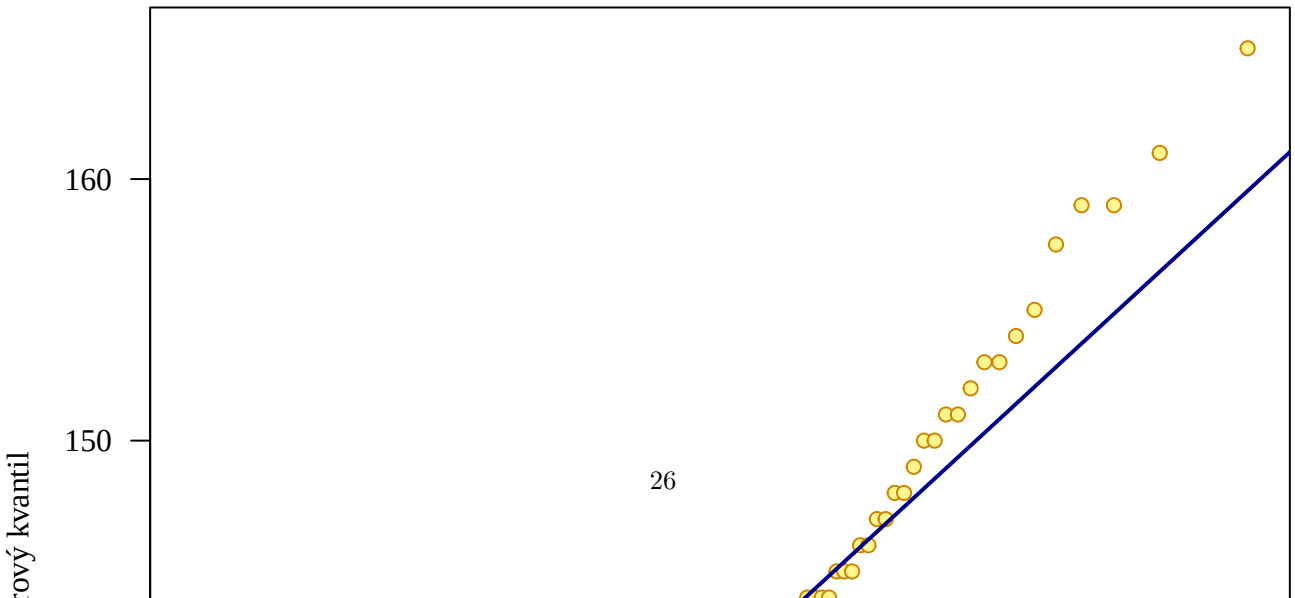
Součástí výstupu je hodnota testovací statistiky $\mathsf{t} = 3.32$, počet stupňů volnosti Studentova rozdělení $\mathsf{num\ df} = 199$, hranice 90% oboustranného intervalu spolehlivosti 2.045579 a 6.100100 a p-hodnota $\mathsf{p\text{-}value} = 0.001071$. Jediné, co musíme stanovit zvlášť, jsou dolní a horní hranice kritického oboru.

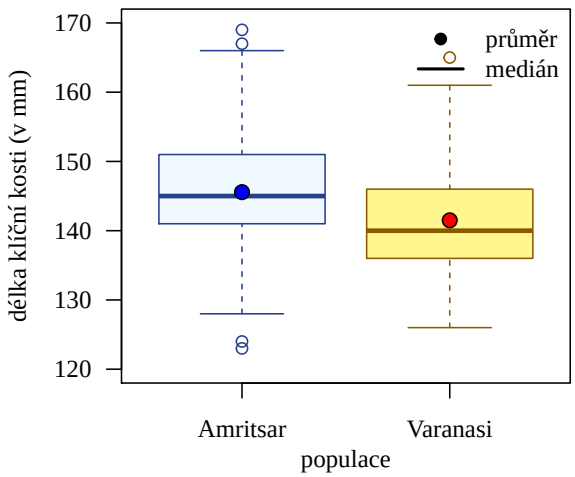
★


0.06
0.05
0.04
0.03
0.02
0.01
0.00
relativní četnost
125
131
137
143
149
155
161
167
délka klíční kosti (v mm)


170
160
150
25


0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
relativní četnost
128
134
140
146
152
158
164
délka klíční kosti (v mm)


160
150
rový kvantil



Obrázek 16: Krabicový diagram délky klíční kosti u mužů z indické populace z Amritsaru a u mužů z indické populace z Varanasi

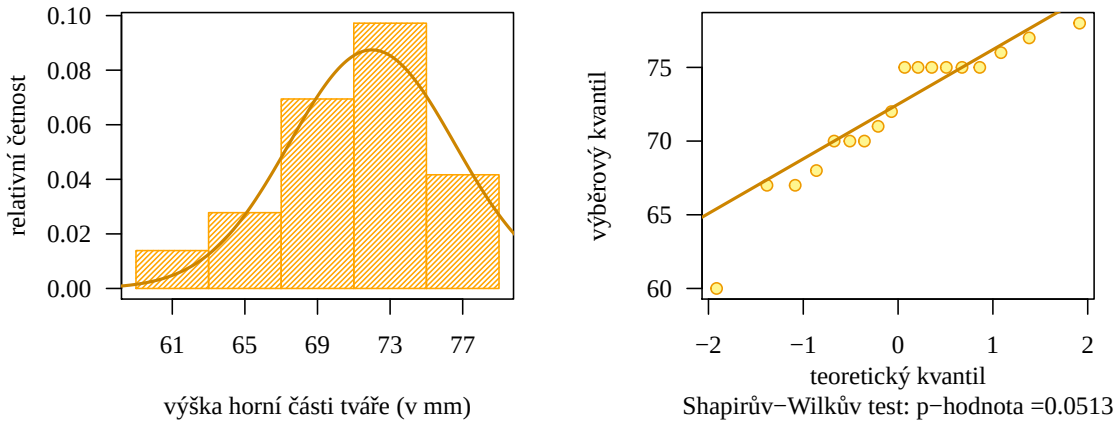
Příklad 10.7. Test o rozdílu středních hodnot (pravostranná alternativa)
Mějme datový soubor 15-anova-means-skull.txt a proměnnou upface.H popisující výšku horní části tváře v mm (viz sekce **??**). Předpokládejme, že náhodná veličina X, popisující *výšku horní části tváře čínské populace*, pochází z 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 $\alpha = 0.05$ prověřte, 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 odtraníme 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.

```
126 data <- read.delim('00-Data//15-anova-means-skull.txt')
127 data <- na.omit(data)
128 #head(data, n = 3)
129 upface.Hc <- data[data$pop == 'cin', 'upface.H']
130 upface.Hm <- data[data$pop == 'mal', 'upface.H']
131
132 n1 <- length(upface.Hc) # 18
133 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 Shapirův-Wilkův test spolu s QQ-diagramem a histogramem (viz obrázek 17). Datový soubor rozdělíme do pěti ekvidistatních intervalů s šířkou 4 mm prostřednictvím stanovených hranic $59, 63, \ldots, 79$.

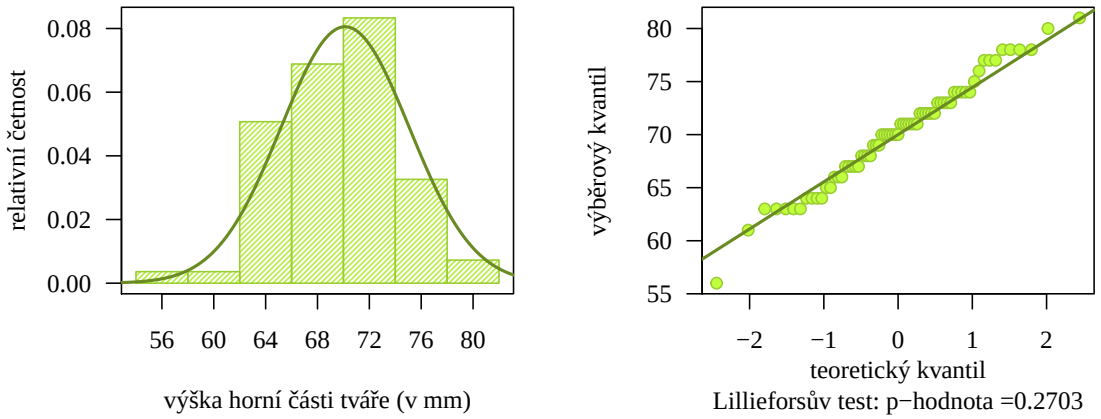


Obrázek 17: Histogram a QQ-diagram výšky horní části tváře mužů čínské populace

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 QQ-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

Shapirova-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 ekvidistatních intervalů s šířkou 4 mm prostřednictvím stanovených hranic $54, 58 \ldots, 82$ (viz obrázek 18).



Obrázek 18: Histogram a QQ-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 QQ-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).

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

```
[1] 0.7340571                                                    135
```

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
  $H_0 : \mu_1 \leq \mu_2 \quad \rightarrow \quad \mu_1 - \mu_2 \leq \mu_0$, kde $\mu_0 = 0$
  $H_1 : \mu_1 > \mu_2 \quad \rightarrow \quad \mu_1 - \mu_2 > \mu_0$, kde $\mu_0 = 0$
  (pravostranná alternativa)

2. **Volba hladiny významnosti**

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

3. **Testování kritickým oborem**

   - **Aritmetický průměr výběrových rozptylů**

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

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

   - **Testovací statistika**

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

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

   - **Kritický obor**

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

```
144 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**

$$\begin{aligned}(d, \infty) &= \left(m_1 - m_2 - s_* \sqrt{\frac{1}{n1} + \frac{1}{n_2}} t_{n1+n_2-2}(1-\alpha)\,;\,\infty\right) \\ &= \left((72 - 70.13043) - 4.877109 \sqrt{\frac{1}{18} + \frac{1}{69}} t_{85}(0.95)\,;\,\infty\right) \\ &= (1.869565 - 4.877109 \times 0.2646664 \times 1.662978\,;\,\infty) \\ &= (-0.277018\,;\,\infty) \doteq (-0.2770\,;\,\infty)\end{aligned}$$

```
145 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**

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

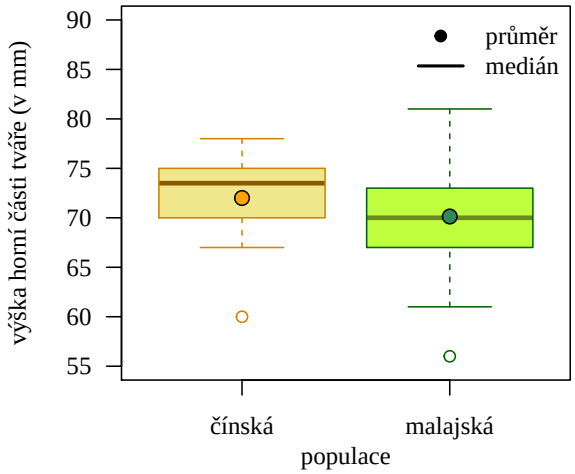
```
146 p.hodnota <- 1 - pt(tw, n1 + n2 -2) # 0.07559631
```

- **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**
   Vhodný 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

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

```
148
149
150
151
152
153
154
155
156
157
158
```

```
        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
```

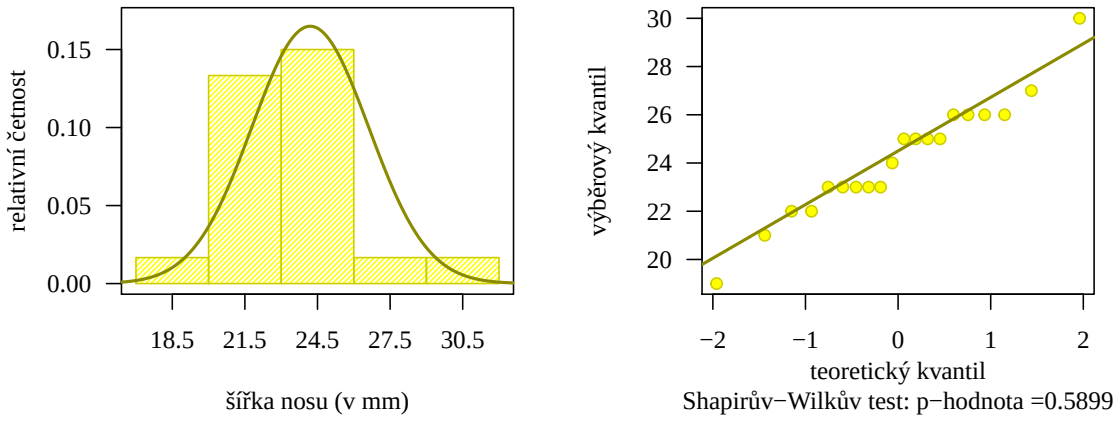
★

Příklad 10.8. Test o rozdílu středních hodnot (levostranná alternativa)
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.

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

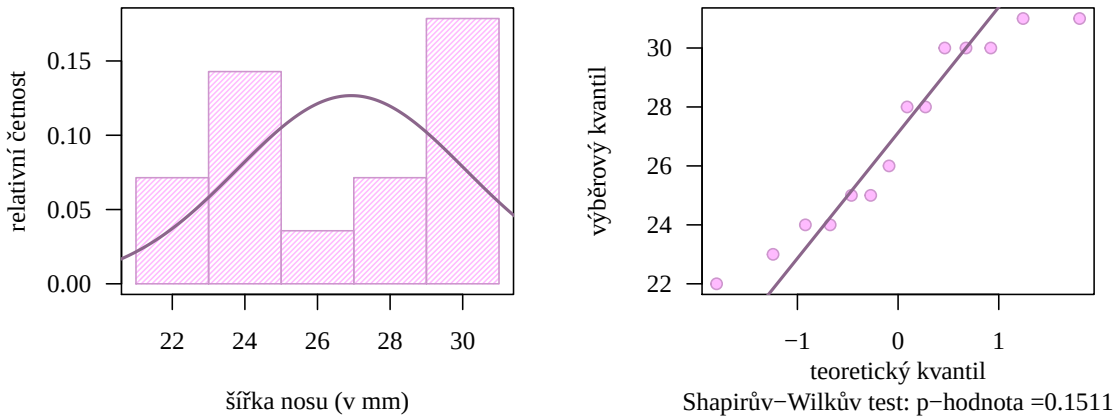
Datový soubor obsahuje naměřené hodnoty šířek nosu 20 jedinců německé populace a 14 jedinců bantuské populace.
Nejprve otestujeme předpoklad normality obou náhodných výběrů. Na hladině významnosti α testujeme hypotézu H_{01}: *Náhodný výběr šířky nosu německé populace pochází z normálního rozdělení.* oproti alternativní hypotéze H_{11}: *Náhodný výběr šířky nosu německé populace nepochází z normálního rozdělení.* K otestování předpokladu normality použijeme Shapirův-Wilkův test, QQ-diagram a histogram. Datový soubor rozdělíme do pěti ekvidistatních intervalů s šířkou 3 mm prostřednictvím stanovených hranic $17, 20, \ldots, 32$ (viz obrázek 20).



Obrázek 20: Histogram a QQ-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 QQ-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 Shapirovým-Wilkovým testem, QQ-diagramem a histogramem. Datový soubor rozdělíme do pěti ekvidistatních intervalů s šířkou 2 mm prostřednictvím stanovených hranic $21, 23 \ldots, 31$ (viz obrázek 21).



Obrázek 21: Histogram a QQ-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 normlání 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).

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

```
[1] 0.2877778                                                    168
```

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
     $H_0 : \mu_1 \geq \mu_2 \quad \rightarrow \quad \mu_1 - \mu_2 \geq \mu_0$, kde $\mu_0 = 0$
     $H_1 : \mu_1 < \mu_2 \quad \rightarrow \quad \mu_1 - \mu_2 < \mu_0$, 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**

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

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

- **Testovací statistika**

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

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

- **Kritický obor**

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

```
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**

$$\begin{aligned}(d,\infty) &= \left(-\infty\,;\, m_1 - m_2 - s_*\sqrt{\frac{1}{n1}+\frac{1}{n_2}}t_{n1+n_2-2}(1-\alpha)\,;\,\infty\right)\\ &= \left(-\infty\,;\,(24.2-26.92857) - 2.739346\sqrt{\frac{1}{20}+\frac{1}{14}}t_{32}(0.01)\right)\\ &= (-\infty\,;\,-2.72857 - 2.739346\times 0.348466\times(-2.448678))\\ &= (-\infty\,;\,-0.39113) \doteq (-\infty\,;\,-0.3911)\end{aligned}$$

```
178 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**

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

```
179 p.hodnota <- 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 $\alpha = 0.01$.

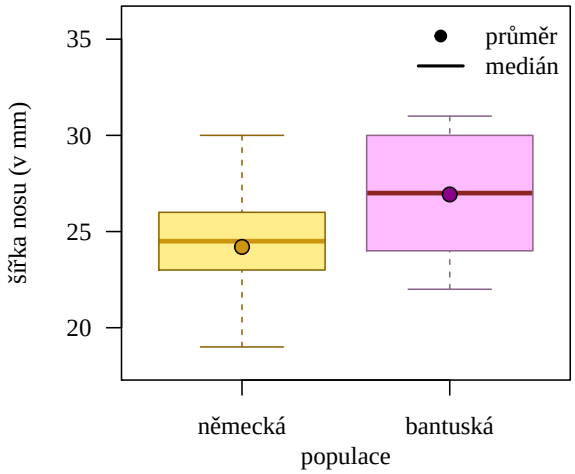
6. **Grafická vizualizace výsledků testování**
   Výsledek testování vizualizujeme pomocí krabicového diagramu (viz obrázek 22).

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.

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

```
181
182        Two Sample t-test
183
184 data:  nose.Bn and nose.Bb
185 t = -2.8584, df = 32, p-value = 0.003715
186 alternative hypothesis: true difference in means is less than 0
187 99 percent confidence interval:
188        -Inf -0.3911394
189 sample estimates:
190 mean of x mean of y
191  24.20000  26.92857
```



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

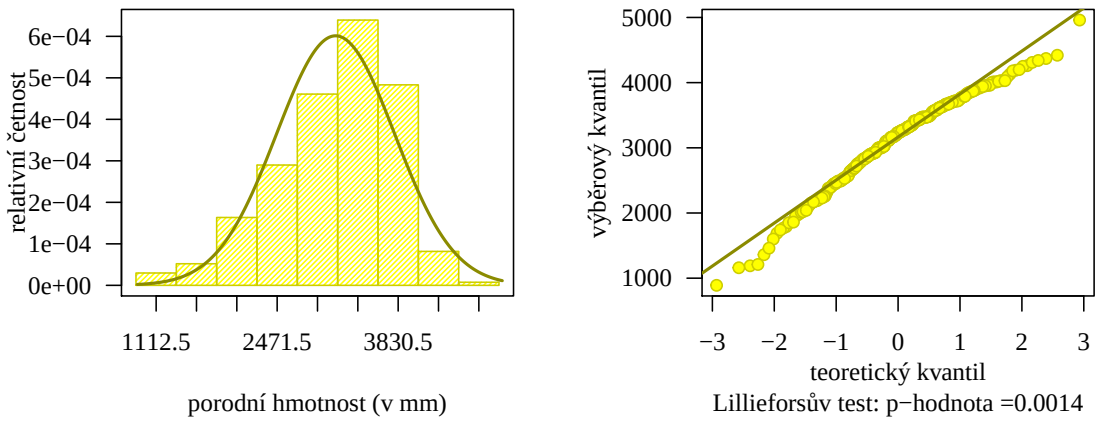
★

Příklad 10.9. Test o rozdílu středních hodnot (nesplnění předpokladů)
Mějme datový soubor 10-two-samples-means-birth.txt a proměnnou weight.W popisující porodní hmotnost novorozence v g (viz sekce **??**). Na hladině významnosti $\alpha = 0.05$ zjistěte, 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.

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

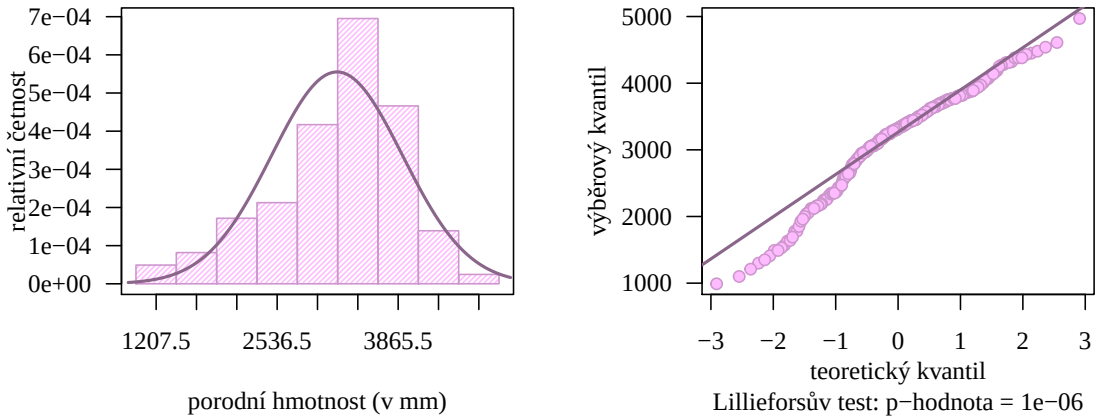
Datový soubor obsahuje naměřené porodní hmotnosti 297 novorozenců bez staršího sourozence a 276 novorozenců s jedním starším sourozencem.
Nejprve otestujeme předpoklad normality obou náhodných výběrů ($\alpha = 0.05$). Protože rozsah náhodného výběru porodních hmotností novorozenců bez staršího sourozence je větší než 30, použijeme na ověření předpokladu normality náhodného výběru Lillieforsův test v kombinaci s QQ-diagramem a histogramem (viz obrázek 23). Datový soubor rozdělíme do devíti ekvidistatních intervalů s šířkou 453 mm prostřednictvím stanovených hranic $886, 1339, \ldots, 4064$ g.



Obrázek 23: Histogram a QQ-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 QQ-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, QQ-diagramem a histogramem. Datový soubor rozdělíme do devíti ekvidistatních intervalů s šířkou 443 mm prostřednictvím stanovených hranic $986, 1429, \ldots, 4973$ g (viz obrázek 24).
Protože p-hodnota $= 1 \times 10^{-6}$ 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



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

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 $X_{11}, \ldots, X_{1n_1}$ je náhodný výběr z $N(\mu_1, \sigma_1^2)$, a $X_{21}, \ldots, X_{2n_2}$ 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.

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

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

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

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.

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

kde

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

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

$H_{11} : \mu \neq \mu_0$	$W = (-\infty\,;\, t_{df}(\alpha/2)\rangle \cup \langle t_{df}(1-\alpha/2)\,;\, \infty)$
$H_{12} : \mu > \mu_0$	$W = \langle t_{df}(1-\alpha)\,;\, \infty)$
$H_{13} : \mu < \mu_0$	$W = (-\infty\,;\, t_{df}(\alpha)\rangle$

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 a implementované funkce `qt()`.

Interval spolehlivosti má podle zvolené alternativní hypotézy jeden z následujících tvarů

$$H_{11} : \mu_1 - \mu_2 \neq \mu_0 \qquad (d, h) = \left(m_1 - m_2 - \sqrt{\tfrac{s_1^2}{n_1} + \tfrac{s_1^2}{n_2}}\, t_{df}(1-\alpha/2)\,;\, m_1 - m_2 - \sqrt{\tfrac{1}{n_1} + \tfrac{1}{n_2}}\, t_{df}(\alpha/2)\right)$$

$$H_{12} : \mu_1 - \mu_2 > \mu_0 \qquad (d, \infty) = \left(m_1 - m_2 - \sqrt{\tfrac{s_1^2}{n_2} + \tfrac{s_2^2}{n_2}}\, t_{df}(1-\alpha)\,;\, \infty\right)$$

$$H_{13} : \mu_1 - \mu_2 < \mu_0 \qquad (-\infty, h) = \left(-\infty\,;\, m_1 - m_2 - \sqrt{\tfrac{s_1^2}{n_1} + \tfrac{s_2^2}{n_2}}\, t_{df}(\alpha)\right)$$

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

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

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 a implementované funkce `pt()`.

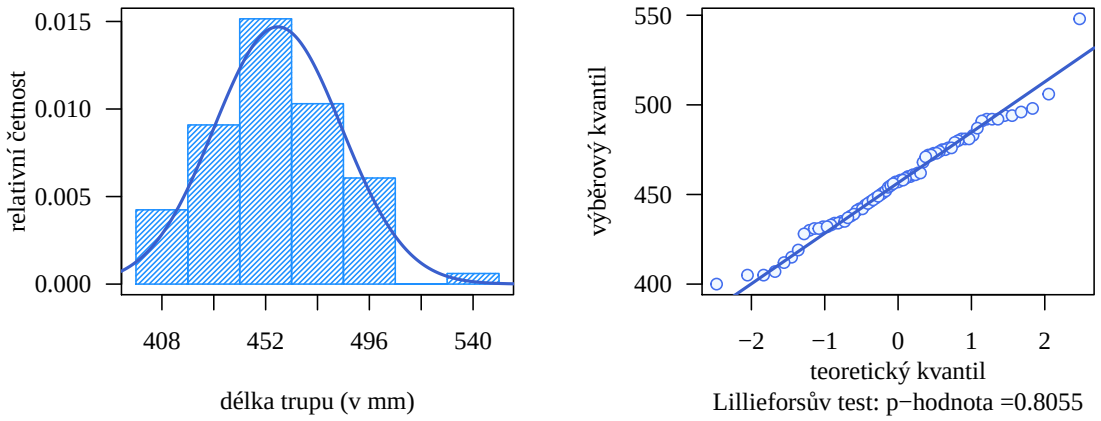
Příklad 10.10. Test o rozdílu středních hodnot (oboustranná alternativa)
Mějme datový soubor 13-two-samples-correlations-trunk.txt a proměnnou tru.L popisující délku trupu v mm (viz sekce **??**). Předpokládejme, že náhodná veličina X, popisující *délku trupu mužů*, pochází z normálního rozdělení $N(\mu_1, \sigma_1^2)$, 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 10.10
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.

```
200 data <- read.delim('00-Data//13-two-samples-correlations-trunk.txt')
201 data <- na.omit(data)
202 #head(data, n = 3)
203 tru.Lm <- data[data$sex == 'm', 'tru.L']
204 tru.Lf <- data[data$sex == 'f', 'tru.L']
205
206 n1 <- length(tru.Lm) # 75
207 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 $\alpha = 0.05$.

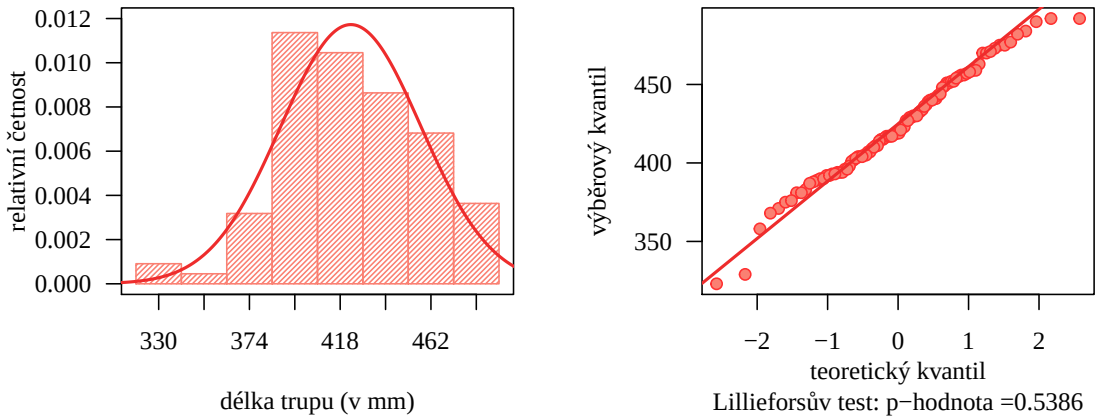
Nejprve testujeme hypotézu H_{01}: *Náhodný výběr délky trupu mužů pochází z normálního rozdělení.* oproti alternativní hypotéze H_{11}: *Náhodný výběr délky trupu mužů nepochází z normálního rozdělení.* Normalitu otestujeme Lillieforsovým testem, QQ-diagramem a histogramem (viz obrázek 25). Naměřené hodnoty rozdělíme do sedmi ekvidistatních intervalů s šířkou 22m mm prostřednictvím stanovených hranic $397, 419, \ldots, 154$.



Obrázek 25: Histogram a QQ-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 QQ-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 $228, 250 \ldots, 495$ (viz obrázek 26).



Obrázek 26: Histogram a QQ-diagram délky trupu žen

Protože p-hodnota $= 0.5386$ je větší než 0.05, nulovou hypotézu nezamítáme na hladině významnosti $\alpha = 0.05$. Náhodný výběrdélky trupu žen pochází z normálního rozdělení. Vykreslený histogram i QQ-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).

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

```
[1] 0.04256397                                                    209
```

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
     $H_0 : \mu_1 = \mu_2 \quad \rightarrow \quad \mu_1 - \mu_2 = \mu_0$, kde $\mu_0 = 0$
     $H_1 : \mu_1 \neq \mu_2 \quad \rightarrow \quad \mu_1 - \mu_2 \neq \mu_0$, kde $\mu_0 = 0$
     (oboustranná alternativa)

2. **Volba hladiny významnosti**

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

3. **Testování kritickým oborem**

- **Testovací statistika**

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

```
alpha <- 0.01
mu0 <- 0
m1 <- mean(tru.Lm)
m2 <- mean(tru.Lf)
s1 <- sd(tru.Lm)
s2 <- sd(tru.Lf)

tw <- ((m1 - m2) - mu0) / (sqrt(s1 ^ 2 / n1 + s2 ^ 2 / n2)) # 7.351989
```

- **Welchova aproximace počtu stupňů volnosti**

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

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

Kritický obor

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

```
qt(alpha/2, df) # -2.604664
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**

$$
\begin{aligned}
(d, h) &= \left(m_1 - m_2 - \sqrt{\frac{s_1^2}{n1} + \frac{s_2^2}{n_2}} t_{df}(1 - \alpha/2) \,;\, m_1 - m_2 - \sqrt{\frac{s_1^2}{n1} + \frac{s2^2}{n_2}} t_{df}(\alpha/2) \right) \\
&= ((457.1867 - 423.17) - \sqrt{\frac{27.15555^2}{75} + \frac{34.02288^2}{100}} t_{172.2998}(0.995) \,;\, (457.1867 - 423.17) - \sqrt{\frac{27.15555^2}{75} + \frac{34.02}{10} \\
&= (34.0167 - 4.626865 \times 2.604664 \,;\, 34.0167 - 4.626865 \times (-2.604664)) \\
&= (21.96527 \,;\, 46.06813) \doteq (21.9653 \,;\, 46.06813)
\end{aligned}
$$

```
221 dh <- (m1 - m2) - sqrt(s1^2 / n1 + s2^2 / n2) * qt(1 - alpha / 2, n1 + n2 - 2) #
    21.96578
222 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**

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

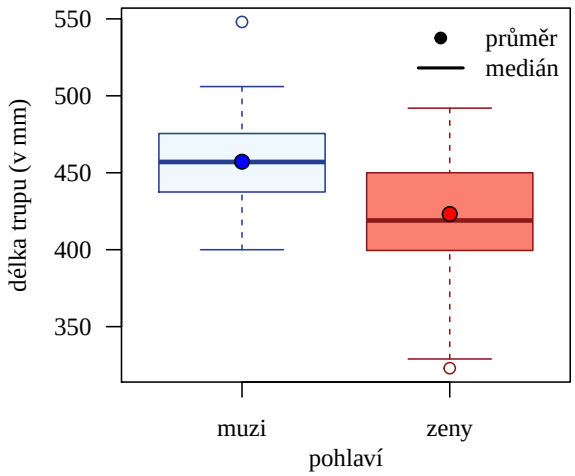
```
223 p.hodnota <- 2*min(pt(tw, df), 1 - pt(tw, df)) # 7.549072e-12
```

   - **Závěr testování**
     Protože p-hodnota $= 1.5490 \times 10^{-12}$ 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 27).

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 27: Krabicový diagram délky trupu mužů a žen

```
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 num df = 172.3, hranice intervalu spolehlivosti 24.88403 a 43.14930 a p-hodnota p-value = 7.549e-12. Jediné, co musíme stanovit zvlášť, jsou dolní a horní hranice kritického oboru. ★

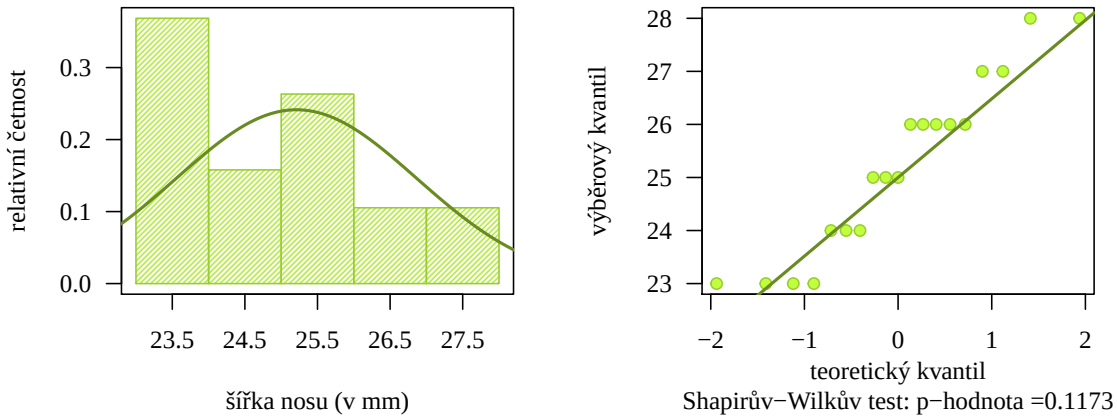
Příklad 10.11. Test o rozdílu středních hodnot (pravostranná alternativa)
Mějme datový soubor 19-more-samples-correlations-skull.txt a proměnnou nose.B popisující šířku nosu v mm (viz sekce **??**). Předpokládejme, že náhodná veličina X, popisující *šířku nosu bantuské populace*, pochází z normálního 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 10.11
Příkazem read.delim() načteme datový soubor a příkazem na.omit() z něj odtraníme NA hodnoty. Pomocí operátoru [] vybereme z datové tabulky nameřené hodnoty šířky nosu (nose.B) čínské populace (pop == 'cin'), resp. bantuské populace (pop == 'ban') a zjistíme rozsahy obou náhodných výběrů.

```
236 data <- read.delim('00-Data//19-more-samples-correlations-skull.txt')
237 data <- na.omit(data)
238 #head(data, n = 3)
239 nose.Bc <- data[data$pop == 'cin', 'nose.B']
240 nose.Bb <- data[data$pop == 'ban', 'nose.B']
241
242 n1 <- length(nose.Bc) # 19
243 n2 <- length(nose.Bb) # 14
```

Datový soubor obsahuje naměřené hodnoty šířky nosu u 19 jedinců čínské populace a 14 jedinců bantusképopulace. 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 Shapirův-Wilkův test, QQ-diagram a histogram (viz obrázek 28). Datový soubor rozdělíme do pěti ekvidistatních intervalů s šířkou 1 mm prostřednictvím stanovených hranic $23, 24 \ldots, 28$.

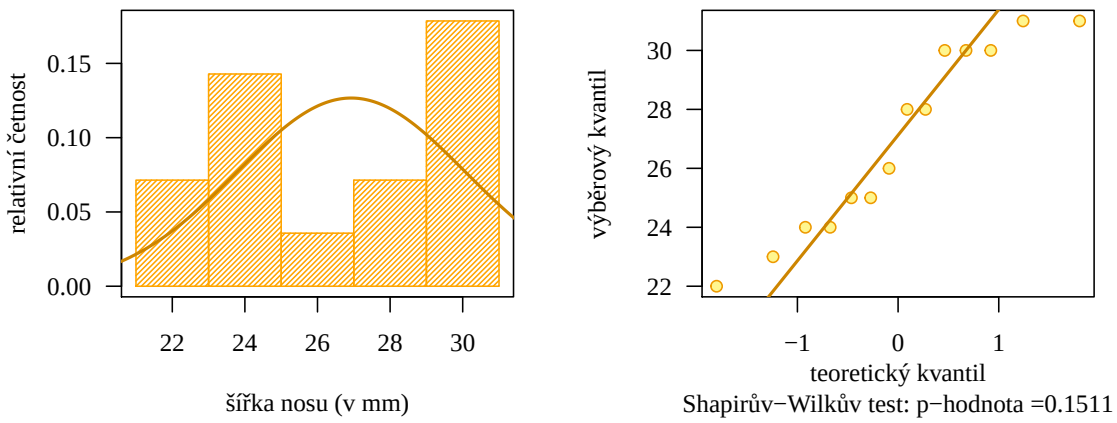


Obrázek 28: Histogram a QQ-diagram šířky nosu mužů čínské populace

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 QQ-diagram 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 Shapirův-Wilkův test, QQ-diagram a histogram. Datový soubor rozdělíme do pěti ekvidistatních intervalů s šířkou 2 mm prostřednictvím stanovených hranic $21, 23, \ldots, 31$ (viz obrázek 29).



Obrázek 29: Histogram a QQ-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. QQ-diagram normalitu naměřených hodnot podporuje. Proto se kloníme k závěru Shapirova-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).

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

```
[1] 0.01257845                                                    245
```

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
     $H_0 : \mu_1 \leq \mu_2 \quad \rightarrow \quad \mu_1 - \mu_2 \leq \mu_0$, kde $\mu_0 = 0$
     $H_1 : \mu_1 > \mu_2 \quad \rightarrow \quad \mu_1 - \mu_2 > \mu_0$, kde $\mu_0 = 0$
     (pravostranná alternativa)

2. **Volba hladiny významnosti**

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

3. **Testování kritickým oborem**

   - **Testovací statistika**

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

```
246 alpha <- 0.05
247 mu0 <- 0
248 m1 <- mean(nose.Bc)
249 m2 <- mean(nose.Bb)
250 s1 <- sd(nose.Bc)
251 s2 <- sd(nose.Bb)
252
253 tw <- ((m1 - m2) - mu0) / (sqrt(s1 ^ 2 / n1 + s2 ^ 2 / n2)) # -1.861145
```

   - **Welchova aproximace počtu stupňů volnosti**

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

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

   - **Kritický obor**

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

```
255 qt(1 - alpha, df) # 1.732689
```

- **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**

$$\begin{aligned}(d, \infty) &= \left(m_1 - m_2 - \sqrt{\frac{s_1^2}{n1} + \frac{s_2^2}{n_2}} t_{df}(1-\alpha)\,;\, \infty \right) \\ &= \left((25.21053 - 26.92857) - \sqrt{\frac{1.652572^2}{19} + \frac{3.149219^2}{14}} t_{18.26753}(0.95)\,;\, \infty \right) \\ &= (-1.71804 - \times 0.9231117 \times 1.732689\,;\, \infty) \\ &= (-3.31751\,;\, \infty) \doteq (-3.3175\,;\, \infty)\end{aligned}$$

```
256 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**

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

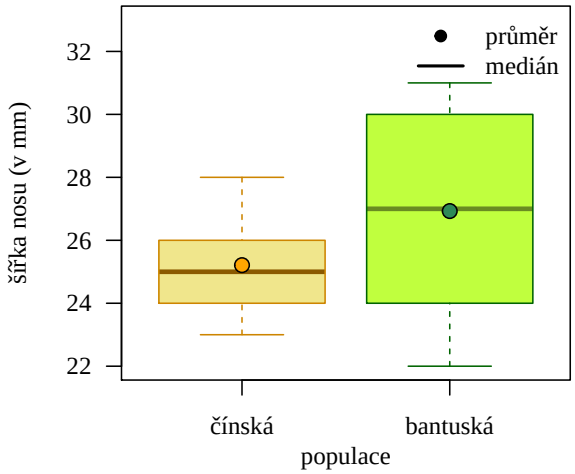
```
257 p.hodnota <- 1 - pt(tw, df) # 0.960552
```

- **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 30).

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 30: Krabicový diagram šířky nosu mužů čínské a bantuské populace

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

```
259
260        Welch Two Sample t-test
261
262 data:  nose.Bb and nose.Bc
263 t = 1.8611, df = 18.268, p-value = 0.03945
264 alternative hypothesis: true difference in means is greater than 0
265 95 percent confidence interval:
266  0.1185797       Inf
267 sample estimates:
268 mean of x mean of y
269  26.92857  25.21053
```

★

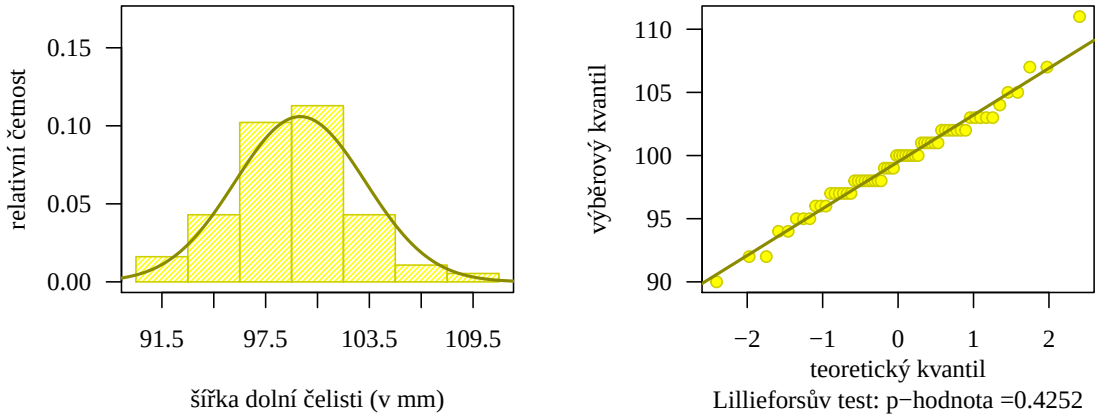
Příklad 10.12. Test o rozdílu středních hodnot (levostranná alternativa)
Mějme datový soubor 16-anova-head.txt a proměnnou bigo.W popisující šířku dolní čelisti v mm (viz sekce **??**). Předpokládejme, že náhodná veličina X, popisující *šířku dolní čelisti u žen orientovaných heterosexuálně*, pochází 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 $\alpha = 0.05$ zjistěte, zda šířka dolní čelisti žen orientovaných heterosexuálně je statisticky významně menší než šířka dolní čelisti žen orientovaných jinak než výhradně heterosexuálně.

Řešení příkladu 10.12
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') .

```
data <- read.delim('00-Data//16-anova-head.txt')
data <- na.omit(data)
#head(data, n = 3)
bigo.Wfo <- data[data$sex == 'f' & data$sexor == 'op', 'bigo.W']
bigo.Wfs <- data[data$sex == 'f' & data$sexor == 'sa', 'bigo.W']

n1 <- length(bigo.Wfo) # 62
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, QQ-diagram a histogram (viz obrázek 31). Datový soubor rozdělíme na základě Sturgesova pravidla do sedmi třídicích intervalů s ekvidistantní šířkou 3 mm prostřednictvím stanovených hranic $90, 93, \ldots, 111$.

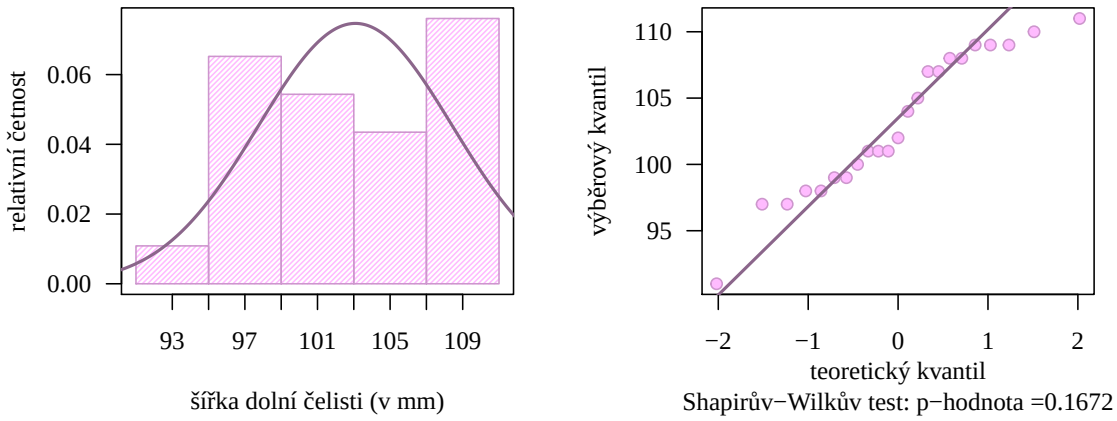


Obrázek 31: Histogram a QQ-diagram šířky dolní čelisti heterosexuálně orientovaných žen

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 QQ-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 Shapirovým-Wilkovým

testem, QQ-diagramem a histogramem. Datový soubor rozdělíme do pěti ekvidistatních intervalů s šířkou 4 mm prostřednictvím stanovených hranic $91, 95 \ldots, 111$ (viz obrázek 32).



Obrázek 32: Histogram a QQ-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 nulovu 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).

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

```
[1] 0.0336774                                                    279
```

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
     $H_0 : \mu_1 \geq \mu_2 \quad \rightarrow \quad \mu_1 - \mu_2 \geq \mu_0$, kde $\mu_0 = 0$
     $H_1 : \mu_1 < \mu_2 \quad \rightarrow \quad \mu_1 - \mu_2 < \mu_0$, kde $\mu_0 = 0$
     (levostranná alternativa)

2. **Volba hladiny významnosti**

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

3. **Testování kritickým oborem**

   - **Testovací statistika**

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

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

   - **Welchova aproximace počtu stupňů volnosti**

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

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

   - **Kritický obor**

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

```
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**

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

```
287 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**

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

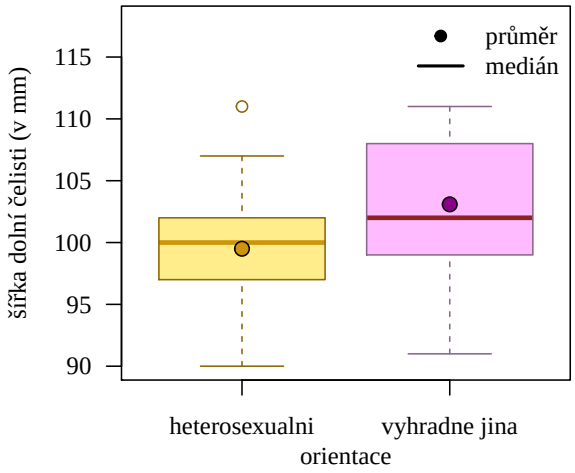
```
288 p.hodnota <- pt(tw, df) # 0.002961084
```

   - **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 33).

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. bigo.Wfo a bigo.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 33: Krabicový diagram šířky dolní čelisti žen orientovaných heterosexuálně a žen orientovaných jinak než heterosexuálně

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

```
290
291         Welch Two Sample t-test
292
293 data:  bigo.Wfo and bigo.Wfs
294 t = -2.9589, df = 30.49, p-value = 0.002961
295 alternative hypothesis: true difference in means is less than 0
296 95 percent confidence interval:
297       -Inf -1.530462
298 sample estimates:
299 mean of x mean of y
300    99.500   103.087
```

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

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.

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

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

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

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

$H_{11} : \rho_1 - \rho_2 \neq \rho_0$	$W = \left(-\infty\,;\, u_{\alpha/2}\right\rangle \cup \left\langle u_{1-\alpha/2}\,;\, \infty\right)$
$H_{12} : \rho_1 - \rho_2 > \rho_0$	$W = \left\langle u_{1-\alpha}\,;\, \infty\right)$
$H_{13} : \rho_1 - \rho_2 < \rho_0$	$W = \left(-\infty\,;\, u_{\alpha}\right\rangle$

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

Interval spolehlivosti má podle zvolené alternativní hypotézy jeden z následujících tvarů

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

kde $s_g = \sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}$ a tanh je hyperbolický tangens, jehož hodnotu získáme pomocí 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

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

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 R a implementované funkce pnorm().

Příklad 10.13. Test o rozdílu dvou korelačních koeficientů $\rho_1 - \rho_2$ (oboustranná alternativa)
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 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.

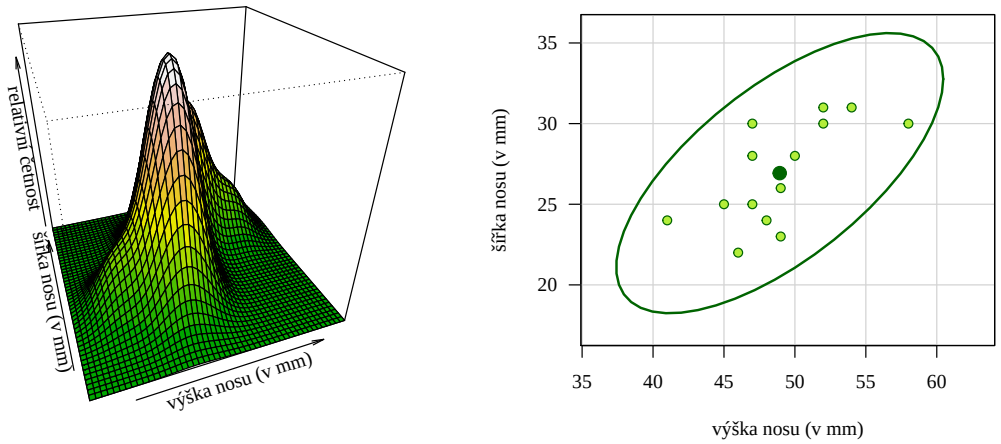
Abychom mohli hypotézu ze zadání otestovat pomocí parametrického testu, musí oba datové soubory splňovat předpoklad dvourozměrné normality. Před samotným testováním hypotézy uvedené v zadání je potřeba ověřit, zda oba výběry tento předpoklad splňují. Závěr o dvourozměrné normalitě obou náhodných výběrů 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 # sikmost: 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 významnsoti $\alpha = 0.05$ hypotézu o dvourozměrné normalitě náhodného výběru výšek a šířek nosu u mužů bantuské populace. Ke stejnému závěru bychom došli také použitím Henzova-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.

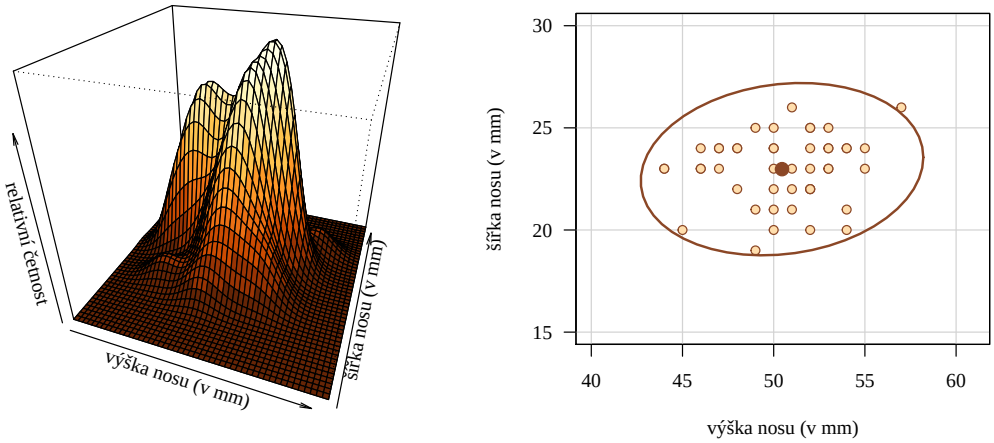
```
316 MVN::mvn(cbind(nose.HP, nose.BP), mvnTest = 'mardia')$multivariateNormality # sikmost:
    0.409373 # spicatost: 0.2925019
```

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 významnsoti $\alpha = 0.05$. Stejný závěr byhcom získali také použitím Henzova-Zirklerova testu (p-hodnota $= 0.1306 > 0.05$). N azákladě 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 odlehlým pozorováním, které se nevešlo do 95 % elipsy spolehlivosti. Vzhledem k tomu, že rozsah náhodného výběru je 46, požadujeme, aby elipsa spolehlivosti pokrývala alepsoň 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 Mardiova a Henzova-Zirklerova testu o dvourozměrné normalitě náhodného výběru výšek a šířek nosu u mužů peruánské populace.

Protože oba výběry pochází z dvourozměnrého 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
     $H_0 : \rho_1 = \rho_2 \quad \rightarrow \quad \rho_1 - \rho_2 = \rho_0$, kde $\rho_0 = 0$
     $H_1 : \rho_1 \neq \rho_2 \quad \rightarrow \quad \rho_1 - \rho_2 \neq \rho_0$, 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 Fisheroy Z-transformace těchto korelačních koeficientů i Fisherovu Z-tranformaci konstanty $\rho_0 = 0$ z nulové hypotézy.

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

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

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

```r
alpha <- 0.05
ksi0 <- 0

r1 <- cor(nose.HB, nose.BB) # 0.6719065
r2 <- cor(nose.HP, nose.BP) # 0.1370691

z1 <- 1 / 2 * log((1 + r1) / (1 - r1)) # 0.8142106
z2 <- 1 / 2 * log((1 + r2) / (1 - r2)) # 0.1379373
```

- **Testovací statistika**

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

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

- **Kritický obor**

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

```r
qnorm(alpha / 2) # -1.959964
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.

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

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

```
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 položce coeffcieints, 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()).

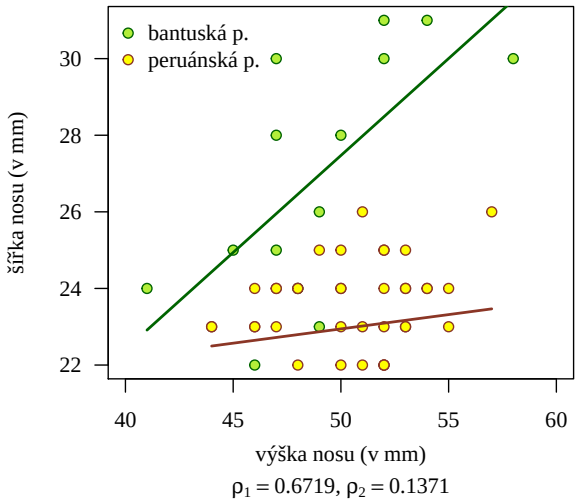
```
par(mar = c(5, 4, 1, 1), family = 'Times')
plot(nose.HB, nose.BB, xlab = '', ylab = 'šířka nosu (v mm)',
     xlim = c(40, 60), las = 1,
     pch = 21, col = 'darkgreen', bg = 'olivedrab2')
k <- lm(nose.BB ~ nose.HB)$coef
x <- seq(min(nose.HB), max(nose.HB), length = 1000)
y <- k[1] + x * k[2]
lines(x, y , col = 'darkgreen', lwd = 2)

points(nose.HP, nose.BP, pch = 21, col = 'tomato4', bg = 'yellow')
k <- lm(nose.BP ~ nose.HP)$coef
x <- seq(min(nose.HP), max(nose.HP), length = 1000)
y <- k[1] + x * k[2]
lines(x, y, col = 'tomato4', lwd = 2)

r1 <- round(r1, digit = 4)
r2 <- round(r2, digit = 4)
mtext('výška nosu (v mm)', side = 1, line = 2.3)
mtext(bquote(paste(rho[1] == . (r1), ', ', rho[2] == .(r2))), side = 1, line = 3.7)
legend('topleft', pch = 21, pt.bg = c('olivedrab2', 'yellow'),
       col = c('darkgreen', 'tomato4'), legend = c('bantuská p.', 'peruánská p.'),
       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 peránské populace existuje statisticky významný rozdíl. Mezi výškou a šíkou 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 $\rho_1 - \rho_2$ (pravostranná alternativa)
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 $\alpha = 0.10$ zjistěte, zda korelační koeficient největší výšky mozkovny a morfologické výšky tváře mužů je větší než korelační koeficient největší výšky mozkovny a morfologické výšky 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 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,

```
354 data <- read.delim('00-Data//05-one-sample-correlation-skull-mf.txt')
355 data <- na.omit(data)
356 head(data)
357 skull.pHM <- data[data$sex == 'm', 'skull.pH']
358 face.HM  <- data[data$sex == 'm', 'face.H']
359 skull.pHF <- data[data$sex == 'f', 'skull.pH']
360 face.HF  <- data[data$sex == 'f', 'face.H']
361
362 n1 <- length(skull.pHM) # 164
363 n2 <- length(skull.pHF) # 78
364 range(skull.pHM) # 127-149
365 range(face.HM)   # 100-136
366 range(skull.pHF) # 120-142
367 range(face.HF)   # 95-119
```

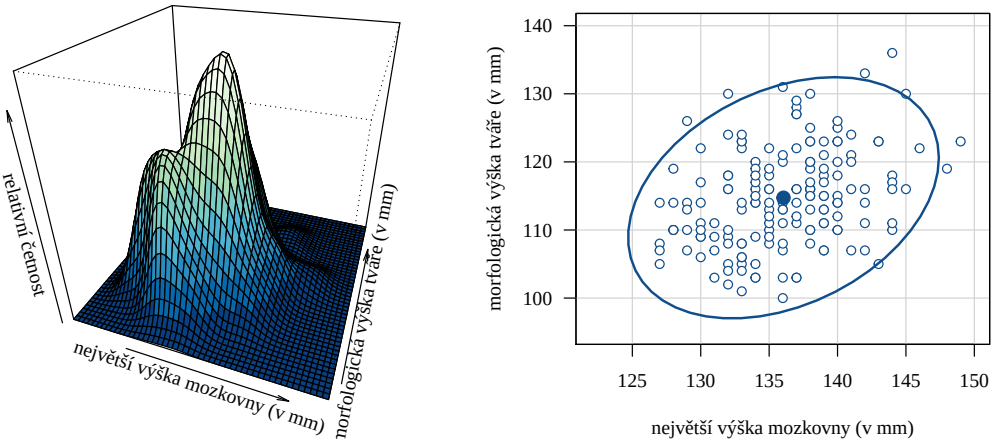
přičemž naměřené výšky mozkovny nabývají hodnot v rozmezí 127–149 mm, naměřené morfologické výšky tváře nabývají hodnot v rozmezí 100–136 mm. Soubor dále obsahuje údaje o největší výšce mozkovny a morfologické výšce tváře u 78 žen, přičemž naměřené výšky mozkovny nabývají hodnot v rozmezí 120–142 mm 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 Henzova-Zirklerova testu ($\alpha = 0.05$) v kombinaci s grafickou vizualizací dat.

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

Protože p-hodnota Henzova-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 37). 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.

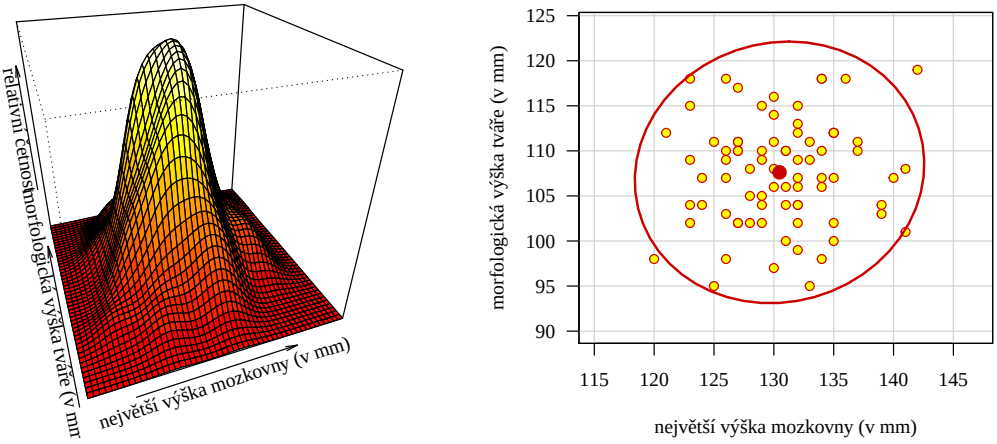
Jelikož p-hodnota Henzova-Zirklerova testu, tj. 0.9997 je větší než 0.05, hypotézu o dvourozměrné normalitě náhodného výběru největší výšky mozkovny a morfologické výšky tváře žen nezamítáme na hladině významnosti



Obrázek 37: 3D graf a tečkový diagram s 95% elipsou spolehlivosti pro největší výšku mozkovny a morfologickou výšku tváře mužů (v mm)

```
369 MVN::mvn(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 38: 3D graf a tečkový diagram s 95% elipsou spolehlivosti pro největší výšku mozkovny a morfologickou výšku tváře žen (v mm)

3D graf nabízí pohled na pospolité dvourozměrné normální rozdělení tvořené jedním souvislým kopcem (viz obrázek 38). Aby data pocházela z dvourozměnrého 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 morfologické 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 morfologické výšky tváře mužů je statisticky významně větší než korelační koeficient největší výšky mozkovny a morfologické 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
     $H_0 : \rho_1 \leq \rho_2 \quad \rightarrow \quad \rho_1 - \rho_2 \leq \rho_0$, kde $\rho_0 = 0$
     $H_1 : \rho_1 > \rho_2 \quad \rightarrow \quad \rho_1 - \rho_2 > \rho_0$, 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 morkovny 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.

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

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

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

```
370 alpha <- 0.10
371 ksi0  <- 0
372 r1 <- cor(skull.pHM, face.HM) # 0.3306431
373 r2 <- cor(skull.pHF, face.HF) # 0.06417166
374 z1 <- 1 / 2 * log((1 + r1) / (1 - r1)) # 0.3435501
375 z2 <- 1 / 2 * log((1 + r2) / (1 - r2)) # 0.06425996
```

- **Testovací statistika**

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

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

- **Kritický obor**

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

```
377 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.

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

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

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

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

5. **Testování p-hodnotou**

   - **p-hodnota**

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

```
1 - pnorm(zw) # 0.02287137
```

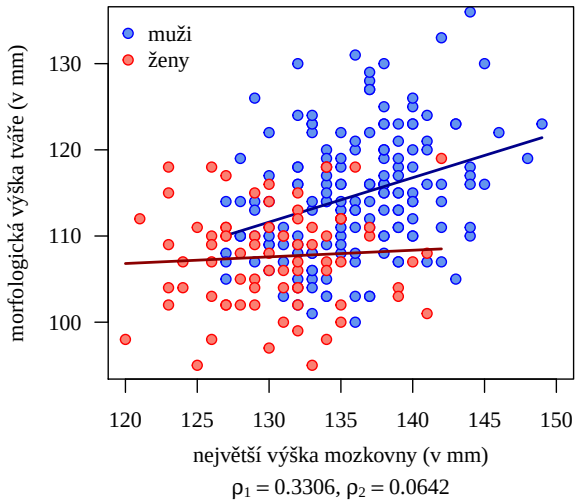
   - **Závěr testování**
     Protože p-hodnota $= 0.02287$ je menší než $\alpha = 0.10$, H_0 zamítáme na hladině významnosti $\alpha = 0.10$.

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 42). Koeficienty 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.

```
par(mar = c(5, 4, 1, 1), family = 'Times')
plot(skull.pHM, face.HM, xlab = '', ylab = 'morfologická výška tváře (v mm)',
     xlim = c(120, 150), ylim = c(95, 135), las = 1,
     pch = 21, col = 'blue', bg = 'cornflowerblue')
k <- lm(face.HM ~ skull.pHM)$coef
x <- seq(min(skull.pHM), max(skull.pHM), length = 1000)
y <- k[1] + x * k[2]
lines(x, y, col = 'darkblue', lwd = 2)

points(skull.pHF, face.HF, pch = 21, col = 'red', bg = 'salmon')
k <- lm(face.HF ~ skull.pHF)$coef
x <- seq(min(skull.pHF), max(skull.pHF), length = 1000)
y <- k[1] + x * k[2]
lines(x, y, col = 'darkred', lwd = 2)

r1 <- round(r1, digit = 4)
r2 <- round(r2, digit = 4)
mtext('největší výška mozkovny (v mm)', side = 1, line = 2.3)
mtext(bquote(paste(rho[1] == . (r1), ', ', rho[2] == .(r2))), side = 1, line = 3.7)
legend('topleft', pch = 21, pt.bg = c('cornflowerblue', 'salmon'),
       col = c('blue', 'red'), legend = c('muži', 'ženy'),
       bty = 'n')
```



Obrázek 39: 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 10.15. Test o rozdílu dvou korelačních koeficientů $\rho_1 - \rho_2$ (levostranná alternativa)
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 $\alpha = 0.01$ zjistěte, 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 10.15
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

```
403 data <- read.delim('00-Data//13-two-samples-correlations-trunk.txt')
404 data <- na.omit(data)
405 lowex.LM <- data[data$sex == 'm', 'lowex.L']
406 tru.LM   <- data[data$sex == 'm', 'tru.L']
407 lowex.LF <- data[data$sex == 'f', 'lowex.L']
408 tru.LF   <- data[data$sex == 'f', 'tru.L']
409
410 n1 <- length(lowex.LM) # 75
411 n2 <- length(lowex.LF) # 100
412 range(lowex.LM) # 915-1114
413 range(tru.LM)   # 400-548
414 range(lowex.LF) # 836-1076
415 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 Henzova-Zirklerova testu ($\alpha = 0.05$) v kombinaci s grafickou vizualizací dat.

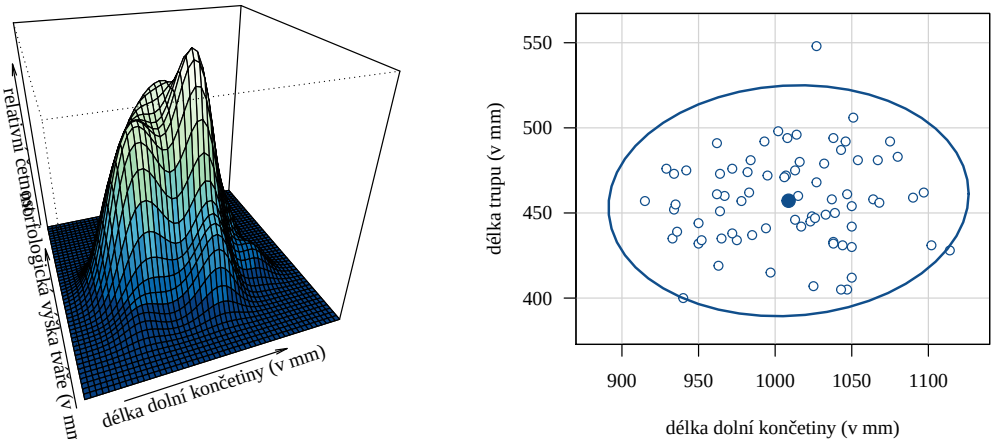
```
416 MVN::mvn(cbind(lowex.LM, tru.LM), mvnTest = 'hz')$multivariateNormality # 0.543489
```

Protože p-hodnota Henzova-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 $= 0.1876 > 0.05$).

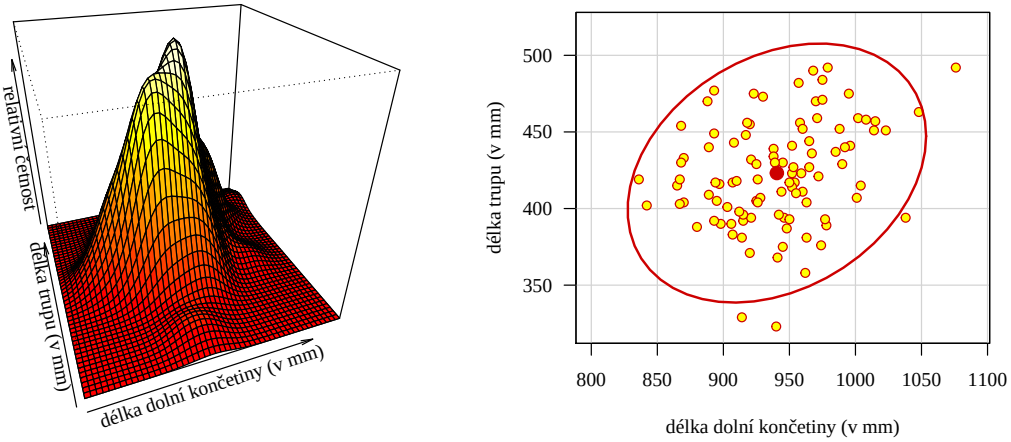
Na 3D grafu vidíme dvourozměné normální rozdělení tvoření jedním kopcem (viz obrázek 40). Aby data pocházela z dvourozměrného normálního rozdělení stačí, aby alespoň 71 bodů leželo uvnitř 95% elipsy spolehlivosti. Zbylé 4 body mohou ležet mimo elipsu spolehlivosti. Z tečkového diagramu vidíme, že mimo 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.

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

Jelikož p-hodnota Henzova-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 40: 3D graf a tečkový diagram s 95% elipsou spolehlivosti pro délku dolní končetiny a délku trupu u mužů (v mm)



Obrázek 41: 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 41). Aby data pocházela z dvourozměnrého 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
  $H_0 : \rho_1 \geq \rho_2 \quad \rightarrow \quad \rho_1 - \rho_2 \geq \rho_0$, kde $\rho_0 = 0$
  $H_1 : \rho_1 < \rho_2 \quad \rightarrow \quad \rho_1 - \rho_2 < \rho_0$, 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.

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

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

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

```
418 alpha <- 0.01
419 ksi0  <- 0
420 r1 <- cor(lowex.LM, tru.LM) # 0.05975781
421 r2 <- cor(lowex.LF, tru.LF) # 0.285256
422 z1 <- 1 / 2 * log((1 + r1) / (1 - r1)) # 0.0598291
423 z2 <- 1 / 2 * log((1 + r2) / (1 - r2)) # 0.2933943
```

   - **Testovací statistika**

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

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

- **Kritický obor**

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

```
425 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.

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

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

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

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

5. **Testování p-hodnotou**

   - **p-hodnota**

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

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

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

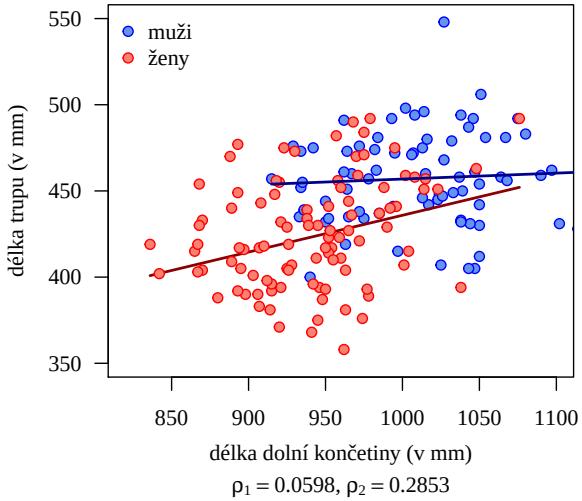
6. **Grafická vizualizace výsledků testování**
   Míru závislosti obou náhodných výběrů vizualizujeme pomocí tečkového diagramu s lineárními křivkami prokládajícími naměřené body (viz obrázek 42). Koeficienty lineární přímky pro délku dolní končetiny a délku trupu mužů, resp. žen, získáme pomocí funkce lm(), analogicky jako v předchozích příkladech.

```
par(mar = c(5, 4, 1, 1), family = 'Times')
plot(lowex.LM, tru.LM, xlab = '', ylab = 'délka trupu (v mm)',
    xlim = c(820, 1100), ylim = c(350, 550), las = 1,
     pch = 21, col = 'blue', bg = 'cornflowerblue')
k <- lm(tru.LF ~ lowex.LF)$coef
x <- seq(min(lowex.LF), max(lowex.LF), length = 1000)
y <- k[1] + x * k[2]
lines(x, y, col = 'darkred', lwd = 2)

points(lowex.LF, tru.LF, pch = 21, col = 'red', bg = 'salmon')
k <- lm(tru.LM ~ lowex.LM)$coef
x <- seq(min(lowex.LM), max(lowex.LM), length = 1000)
y <- k[1] + x * k[2]
lines(x, y, col = 'darkblue', lwd = 2)

r1 <- round(r1, digit = 4)
r2 <- round(r2, digit = 4)
mtext('délka dolní končetiny (v mm)', side = 1, line = 2.3)
mtext(bquote(paste(rho[1] == . (r1), ', ', rho[2] == .(r2))), side = 1, line = 3.7)
legend('topleft', pch = 21, pt.bg = c('cornflowerblue', 'salmon'),
       col = c('blue', 'red'), legend = c('muži', 'ženy'),
       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 42: 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 ($\rho_1 = 0.2853$).

★

Příklad 10.16. Test o rozdílu dvou korelačních koeficientů $\rho_1 - \rho_2$ (nesplnění předpokladů)
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 10.16
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,

```
451 data <- read.delim('00-Data//16-anova-head.txt')
452 data <- na.omit(data)
453 head.WM <- data[data$sex == 'm', 'head.W']
454 bizyg.WM   <- data[data$sex == 'm', 'bizyg.W']
455 head.WF <- data[data$sex == 'f', 'head.W']
456 bizyg.WF   <- data[data$sex == 'f', 'bizyg.W']
457
458 n1 <- length(head.WM) # 63
459 n2 <- length(head.WF) # 85
460 range(head.WM)  # 142-170
461 range(bizyg.WM) # 113-155
462 range(head.WF)  # 135-162
463 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 normalního rozdělení. Před samotným testem hypotézy 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

```
464 MVN::mvn(cbind(head.WM, bizyg.WM), mvnTest = 'royston')$multivariateNormality #
        0.003572831
```

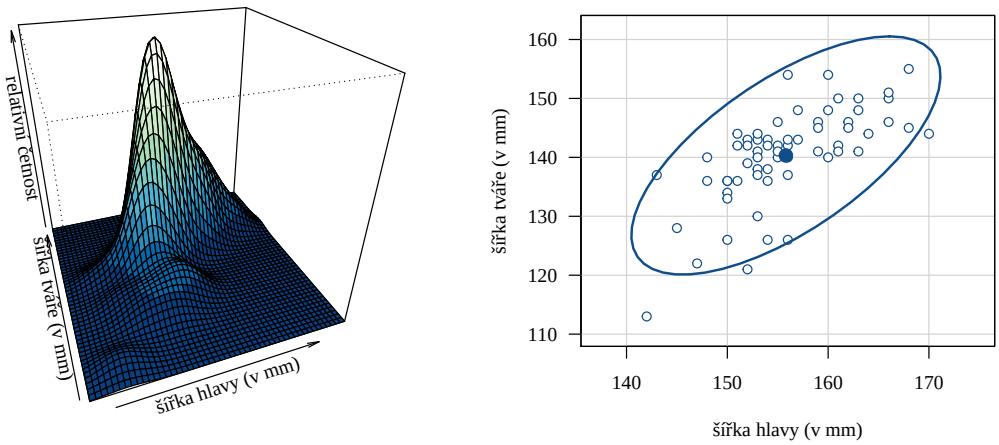
Protože p-hodnota Roystonova testu, tj. 0.003573 je menší než 0.05, hypotézu o dvourozměrné normalitě náhodného výběru šířky lebky a šířky tváře u mužů zamítáme na hladině významnosti $\alpha = 0.05$. Ke stejnému závěru dojdeme také použitím Henzova-Zirklerova testu (p-hodnota $= 0.002498 < 0.05$) i 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 43).

```
Loading required package: viridisLite                                    465
```

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 normalizu 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.

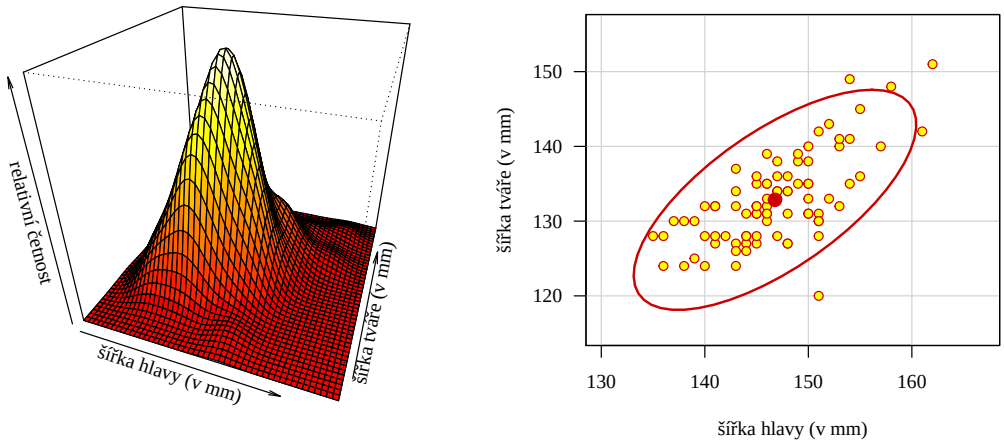
Jelikož p-hodnota Roystonova testu, tj. 0.02646746, je menší než 0.05, hypotézu o dvourozměrné normalitě náhodného výběru šířky hlavy a šířky tváře žen zamítáme na hladině významnosti $\alpha = 0.05$. Ke stejnému závěru



Obrázek 43: 3D graf a tečkový diagram s 95% elipsou spolehlivosti pro šířku hlavy a šířku tváře mužů (v mm)

```
466 MVN::mvn(cbind(head.WF, bizyg.WF), mvnTest = '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 Henzova-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 44).



Obrázek 44: 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 dvourozměné normality obou náhodných výběrů je porušen, nemůžeme k testování hypotézy ze zadání použít parametrický test o rozdílu korelačních koeficientů $\rho_1 - \rho_2$. ★

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

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

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

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

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.

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

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

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

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

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

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

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

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

Příklad 10.17. Test o rozdílu parametrů p_1-p_2 (oboustranná alternativa)
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 10.17
Datový soubor načteme příkazem read.delim() a odstraníme z něj NA hodnoty příkazem na.omit().

```
467 data <- read.delim('00-Data/14-two-samples-probabilities-sexratio.txt')
468 data <- na.omit(data)
469 head(data)
```

```
  sex o.sib.N                                                                    470
1   m       0                                                                    471
2   m       0                                                                    472
3   f       0                                                                    473
4   m       0                                                                    474
5   m       0                                                                    475
6   m       1                                                                    476
```

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 satarší 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

```
477 data0 <- data[data$o.sib.N == 0, ]
478 data1 <- data[data$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

```
479 x1 <- sum(data0$sex == 'm')
480 x2 <- sum(data1$sex == 'm')
481 N1 <- length(data0$sex)
482 N2 <- length(data1$sex)
483 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 zvlášť 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.

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

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

```
484 m1 <- x1 / N1 # 0.5210084
485 m2 <- x2 / N2 # 0.534749
```

Podmínky dobré aproximace potom dopočítáme následujícím způsobem

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

```
486 N1 * m1 * (1 - m1) # 5.804348
```

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

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

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

Jelikož číslo $128.8745 > 5$, je podmínka dobré aproximace populaci novorozenců s jendím 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
     $H_0 : p_1 = p_2 \quad \rightarrow \quad p_1 - p_2 = p_0$, kde $p_0 = 0$
     $H_1 : p_1 \neq p_2 \quad \rightarrow \quad p_1 - p_2 \neq p_0$, kde $p_0 = 0$
     (oboustranná alternativa)

2. **Volba hladiny významnosti**

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

3. **Testování kritickým oborem**

   - **Testovací statistika**

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

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

   - **Kritický obor**

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

```
qnorm(alpha / 2) # -2.575829
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**

$$\begin{aligned} (d,h) &= \left(m_1 - m_2 - \sqrt{\frac{m_1(1-m_1)}{N_1} + \frac{m_2(1-m_2)}{N_2}}u_{1-\alpha/2}\,;\, m_1 - m_2 - \sqrt{\frac{m_1(1-m_1)}{N_1} + \frac{m_2(1-m_2)}{N_2}}u_{\alpha/2}\right) \\ &= \left(0.5210084 - 0.534749 - \sqrt{\frac{0.5210084(1-0.5210084)}{595} + \frac{0.534749(1-0.534749)}{518}}u_{1-0.01/2}\,;\, 0.5210084 - 0.534\right. \\ &= \left(-0.0137406 - \sqrt{0.0004194263 + 0.0004802944}u_{0.975}\,;\, -0.0137406 - \sqrt{0.0004194263 + 0.0004802944}u_{0.005}\right) \\ &= (-0.0137406 - 0.02999534 \times 2.575829\,;\, -0.0137406 - 0.02999534 \times (-2.575829)) \\ &= (-0.09100347\,;\, 0.06352227) \end{aligned}$$

   - **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 $\alpha = 0.01$.

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

5. **Testování p-hodnotou**

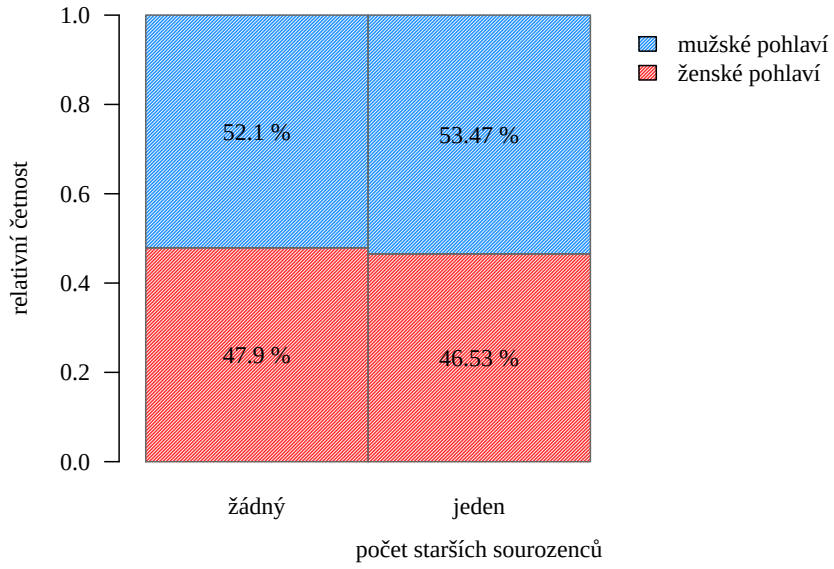
   - **p-hodnota**

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

```r
p.hodnota <- 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 45).



Obrázek 45: 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 10.18. Test o rozdílu parametrů p_1-p_2 (pravostranná alternativa)
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 mechské populace a 87 žen z populace Rajbanshi (viz sekce **??** a tabulky 1 a 2). Na hladině významnosti $\alpha = 0.05$ zjistěte, 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.

Tabulka 1: Frekvence zakončení dlaňových linií u žen z mechské populace

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

Tabulka 2: 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 10.18
Z obou tabulek si vytáhneme pouze relevantní údaje, a sice počet výskytů nízkého zakončení tří dlaňových linií na levé straně u žen mechské 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 mechské populace $N_1 = 50$ a rozsah náhodného výběru žen z populace Rajbanshi $N_2 = 87$.

```
496 x1 <- 19
497 x2 <- 21
498 N1 <- 50
499 N2 <- 87
500 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 mechské 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 mechské 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_1 je pravděpodobnost výskytu nízkého zakončení tří dlaňových linií na levé straně. Řešení příkladu vede na situaci, kdy pravděpodobnost $p1$ porovnáváme s pravděpodobností p_2, tedy na dvouvýběrový test o rozdílu $p_1 - p_2$.Před použitím tetsu 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.

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

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

```
501  m1 <- x1 / N1 # 0.38
502  m2 <- x2 / N2 # 0.2413793
```

Podmínky dobré aproximace potom dopočítáme následujícím způsobem

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

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

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

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

```
504  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 frekveni 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
     $H_0 : p_1 \leq p_2 \quad \rightarrow \quad p_1 - p_2 \leq p_0$, kde $p_0 = 0$
     $H_1 : p_1 > p_2 \quad \rightarrow \quad p_1 - p_2 > p_0$, 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**

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

```
505 alpha <- 0.05
506 p0 <- 0
507 zw <- ((m1 - m2) - p0) / (sqrt(m1 * (1 - m1) / N1 + m2 * (1 - m2) / N2)) # 1.678953
```

- **Kritický obor**

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

```
508 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**

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

```
509 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 levostranného intervalu spolehlivosti, tj. $p_0 = 0 \notin IS$, H_0 zamítáme na hladině významnosti $\alpha = 0.05$.

5. **Testování p-hodnotou**

   - **p-hodnota**

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

```
510 p.hodnota <- 1 - pnorm(zw) # 0.04658058
```

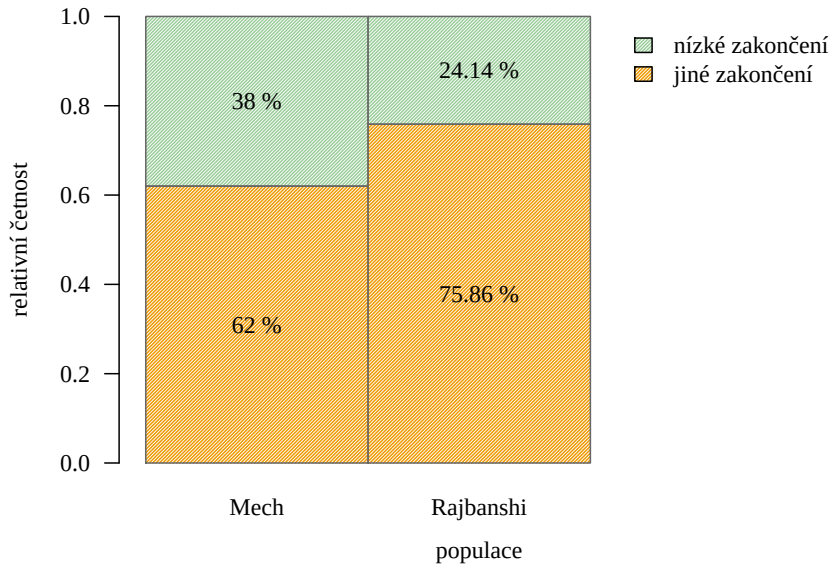
   - **Závěr testování**
     Protože p-hodnota $= 0.04658$ je menší než 0.05, H_0 zamítáme na hladině významnosti $\alpha = 0.05$.

6. **Grafická vizualizace výsledků testování**
   Rozdíly mezi oběma populacemi vizualizujeme sloupcovým diagramem relativních četností (viz obrázek 46).

```
511 source('Sbirka-AS-I-2018-funkce.R')
512 data <- cbind(c(19, N1 - 19), c(21, N2 - 21))
513 par(mar = c(5, 4, 1, 1), family = 'Times')
514 rel.barplot.two(data, col = c('orange2', 'darkseagreen3'),
515                 border = 'grey40',
516                 names = c('Mech', 'Rajbanshi'),
517                 legend = c('nízké zakončení', 'jiné zakončení'),
518                 density = 60, col.number = 'black', main = '',
519                 xlab = 'populace', ylab = 'relativní četnost')
```

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 mechské populace je statisticky významně vyšší než frekvence výskytu nízkého zakončení dlaňových linií na levé straně u žen z populace Rajbanshi.

★



Obrázek 46: Sloupcový diagram relativních četností zakončení dlaňových linií z levé strany u žen z mechské populace a žen z populace v Rajbanshi

Příklad 10.19. Test o rozdílu parametrů p_1-p_2 (levostranná alternativa)
Mějme k dispozici údaje o frekvenci výskytu epigenetického znaku *sutura metopica* (binomické proměnná) na lebkách jedinců z Anatolské populace (viz tabulka 3; (Eroglu, 2008)).

Tabulka 3: Výskyt epigenetického znaku *sutura metopica* u jedinců Anatolské populace

	sutura metopica	$\sum$
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 10.19
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().

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

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

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.

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

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

```r
N1 <- 334
N2 <- 153
x1 <- 26
x2 <- 15

m1 <- x1 / N1 # 0.07784431
m2 <- x2 / N2 # 0.09803922
```

Podmínky dobré aproximace potom dopočítáme následujícím způsobem

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

```r
N1 * m1 * (1 - m1) # 23.97605
```

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

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

```r
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
     $H_0 : p_1 \geq p_2 \quad \rightarrow \quad p_1 - p_2 \geq p_0$, kde $p_0 = 0$
     $H_1 : p_1 < p_2 \quad \rightarrow \quad p_1 - p_2 < p_0$, kde $p_0 = 0$
     (levostranná alternativa)

2. **Volba hladiny významnosti**

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

3. **Testování kritickým oborem**

- **Testovací statistika**

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

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

- **Kritický obor**

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

```
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**

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

```
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**

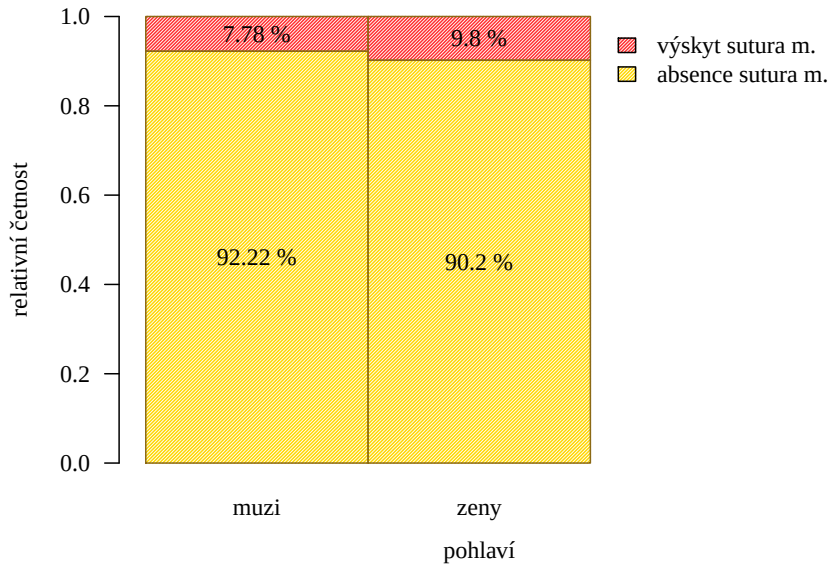
- **p-hodnota**

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

```
539 p.hodnota <- 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 47), 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 47: 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 10.20. Test o rozdílu parametrů p_1-p_2 (nesplnění předpokladů)
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 10.20
Datový soubor načteme příkazem read.delim() a odstraníme z něj NA hodnoty příkazem na.omit().

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

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

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 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%.

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

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

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 zvlášť 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

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

```
553 N1 * m1 * (1 - m1) # 24.55455
```

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

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

```
554 N2 * m2 * (1 - m2) # 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 nulovu 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 10.20 vypořádat. První možností je zamyslet se, zda by bylo možné doplnit datový soubor o další údaje. Doplnění údajů není v příkladě 10.20 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 10.20 použili. Je však nezbytné pracovat s tímto postupem velmi opatrně, a při závažnějším porušení byť jen jedné podmínky dobré aproximace volit raději jiné, vhodnější metody, nebo zauvažovat o rozšíření datového souboru.

V případě příkladu 10.20 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 **??**.

10.6 Test o logaritmu podílu šancí

V této kapitole si ukážeme další přístup, pomocí kterého můžeme porovnávat dva náhodné výběry z alternativních rozdělení. Myšlenka zmíněného přístupu tkví v tom, že máme dva náhodné výběry z alternativních rozdělení, o kterých předpokládáme, že popisují výskyt stejné sledované události, ovšem za dvou různých okolností (I a II). První náhodný výběr popisuje výskyt sledované události za okolnosti I, druhý náhodný výběr popisuje výskyt tentýž sledované události ovšem za okolnosti II.

Předpokládejme tedy nyní, že $X_{11}, \dots X_{1N_1}$ je náhodný výběr z alternativního rozdělení $\text{Alt}(p_1)$. Tento výběr popisuje výskyt sledované události za okolnosti I. Zde je důležité si uvědomit, že náhodná veličina X_{1i}, $i = 1, \dots, N_1$, jakožto náhodná veličina pocházející z alternativního rozdělení, nabývá buď hodnoty 1 (v případě, že sledovaná událost nastala) nebo hodnoty 0 (v případě, že sledovaná událost nenastala). Celý náhodný výběr je potom soubor n_{11} jedniček, určujících, kolikrát v N_1 pokusech sledovaná událost nastala, a n_{12} nul, určujících, kolikrát v N_1 pokusech sledovaná událost nenastala. Součet celkového počtu jedniček n_{11} a celkového počtu nul n_{12} je potom rovný celkovému počtu pokusů N_1, tj. $n_{11} + n_{12} = N_1$.

Analogicky předpokládejme, že $X_{21}, \dots X_{2N_2}$ je na náhodný výběr z alternativního rozdělení $\text{Alt}(p_2)$, který je nezávislý na náhodném výběru $X_{11}, \dots X_{1N_1}$. Tento výběr popisuje výskyt sledované události za okolnosti II. Také každá náhodná veličina $X_{2j}, j = 1, \dots, N_2$, jakožto náhodná veličina pocházející z alternativního rozdělení, nabývá buď hodnoty 1 (v případě, že sledovaná událost nastala) nebo hodnoty 0 (v případě, že sledovaná událost nenastala). Celý náhodný výběr je potom soubor n_{21} jedniček, určujících, kolikrát v N_2 pokusech sledovaná událost nastala, a n_{22} nul, určujících, kolikrát v N_2 pokusech sledovaná událost nenastala. Součet celkového počtu jedniček n_{21} a celkového počtu nul n_{22} je potom rovný celkovému počtu pokusů N_2, tj. $n_{21} + n_{22} = N_2$.

Celkový počet pokusů provedených za okolnosti I nebo za okolnosti II značíme jako N, tj. $N_1 + N_2 = N$. Sečtením počtu výskytů sledované události za okolnosti I a počtu výskytů sledované události za okolnosti II, tj. $n_{11} + n_{21}$, získáme celkový počet výskytů sledované události ve všech N pokusech. Sečtením počtu nenastání sledované události za okolnosti I a počtu nenastání sledované události za okolnosti II, tj. $n_{12} + n_{22}$, získáme celkový počet nenastání sledované události ve všech N pokusech.

Počet nastání a nenastání sledované události za nastání okolnosti I nebo okolnosti II můžeme přehledně zavést do čtyřpolní kontingenční tabulky.

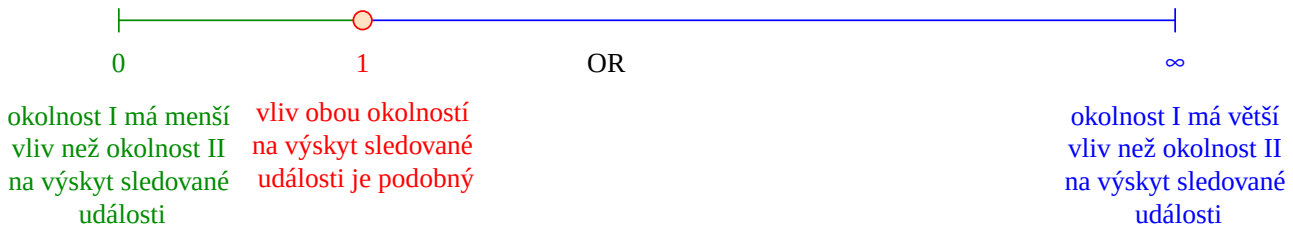
Okolnost	Sledovaná událost		$\sum$
	nastala	nenastala	
I	n_{11}	n_{12}	$n_{11} + n_{12}$
II	n_{21}	n_{22}	$n_{21} + n_{22}$
$\sum$	$n_{11} + n_{21}$	$n_{12} + n_{22}$	N

Podíl počtu úspěchů ku počtu neúspěchů n_{11}/n_{12} vyjadřuje šanci, že sledovaná událost za okolnosti I nastane. Analogicky podíl počtu úspěchů ku počtu neúspěchů n_{21}/n_{22} vyjadřuje šanci, že sledovaná událost za okolnosti II nastane. Konečně,

$$OR = \frac{n_{11}/n_{12}}{n_{21}/n_{22}} = \frac{n_{11}n_{22}}{n_{12}n_{21}} \tag{10.9}$$

popisuje podíl šancí nastání sledované události za okolnosti I ku nastání sledované události za okolnosti II. Podíl OR vyjádřený vztahem 10.9 je statistikou a nazývá se *podíl šancí*. Protože podíl šancí OR je sestaven pouze z početností, které jsou vždy kladné nebo nulové (není možné, aby sledovaná událost nastala $-2\times$, apod.), bude se jeho hodnota vždy pohybovat v intervalu od nuly do nekonečna, tj. $OR \in (0\,;\,\infty)$.

Zaměřme si nyní na to, co nám hodnota podílu šancí OR říká o nastání sledované události. Pohybuje-li se hodnota podílu šancí okolo 1, znamená to, že čitatel a jmenovatel nabývají podobné hodnoty, a tedy, že šance, že sledovaná událost nastane za okolnosti I je podobná šanci, že sledovaná událost nastane za okolnosti II. Je-li tedy hodnota podílu šancí OR blízká 1, znamená to, že okolnosti I a II nemají na sledovanou událost významný vliv, protože šance na nastání události za okolnosti I a II si jsou podobné. Naopak, čím více se hodnota podílu šancí OR vzdaluje od 1 (tj. čím více se OR blíží k nule nebo k nekonečnu), tím více se liší šance na nastání sledované události za okolnosti I a za okolnosti II. Zmíněné poznatky přehledně znázorníme na obrázku 48.
Z obrázku 48 je patrná jedna nežádoucí vlastnost podílu šancí OR, a sice asymetrie okolo 1. Z grafu na první pohled vidíme, že oblast vyjadřující větší vliv okolnosti I na výskyt sledované události se realizuje v úzkém intervalu $(0\,;\,1)$. Naopak oblast vyjadřující větší vliv okolnosti II na výskyt sledované události se realizuje v širokém intervalu $(1\,;\,\infty)$. Tuto nežádoucí vlastnost odstraníme aplikováním funkce přirozeného logaritmu $\ln(x)$ na podíl šancí OR.



Obrázek 48: Vizualizace podílu šancí OR

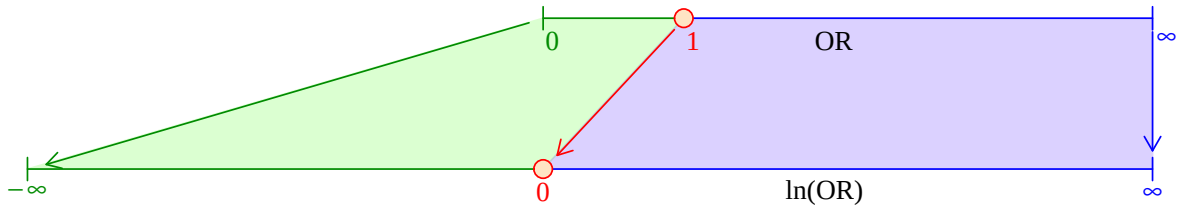
Poznámka: Funkce přirozeného logaritmu $f(x) = \ln(x)$ je funkce proměnné x s definičním oborem $D(f) = (0\,;\,\infty)$ a s oborem hodnot $H(f) = (-\infty\,;\,\infty)$. Jednoduše řečeno je to funkce, která tranformuje hodnoty z intervalu $(0\,;\,\infty)$ do intervalu $(-\infty\,;\,\infty)$ podle následujícího předpisu:

$$\begin{aligned} x \in (0\,;\,1) &\rightarrow f(x) \in (-\infty\,;\,0) \\ x = 1 &\rightarrow f(x) = 0 \\ x = (1\,;\,\infty) &\rightarrow f(x) \in (0\,;\,\infty) \end{aligned}$$

Zmíněným aplikováním funkce přirozeného logaritmu $\ln(x)$ na podíl šancí OR získáváme tzv. *logaritmus podílu šancí* $\ln OR$

$$\ln OR = \ln(OR) = \ln\left(\frac{n_{11}/n_{12}}{n_{21}/n_{22}}\right) = \ln\left(\frac{n_{11}n_{22}}{n_{12}n_{21}}\right) \tag{10.10}$$

Z obrázku 49 je potom patrné, jak jsme pomocí funkce $f(x) = \ln(x)$ převedli podíl šancí OR na logaritmus podílu šancí $\ln OR$, a získali tak statistiku $\ln OR$, která je symetrická okolo nuly.



Obrázek 49: Vizualizace transformace podílu šancí OR na logaritmus podílu šancí $\ln OR$ prostřednictvním funkce přirozeného logaritmu $\ln(x)$

Protože statistika $\ln OR$ má lepší vlastnosti než statistika OR, využijeme právě ji při testování rozdílu mezi šancemi na výskyt sledované události za dvou různých okolností I a II. Výše jsme si uvedli, že okolnosti I a II nemají odlišný vliv na sledovanou událost, pokud se podíl šancí OR pohybuje okolo hodnoty 1. Protože nyní již víme, že logaritmická transformace zobrazí hodnotu 1 na hodnotu 0, můžeme tvrzení z předchozí věty rozšířit a konstatovat, že okolnosti I a II nemají odlišný vliv na sledovanou událost, pokud se logaritmus podílu šancí $\ln OR$ pohybuje okolo hodnoty 0.

Nechť $\ln o\rho$ je teoretická hodnota logaritmu podílu šancí a $\ln o\rho_0$ je konstanta. Na hladině významnosti α testujeme jednu z následujících tří hypotéz oproti příslušné alternativní hypotéze.

$H_{01} : \ln o\rho = \ln o\rho_0$	oproti	$H_{11} : \ln o\rho \neq \ln o\rho_0$	(oboustranná alt.)
$H_{02} : \ln o\rho \leq \ln o\rho_0$	oproti	$H_{12} : \ln o\rho > \ln o\rho_0$	(pravostranná alt.)
$H_{03} : \ln o\rho \geq \ln o\rho_0$	oproti	$H_{13} : \ln o\rho < \ln o\rho_0$	(levostranná alt.)

Test nazýváme dvouvýběrovým Z-testem o logaritmu podílu šancí. Testovací statistika má tvar

$$Z_W = \frac{\ln OR - \ln o\rho_0}{\sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}}}, \tag{10.11}$$

kde $\ln OR = \ln\left(\frac{n_{11}n_{22}}{n_{12}n_{21}}\right)$ je logaritmus podílu šancí. Testovací statistika Z_W pochází asymptoticky ze standardizovaného normálního rozdělení, tj.

$$Z_W = \frac{\ln OR - \ln o\rho_0}{\sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}}} \overset{\mathcal{A}}{\sim} N(0,1).$$

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.

Kritický obor podle zvolené alternativní hypotézy má tvar

$$\begin{array}{ll} H_{11} : \ln o\rho \neq \ln o\rho_0 & W = \left(-\infty\,;\, u_{\alpha/2}\right\rangle \cup \left\langle u_{1-\alpha/2}\,;\, \infty\right) \\ H_{12} : \ln o\rho > \ln o\rho_0 & W = \left\langle u_{1-\alpha}\,;\, \infty\right) \\ H_{13} : \ln o\rho < \ln o\rho_0 & W = \left(-\infty\,;\, u_{\alpha}\right\rangle \end{array}$$

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

Interval spolehlivosti má podle zvolené alternativní hypotézy jeden z následujících tvarů

$$\begin{array}{ll} H_{11} : \ln o\rho \neq \ln o\rho_0 & (d,h) = \left(\ln OR - \sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}}u_{1-\alpha/2}\,;\, \ln OR - \sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}}u_{\alpha/2}\right) \\ H_{12} : \ln o\rho > \ln o\rho_0 & (d,1) = \left(\ln OR - \sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}}u_{1-\alpha}\,;\, \infty\right) \\ H_{13} : \ln o\rho < \ln o\rho_0 & (-1,h) = \left(-\infty\,;\, \ln OR - \sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}}u_{\alpha}\right) \end{array}$$

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

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

kde Z_W je náhodná veličina, z_W je realizace testovací statistiky Z_W (viz vzorec 10.11), tedy konkrétní číslo, a $\Pr(Z_W \leq z_W)$ je distribuční funkce standardizovaného normálního rozdělení, jejíž hodnotu získáme pomocí R a implementované funkce pnorm().

Poznámka: Pomocí testu o podílu šancí můžeme také testovat hypotézu H_{01}: *Náhodné veličiny X a Y jsou o stochasticky nezávislé.* Tato hypotéza odpovídá znění nulové hypotézy $H_{01} : \ln o\rho = 0$. Znění alternativní hypotézy H_{11} : *Náhodné veličiny X a Y nejsou stochasticky nezávislé.* potom odpovídá znění alternativní hypotézy H_{01} : $\ln o\rho \neq 0$. Poznamenejme, že jednostranné hypotézy v tomto případě testovat nemůžeme, protože X a Y jako znaky nominálního typu neumožňují diagnostikovat přímou nebo nepřímou závislost (viz. kapitola **??**).

Příklad 10.21. Test o logaritmu podílu šancí $\ln OR$ (oboustranná alternativa)
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 šance na narození chlapce je stejná u novorozenců s žádným sourozencem a u novorozenců s jedním sourozencem.

Řešení příkladu 10.21
Zadání příkladu je analogická zadání příkladu 10.17. Hypotéza uvedená v zadání byla akorát mírně modifikována tak, aby odpovídala testu o logaritmu podílu šancí. Sledovanou událostí je v tomto případě narození novorozence mužského pohlaví, přičemž výskyt této události nás zajímá jednak za situace, kdy novorozenec nemá žádného sourozence (okolnost I) a jednak za situace, kdy novorozenec má jednoho staršího sourozence (okolnost II). V příkladu 10.17 jsme si již udělali rozbor zadání, v rámci kterého jsme zjistili, že z celkového počtu 595 novorozenců s žádným starším sourozencem je 310 novorozenců mužského pohlaví a zbylých 285 novorozenců je ženského pohlaví. Dále jsme zjistili, že z celkového počtu 518 novorozenců s jedním starším sourozencem je 277 novorozenců mužského pohlaví a zbylých 241 novorozenců je ženského pohlaví. Zjištěné údaje můžeme poskládat do čtyřpolní kontingenční tabulky 4.

Tabulka 4: Kontingenční tabulka absolutních četností narození chlapce s žádným nebo jedním starším sourozencem

Počet starších sourozenců	Narození chlapce		$\sum$
	ano	ne	
žádný (okolnost I)	310	285	595
jeden (okolnost II)	277	241	518
$\sum$	587	526	1113

Kontingenční tabulku vytvoříme elegantně pomocí softwaru R. Datový soubor nejprve načteme, příkazem na.omit() odstraníme řádky s chybějícími údaji, a pomocí funkce head() vypíšeme první tři řádky tabulky.

```
data <- read.delim('00-Data/14-two-samples-probabilities-sexratio.txt')
data <- na.omit(data)
head(data, n = 3)
```

```
  sex o.sib.N
1   m       0
2   m       0
3   f       0
```

Datový soubor obsahuje dvě proměnné sex a o.sib.N. Podíváme se nyní blíže na proměnnou sex.

```
[1] m m f m m m
Levels: f m
```

Z výstupu příkazu head(data$sex) vidíme, že software R automaticky vyhodnotil proměnnou sex jako proměnnou typu factor se dvěma úrovněmi f a m. Jako referenční vybral software R úroveň f, protože je první v abecedě. To se nám však příliš nehodí, protože v kontextu našeho příkladu chceme mít jako referenční úroveň mužské pohlaví, tedy úroveň m. Pomocí funkce factor() se specifikací argumentu levels změníme referenční úroveň z úrovně f na m. Proměnná o.sib.N byla vyhodnocena jako proměnná typu integer, což je v pořádku, protože obsahuje číselné hodnoty 0 a 1.

```
data$sex <- factor(data$sex, levels = c('m', 'f'))
head(data$sex)
```

```
[1] m m f m m m
Levels: m f
```

Nyní vytvoříme kontingenční tabulku simultánních absolutních četností pomocí příkazu table. Vstupními argumenty budou vektory o.sib.N a sex z tabulky data, čímž si zajistíme, že ve sloupcích kontingenční tabulky budou údaje dělené podle pohlaví novorozenců a v řádcích budou údaje dělené podle počtu starších sourozenců. Vytvoříme si tak řízeně tabulku analogickou tabulce ??. Výstup funkce table() je tabulka absolutních četností typu table. Tento typ je však nepoužitelý pro další výpočty, proto použijeme funkci as.data.frame.matrix(), která převede tabulku typu table na tabulku typu data.frame. Do tabulky následně příkazem names doplníme názvy sloupců a příkazem row.names() doplníme názvy řádků.

```
tab <- table(data$o.sib.N, data$sex)
tab <- as.data.frame.matrix(tab)
names(tab) <- c('chlapec', 'devce')
row.names(tab) <- c('zadny', 'jeden')
tab
```

```
      chlapec devce
zadny     310   285
jeden     277   241
```

Nyní si vypočítáme hodnotu podílu šancí OR a logaritmu podílu šancí. Z kontingenční tabulky 4 známe konkrétní hodnoty simultánních absolutních četností, zde konkrétně $n_{11} = 285$, $n_{12} = 310$, $n_{21} = 241$ a $n_{22} = 277$. Dosazením hodnot do vztahu 10.9 získáme hodnotu podílu šancí.

$$OR = \frac{n_{11}n_{22}}{n_{12}n_{21}} = \frac{285 \times 277}{310 \times 241} = \frac{78945}{74710} = 1.056686 \doteq 1.05669$$

```
n11 <- tab[1, 1]
n12 <- tab[1, 2]
n21 <- tab[2, 1]
n22 <- tab[2, 2]
OR <- (n11 * n22) / (n12 * n21) # 1.056686
```

Interpretace podílu šancí: Hodnota podílu šancí $OR = 1.05669$. Šance, že novorozenec bude mužského pohlaví, je 1.05669 krát větší u novorozence, který nemá žádného staršího sourozence, než u novorozence, který má jednoho staršího sourozence.

Aplikováním přirozeného logaritmu na podíl šancí OR získáme hodnotu logaritmu podílu šancí $\ln OR$.

$$\ln OR = \ln(OR) = \ln(1.056686) = 0.05513746 \doteq 0.05514.$$

```
lnOR <- log(OR) # 0.05513746
```

Hodnota logaritmu podílu šancí $\ln OR = 0.05514$.

Nyní již můžeme přistoupit k procesu testování nulové hypotézy. V zadání příkladu máme uvedeno přesné znění nulové hypotézy. Zbývá tedy stanovit alternativní hypotézu tak, aby byla doplňkem k hypotéze nulové. K vytvoření správného matematického zápisu nulové hypotézy si stačí uvědomit, že shodě šancí na výskyt události za okolnosti I a okolnosti II odpovídá nulová hodnota logaritmu podílu šancí $\ln o\rho$, tj. $\ln o\rho = 0$.

1. **Stanovení hypotéz**
   - **slovní formulace** nulové a alternativní hypotézy
     H_0 : *Šance na narození chlapce u novorozenců s žádným sourozencem a u novorozenců s jedním sourozencem jsou stejné.*
     H_1 : *Šance na narození chlapce u novorozenců s žádným sourozencem a u novorozenců s jedním sourozencem nejsou stejné.*

- **matematická formulace** nulové a alternativní hypotézy
  $H_0 : \ln o\rho = \ln o\rho_0$, kde $\ln o\rho_0 = 0$
  $H_1 : \ln o\rho \neq \ln o\rho_0$, kde $\ln o\rho_0 = 0$
  (oboustranná alternativa)

2. **Volba hladiny významnosti**

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

3. **Testování kritickým oborem**

   - **Testovací statistika**

$$\begin{aligned} Z_W &= \frac{\ln OR - \ln o\rho_0}{\sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}}} \\ &= \frac{0.05513746 - 0}{\sqrt{\frac{1}{285} + \frac{1}{310} + \frac{1}{241} + \frac{1}{277}}} \\ &= \frac{0.05513746}{\sqrt{0.003508772 + 0.003225806 + 0.004149378 + 0.003610108}} \\ &= \frac{0.05513746}{\sqrt{0.01449406}} \\ &= \frac{0.05513746}{0.1203913} = 0.4579854 \doteq 0.4580 \end{aligned}$$

```
582 alpha <- 0.01
583 ln.orho0 <- 0
584 zw <- (lnOR - ln.orho0) / sqrt(1 / n11 + 1 / n12 + 1 / n21 + 1 / n22) # 0.4579854
```

   - **Kritický obor**

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

```
585 qnorm(alpha / 2) # -2.575829
586 qnorm(1 - alpha / 2) # 2.575829
```

   - **Závěr testování**
     Protože realizace testovací statistiky $z_W = 0.4580$ nenáleží do kritického oboru, tj. $z_W \notin W$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

4. **Testování intervalem spolehlivosti**

- **Interval spolehlivosti**

$$
\begin{aligned}
(d,h) &= \left(\ln OR - \sqrt{\frac{1}{n_{11}}+\frac{1}{n_{12}}+\frac{1}{n_{21}}+\frac{1}{n_{22}}}u_{1-\alpha/2}\,;\,\ln OR - \sqrt{\frac{1}{n_{11}}+\frac{1}{n_{12}}+\frac{1}{n_{21}}+\frac{1}{n_{22}}}u_{\alpha/2}\right)\\
&= \left(0.05513746 - \sqrt{\frac{1}{285}+\frac{1}{310}+\frac{1}{241}+\frac{1}{277}}u_{1-0.01/2}\,;\,0.05513746 - \sqrt{\frac{1}{285}+\frac{1}{310}+\frac{1}{241}+\frac{1}{277}}u_{0.01/2}\right)\\
&= \left(0.05513746 - \sqrt{0.003508772+0.003225806+0.004149378+0.003610108}u_{0.995}\,;\,0.05513746 - \sqrt{0.003508772}\right.\\
&= \left(0.05513746 - \sqrt{0.01449406}\times 2.575829\,;\,0.05513746 - \sqrt{0.01449406}\times(-2.575829)\right)
\end{aligned}
$$

```
dh <- lnOR - sqrt(1 / n11 + 1 / n12 + 1 / n21 + 1 / n22) * qnorm(1 - alpha / 2) # -0.25497
hh <- lnOR - sqrt(1 / n11 + 1 / n12 + 1 / n21 + 1 / n22) * qnorm(alpha / 2) # 0.3652449
```

- **Závěr testování**
  Protože $\ln o\rho_0 = 0$ náleží do Waldova 99% empirického oboustranného intervalu spolehlivosti, tj. $\ln o\rho_0 = 0 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

5. **Testování p-hodnotou**

- **p-hodnota**

$$
\begin{aligned}
p\text{-hodnota} &= 2\min\{\Pr(Z_W \le z_W), \Pr(Z_W > z_W)\}\\
&= 2\min\{\Pr(Z_W \le z_W), 1 - \Pr(Z_W \le z_W)\}\\
&= 2\min\{\Pr(Z_W \le 0.4579854), 1 - \Pr(Z_W \le 0.4579854)\}\\
&= 2\min\{0.6765185, 0.3234815\}\\
&= 2 \times 0.3234815\\
&= 0.646963 \doteq 0.6470
\end{aligned}
$$

```
p.hodnota <- 2 * min(pnorm(zw), 1 - pnorm(zw)) # 0.6469629
```

- **Závěr testování**
  Protože p-hodnota $= 0.6470$ je větší než $\alpha = 0.01$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

6. **Interpretace výsledků:** Na základě všech tří způsobů testování nezamítáme hypotézu H_0. Hodnota podílu šancí OR vyšla 1.05669. Šance, že novorozenec bude mužského pohlaví, je 1.05669 krát větší u novorozence, který nemá žádného staršího sourozence, než u novorozence, který má jednoho staršího sourozence. Tento rozdíl v šancích však nevyšel jako statisticky významný. Ke stejnému závěru testování jsme dospěli taktéž v řešení příkladu 10.17.

★

Příklad 10.22. Test o logaritmu podílu šancí $\ln OR$ (pravostranná alternativa)
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 mechské populace a 87 žen z populace Rajbanshi (viz sekce **??** a tabulky 5 a 6). Na hladině významnosti $\alpha = 0.05$ zjistěte, zda je šance na výskyt nízkého zakončení tří dlaňových linií na levé straně u žen z mechské populace vyšší než šance na výskyt nízkého zakončení tří dlaňových linií na levé straně u žen z populace Rajbanshi.

Tabulka 5: Frekvence zakončení dlaňových linií u žen z mechské populace

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

Tabulka 6: 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 10.22
Zadání příkladu je analogické zadání příkladu 10.18. Otázka uvedená v zadání byla pouze mírně modifikovaná tak, aby odpovídala hypotéze příslušné k testu o logaritmu podílu šancí. Sledovanou událostí je v tomto případě nízké zakončení tří dlaňových linií na levé straně, přičemž výskyt nízkého zakončení dlaňových linií sledujeme jednak u žen z mechské populace (okolnost I) a jednak u žen z populace Rajbanshi (okolonost II). V úvodu příkladu 10.18 jsme provedli rozbor zadání, kde jsme zjistili, že z celkového počtu 50 žen z mechské populace byl zaznamenán výskyt nízkého zakončení tří dlaňových linií u 19 žen. U zbylých 31 žen byl zaznamenán výskyt vysokého, středního nebo jiného zakončení. Dále jsme zjistili, že z celkového počtu 87 žen z populace Rajbanshi byl zaznamenán výskyt nízkého zakončení dlaňových linií u 21 žen, zatímco u zbývajících 66 žen byl opět zaznamenán výskyt vysokého, středního nebo jiného zakončení. Zjištěné údaje přehledně shrneme do čtyřpolní kontingenční tabulky 7.

Tabulka 7: Kontingenční tabulka absolutních četností výskytu nízkého zakončení dlaňových linií u mechské ženské populace a u žen z populace Rajbanshi

Populace žen	Nízké zakončení dlaňových linií		$\sum$
	ano	ne	
mechská (okolnost I)	19	31	50
Rajbanshi (okolnost II)	21	66	87
$\sum$	40	97	137

Protože nemáme k dispozici datový soubor, vytvoříme čtyřpolní kontingenční tabulku simultíánních absolutních četností pomocí příkazu data.frame().

```
          ano ne
mechska    19 31
Rajbanshi  21 66
```

Nyní si vypočítáme hodnotu podílu šancí OR a logaritmu podílu šancí. Z kontingenční tabulky 7 známe konkrétní hodnoty simultánních absolutních četností, zde konkrétně $n_{11} = 19$, $n_{12} = 31$, $n_{21} = 21$ a $n_{22} = 66$. Dosazením hodnot do vztahu 10.9 získáme hodnotu podílu šancí.

$$OR = \frac{n_{11}n_{22}}{n_{12}n_{21}} = \frac{19 \times 66}{31 \times 21} = \frac{1254}{651} = 1.926267 \doteq 1.9263$$

```r
(tab <- data.frame(ano = c(19, 21), ne = c(31, 66),
                   row.names = c('mechska', 'Rajbanshi')))
```

```r
n11 <- tab[1, 1]
n12 <- tab[1, 2]
n21 <- tab[2, 1]
n22 <- tab[2, 2]
OR <- (n11 * n22) / (n12 * n21) # 1.926267
```

Interpretace podílu šancí: Hodnota podílu šancí $OR = 1.9263$. Šance na výskyt nízkého zakončení dlaňových linií u žen mechské populace je 1.9263 krát větší než u žen z populace Rajbanshi.

Aplikováním přirozeného logaritmu na podíl šancí OR získáme hodnotu logaritmu podílu šancí $\ln OR$.

$$\ln OR = \ln(OR) = \ln(1.926267) = 0.6555839 \doteq 0.6556.$$

```r
lnOR <- log(OR) # 0.6555841
```

Hodnota logaritmu podílu šancí $\ln OR = 0.6556$.

V tomto okamžiku se již můžeme zaměřit na otázku v zadání. Naším úkolem je zjistit, zda je šance na výskyt nízkého zakončení tří dlaňových linií na levé straně u žen z mechské populace vyšší než u žen z populace Rajbanshi. Protože v zadání není zmínka o nulové hypotéze ani o rovnosti, je tvrzení z předchozí věty zněním alternativní hypotézy. Matematický ekvivalent ke slovní formulaci alternativní hypotézy potom odpovídá výroku, že logaritmus podílu šancí je větší než nula. Zbývá tedy dodefinovat slovní a matematickou formulaci nulové hypotézy tak, aby byla doplňkem hypotézy alternativní.

1. **Stanovení hypotéz**
   - **slovní formulace** nulové a alternativní hypotézy
     H_0 : *Šance na výskyt nízkého zakončení tří dlaňových linií na levé straně u žen z mechské populace je nižší nebo rovna šanci na výskyt nízkého zakončení tří dlaňových linií na levé straně u žen z populace Rajbanshi.*
     H_1 : *Šance na výskyt nízkého zakončení tří dlaňových linií na levé straně u žen z mechské populace je vyšší než šance na výskyt nízkého zakončení tří dlaňových linií na levé straně u žen z populace Rajbanshi.*
   - **matematická formulace** nulové a alternativní hypotézy
     $H_0 : \ln o\rho \leq \ln o\rho_0$, kde $\ln o\rho_0 = 0$
     $H_1 : \ln o\rho > \ln o\rho_0$, kde $\ln o\rho_0 = 0$ (pravostranná alternativa)
2. **Volba hladiny významnosti**
   - Hladinu významnosti volíme v souladu se zadáním $\alpha = 0.05$.
3. **Testování kritickým oborem**

- **Testovací statistika**

$$\begin{aligned} Z_W &= \frac{\ln OR - \ln o\rho_0}{\sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}}} \\ &= \frac{0.6555841 - 0}{\sqrt{\frac{1}{19} + \frac{1}{31} + \frac{1}{21} + \frac{1}{66}}} \\ &= \frac{0.6555841}{\sqrt{0.05263158 + 0.03225806 + 0.04761905 + 0.01515152}} \\ &= \frac{0.6555841}{\sqrt{0.1476602}} \\ &= \frac{0.6555841}{0.3842658} = 1.706069 \doteq 1.706069 \end{aligned}$$

```
601 alpha <- 0.05
602 ln.orho0 <- 0
603 zw <- (lnOR - ln.orho0) / sqrt(1 / n11 + 1 / n12 + 1 / n21 + 1 / n22) # 1.706069
```

- **Kritický obor**

$$\begin{aligned} W &= \langle u_{1-\alpha}\,;\,\infty) \\ &= \langle u_{1-0.05}\,;\,\infty) \\ &= \langle u_{0.95}\,;\,\infty) \\ &= \langle 1.644854\,;\,\infty) \end{aligned}$$

```
604 qnorm(1 - alpha) # 1.644854
```

- **Závěr testování**
  Protože realizace testovací statistiky $z_W = 1.706069$ náleží do kritického oboru, tj. $z_W \in W$, H_0 zamítáme na hladině významnosti $\alpha = 0.05$.

4. **Testování intervalem spolehlivosti**
   Proti pravostranné alternativní hypotéze postavíme levostranný interval spolehlivosti.

- **Interval spolehlivosti**

$$\begin{aligned} (d, h) &= \left(\ln OR - \sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}}u_{1-\alpha/2}\,;\,\infty\right) \\ &= \left(0.6555841 - \sqrt{\frac{1}{19} + \frac{1}{31} + \frac{1}{21} + \frac{1}{66}}u_{1-0.01/2}\,;\,\infty\right) \\ &= \left(0.6555841 - \sqrt{0.05263158 + 0.03225806 + 0.04761905 + 0.01515152}u_{0.95}\,;\,\infty\right) \\ &= \left(0.6555841 - \sqrt{0.1476602} \times 1.644854\,;\,\infty\right) \\ &= (0.6555841 - 0.3842658 \times 1.644854\,;\,\infty) \\ &= (0.02352296\,;\,\infty) \end{aligned}$$

```
605 dh <- lnOR - sqrt(1 / n11 + 1 / n12 + 1 / n21 + 1 / n22) * qnorm(1 - alpha) # 0.02352308
```

- **Závěr testování**
  Protože $\ln o\rho_0 = 0$ nenáleží do Waldova 95% empirického levostranného intervalu spolehlivosti, tj. $\ln o\rho_0 = 0 \notin IS$, H_0 zamítáme na hladině významnosti $\alpha = 0.05$.

5. **Testování p-hodnotou**

   - p**-hodnota**

$$p\text{-hodnota} = \Pr(Z_W > z_W) = 1 - \Pr(Z_W \leq z_W) = 1 - \Pr(Z_W \leq 1.706069) = 0.04399759 \doteq 0.0440$$

```
606 p.hodnota <- 1 - pnorm(zw) # 0.04399759
```

   - **Závěr testování**
     Protože p-hodnota $= 0.0440$ 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 H_0. Ke stejnému závěru jsme dospěli také v příkladu 10.18. Šance na výskyt nízkého zakončení dlaňových linií u žen mechské populace je statisticky významně větší než u žen z populace Rajbanshi. Konkrétně potom, šance na výskyt nízkého zakončení dlaňových linií u žen mechské populace je 1.9263 krát větší než u žen z populace Rajbanshi.

★

Příklad 10.23. Test o logaritmu podílu šancí $\ln OR$ (levostranná alternativa)
Mějme k dispozici údaje o frekvenci výskytu epigenetického znaku *sutura metopica* (binomické proměnná) na lebkách jedinců z Anatolské populace (viz tabulka 8; (Eroglu, 2008)).

Tabulka 8: Výskyt epigenetického znaku *sutura metopica* u jedinců Anatolské populace

	sutura metopica	$\sum$
muži	26	334
ženy	15	153

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

Řešení příkladu 10.27
Zadání příkladu je analogické zadání příkladu 10.19. Sledovanou událostí je zde výskyt epigenetického znaku *sutura metopika*, a to jednak u mužů Anatolské populace (okolnost I) a jednak u žen Anatolské populace (okolnost II). Z tabulky 8 ihned vidíme, že u 26 mužů celkového počtu 334 mužů byl zaznamenán výskyt epigenetického znaku *sutura metopika*, u zbylých 308 mužů nikoliv. Dále vidíme, že u 15 žen z celkového počtu 153 žen byl rovněž zaznamenán výskyt epigenetického znaku *sutura metopika* a u zbylých 138 nikoliv. Tabulku uvedenou v zadání upravíme do tvaru čtyřpolní kontingenční tabulky 9.

Tabulka 9: Kontingenční tabulka absolutních četností výskytu epigenetického znaku *sutura metopica* u mužů a žen z Anatolské populace

Pohlaví	Epigenetický znak *sutura metopica* ano	ne	$\sum$
muži (okolnost I)	26	308	334
ženy (okolnost II)	15	138	153
$\sum$	41	446	487

Jelikož nemáme k dispozici datový soubor, vytvoříme se čtyřpolní kontingenční tabulku příkazem data.frame().

```
(tab <- data.frame(ano = c(26, 15), ne = c(308, 138),
                   row.names = c('muzi', 'zeny')))
```

```
     ano  ne
muzi  26 308
zeny  15 138
```

Nyní vypočítáme hodnotu podílu šancí OR a logaritmu podílu šancí $\ln OR$. Z kontingenční tabulky 9 známe konkrétní hodnoty simultánních absolutních četností, zde konkrétně $n_{11} = 26$, $n_{12} = 308$, $n_{21} = 15$ a $n_{22} = 334$. Dosazením hodnot do vztahu 10.9 získáme hodnotu podílu šancí.

$$OR = \frac{n_{11}n_{22}}{n_{12}n_{21}} = \frac{26 \times 138}{308 \times 15} = \frac{3588}{4620} = 0.7766234 \doteq 0.7766.$$

Interpretace podílu šancí: Hodnota podílu šancí $OR = 0.7766$. Šance na výskyt epigenetického znaku *sutura metopica* u mužů Anatolské populace je 0.7766 krát větší než u žen Anatolské populace.

Poznámka: Výše uvedená interpretace *Šance na výskyt epigenetického znaku sutura metopica u mužů Anatolské populace je 0.7766 krát větší než u žen Anatolské populace.* vlastně znamená, že šance na výskyt epigenetického znaku *sutura metopica* u mužů je menší než u žen. Uvědomme si, že číslo x, které je 0.7766 krát větší než číslo y, je ve skutečnosti menší než y. Lépe představitelné je to v situaci, kdy tvrdíme, že číslo x je 0.5 krát větší než číslo y, což v překladu znamená, že x nabývá poloviční hodnoty čísla y.
Aplikováním přirozeného logaritmu na podíl šancí OR získáme hodnotu logaritmu podílu šancí $\ln OR$.

$$\ln OR = \ln(OR) = \ln(0.7766234) = -0.2527997 \doteq -0.2528.$$

```r
n11 <- tab[1, 1]
n12 <- tab[1, 2]
n21 <- tab[2, 1]
n22 <- tab[2, 2]
OR <- (n11 * n22) / (n12 * n21) # 0.7766234
```

```r
lnOR <- log(OR) #  -0.2527998
```

Hodnota logaritmu podílu šancí $\ln OR = -0.2528$.

Nyní se zaměříme na testování nulové hypotézy. Protože v zadání příkladu není zmínka o znění nulové hypotézy, pouze o záměru zjistit, zda je šance na výskyt znaku *sutura metopica* u mužů nižší než u žen, bude otázka ze zadání slovní formulací alternativní hypotézy. Matematický ekvivalent k slovní formulaci alternativní hypotézy odpovídá výroku, že logaritmus podílu šancí je menší než nula. Slovní a matematickou formulaci nulové hypotézy definujeme jako doplněk k alternativní hypotéze.

1. **Stanovení hypotéz**

   - **slovní formulace** nulové a alternativní hypotézy
     H_0 : *Šance na výskyt epigenetického znaku sutura metopica u mužů Anatolské populace je vyšší nebo rovna šanci na výsky epigenetického znaku sutura metopica u žen Anatolské populace.*
     H_1 : *Šance na výskyt epigenetického znaku sutura metopica u mužů Anatolské populace je nižší než šance na výsky epigenetického znaku sutura metopica u žen Anatolské populace.*
   - **matematická formulace** nulové a alternativní hypotézy
     $H_0 : \ln o\rho \geq \ln o\rho_0$, kde $\ln o\rho_0 = 0$
     $H_1 : \ln o\rho < \ln o\rho_0$, kde $\ln o\rho_0 = 0$
     (levostranná alternativa)

2. **Volba hladiny významnosti**

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

3. **Testování kritickým oborem**

   - **Testovací statistika**

$$\begin{aligned} Z_W &= \frac{\ln OR - \ln o\rho_0}{\sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}}} \\ &= \frac{-0.2527997 - 0}{\sqrt{\frac{1}{26} + \frac{1}{308} + \frac{1}{15} + \frac{1}{138}}} \\ &= \frac{-0.2527997}{\sqrt{0.03846154 + 0.003246753 + 0.06666667 + 0.007246377}} \\ &= \frac{-0.2527997}{\sqrt{0.1156213}} \\ &= \frac{-0.2527997}{0.3400313} = -0.7434601 \doteq -0.7435 \end{aligned}$$

```r
alpha <- 0.01
ln.orho0 <- 0
zw <- (lnOR - ln.orho0) / sqrt(1 / n11 + 1 / n12 + 1 / n21 + 1 / n22) # -0.7434601
```

- **Kritický obor**

$$\begin{aligned} W &= (-\infty\,;\, u_\alpha) \\ &= (-\infty\,;\, u_{0.01}) \\ &= (-\infty\,;\, -2.326348) \end{aligned}$$

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

- **Závěr testování**
  Protože realizace testovací statistiky $z_W = -0.7435$ nenáleží do kritického oboru, tj. $z_W \notin W$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

4. **Testování intervalem spolehlivosti**
   Proti levostranné alternativní hypotézy postavíme pravostranný interval spolehlivosti.

   - **Interval spolehlivosti**

$$\begin{aligned} (d, h) &= \left(-\infty\,;\, \ln OR - \sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}}\, u_\alpha\right) \\ &= \left(-\infty\,;\, -0.2527998 - \sqrt{\frac{1}{285} + \frac{1}{310} + \frac{1}{241} + \frac{1}{277}}\, u_{0.01/2}\right) \\ &= \left(-\infty\,;\, -0.2527998 - \sqrt{0.03846154 + 0.003246753 + 0.06666667 + 0.007246377}\, u_{0.005}\right) \\ &= \left(-\infty\,;\, -0.2527998 - \sqrt{0.1156213} \times (-2.326348)\right) \qquad = (-\infty\,;\, -0.2527998 - \end{aligned}$$

```
622 hh <- lnOR - sqrt(1 / n11 + 1 / n12 + 1 / n21 + 1 / n22) * qnorm(alpha) # 0.5382315
```

   - **Závěr testování**
     Protože $\ln o\rho_0 = 0$ náleží do Waldova 99% empirického pravostranného intervalu spolehlivosti, tj. $\ln o\rho_0 = 0 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

5. **Testování p-hodnotou**

   - **p-hodnota**

$$p\text{-hodnota} = \Pr(Z_W \leq z_W) = \Pr(Z_W \leq -0.7434601) = 0.2286016 \doteq 0.2286$$

```
623 p.hodnota <- pnorm(zw) # 0.9706925
```

   - **Závěr testování**
     Protože p-hodnota $= 0.2286$ je větší než $\alpha = 0.01$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

6. **Interpretace výsledků:** Šance na výskyt epigenetického znaku *sutura metopica* u mužů Anatolské populace je 0.7766 krát větší než u žen Anatolské populace. Na základě všech tří způsobů testování nezamítáme hypotézu H_0. Šance na výskytu epigenetického znaku *sutura metopica* u mužů Anatolské populace není statisticky významně nižší než u žen Anatolské populace. Ke stejnému závěru jsme dospěli také v příkladu 10.19.

★

Příklad 10.24. Testování nezávislosti pomocí testu o logaritmu podílu šancí $\ln OR$
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 o nezávislosti pohlaví a počtu starších sourozenců.

Řešení příkladu 10.24
Zadání příkladu je analogické zadání příkladů 10.17 a 10.21. Hypotéza uvedená v zadání byla akorát modifikovaná tak, aby odpovídala testu o nezávislosti dvou náhodných veličin. Náhodná veličina X, popisující pohlaví novorozence, pochází z alternativního rozdělení s parametrem p_1, který popisuje pravděpodobnost narození chlapce, tj. $X \sim \mathrm{Alt}(p_1)$. Náhodná veličina Y, popisující počet starších sourozenců, pochází z alternativního rozdělení s parametrem p_2, který popisuje pravděpodobnost, že novorozenec nemá žádného staršího sourozence, tj. $Y \sim \mathrm{Alt}(p_2)$. Zadání příkladu ukazuje na test o nezávislosti dvou náhodných veličin X a Y, který je analogický testu o nulové hodnotě logaritmu podílu šancí $\ln OR = 0$.

V příkladu 10.17 jsme si již udělali rozbor zadání, v rámci kterého jsme zjistili, že z celkového počtu 595 novorozenců s žádným starším sourozencem je 310 novorozenců mužského pohlaví a zbylých 285 novorozenců je ženského pohlaví. Dále jsme zjistili, že z celkového počtu 518 novorozenců s jedním starším sourozencem je 277 novorozenců mužského pohlaví a zbylých 241 novorozenců je ženského pohlaví. Zjištěné údaje můžeme poskládat do čtyřpolní kontingenční tabulky 10 upravené pro test o nezávislosti.

Tabulka 10: Kontingenční tabulka absolutních četností narození chlapce a dívky s žádným nebo jedním starším sourozencem

Počet starších sourozenců	Pohlaví novorozence		$\sum$
	muž	žena	
žádný	310	285	595
jeden	277	241	518
$\sum$	587	526	1113

Kontingenční tabulku vytvoříme elegantně pomocí softwaru R analogicky jako v příkladu 10.21.

```
data <- read.delim('00-Data/14-two-samples-probabilities-sexratio.txt')
data <- na.omit(data)
data$sex <- factor(data$sex, levels = c('m', 'f'))
tab <- table(data$o.sib.N, data$sex)
tab <- as.data.frame.matrix(tab)
names(tab) <- c('chlapec', 'devce')
row.names(tab) <- c('zadny', 'jeden')
tab
```

```
      chlapec devce
zadny     310   285
jeden     277   241
```

Protože zadání příkladu vede na test o nezávislosti dvou náhodných veličin z alternativních rozdělení (znaky nominálního typu), vyjádříme si míru závislosti mezi těmito náhodnými veličinami Cramérovým koeficientem. Hodnotu Cramérova koeficientu vypočítáme pomocí funkce cramersV() z knihovny lsr.

```
lsr::cramersV(tab) # 0.02579428
```

Interpretace Cramérova koeficientu: Hodnota Cramérova koeficientu $V = 0.02579$. Mezi pohlavím novorozence a počtem starších sourozenců existuje zanedbatelný stupeň závislosti.
Nyní si vypočítáme hodnotu podílu šancí OR a logaritmu podílu šancí.
Hodnota podílu šancí $OR = 1.056686$, hodnota logaritmu podílu šancí $\ln OR = 0.05514$.

```r
n11 <- tab[1, 1]
n12 <- tab[1, 2]
n21 <- tab[2, 1]
n22 <- tab[2, 2]
OR <- (n11 * n22) / (n12 * n21) # 1.056686
lnOR <- log(OR) # 0.05513746
```

V tomto okamžiku již můžeme přistoupit k procesu testování nulové hypotézy. V zadání příkladu máme uvedeno přesné znění nulové hypotézy. Zbývá tedy stanovit alternativní hypotézu tak, aby byla doplňkem k hypotéze nulové. K vytvoření správného matematického zápisu nulové hypotézy si stačí uvědomit, že hypotéze o nezávislosti dvou náhodných veličin X a Y odpovídá nulová hodnota logaritmu podílu šancí $\ln o\rho$, tj. $\ln o\rho = 0$.

1. **Stanovení hypotéz**

   - **slovní formulace** nulové a alternativní hypotézy
     H_0 : *Pohlaví novorzence a počet starších sourozenců jsou stochasticky nezávislé.*
     H_1 : *Pohlaví novorzence a počet starších sourozenců nejsou stochasticky nezávislé. (Tj. jsou stochasicky závislé.)*
   - **matematická formulace** nulové a alternativní hypotézy
     $H_0 : \ln o\rho = \ln o\rho_0$, kde $\ln o\rho_0 = 0$
     $H_1 : \ln o\rho \neq \ln o\rho_0$, kde $\ln o\rho_0 = 0$ (oboustranná alternativa)

2. **Volba hladiny významnosti**

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

3. **Testování kritickým oborem**

   - **Testovací statistika**

$$\begin{aligned} Z_W &= \frac{\ln OR - \ln o\rho_0}{\sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}}} \\ &= \frac{0.05513746 - 0}{\sqrt{\frac{1}{285} + \frac{1}{310} + \frac{1}{241} + \frac{1}{277}}} \\ &= 0.4579854 \doteq 0.4580 \end{aligned}$$

```r
alpha <- 0.01
ln.orho0 <- 0
zw <- (lnOR - ln.orho0) / sqrt(1 / n11 + 1 / n12 + 1 / n21 + 1 / n22) # 0.4579854
```

   - **Kritický obor**

$$\begin{aligned} W &= \left(-\infty\,;\, u_{\alpha/2}\right\rangle \cup \left\langle u_{1-\alpha/2}\,;\, \infty\right) \\ &= \left(-\infty\,;\, u_{0.01/2}\right\rangle \cup \left\langle u_{1-0.01/2}\,;\, \infty\right) \\ &= \left(-\infty\,;\, -2.5758\right\rangle \cup \left\langle 2.575829\,;\, \infty\right) \end{aligned}$$

```r
qnorm(alpha / 2) # -2.575829
qnorm(1 - alpha / 2) # 2.575829
```

- **Závěr testování**
  Protože realizace testovací statistiky $z_W = 0.4580$ nenáleží do kritického oboru, tj. $z_W \notin W$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

4. **Testování intervalem spolehlivosti**

   - **Interval spolehlivosti**

$$\begin{aligned}(d,h) &= \left(\ln OR - \sqrt{\frac{1}{n_{11}}+\frac{1}{n_{12}}+\frac{1}{n_{21}}+\frac{1}{n_{22}}}u_{1-\alpha/2}\,;\,\ln OR - \sqrt{\frac{1}{n_{11}}+\frac{1}{n_{12}}+\frac{1}{n_{21}}+\frac{1}{n_{22}}}u_{\alpha/2}\right)\\ &= \left(0.05513746 - \sqrt{\frac{1}{285}+\frac{1}{310}+\frac{1}{241}+\frac{1}{277}}u_{1-0.01/2}\,;\,0.05513746 - \sqrt{\frac{1}{285}+\frac{1}{310}+\frac{1}{241}+\frac{1}{277}}u_{0.01/2}\right)\\ &= (-0.2549699\,;\,0.3652449)\end{aligned}$$

```
647 dh <- lnOR - sqrt(1 / n11 + 1 / n12 + 1 / n21 + 1 / n22) * qnorm(1 - alpha / 2) #
        -0.25497
648 hh <- lnOR - sqrt(1 / n11 + 1 / n12 + 1 / n21 + 1 / n22) * qnorm(alpha / 2) # 0.3652449
```

   - **Závěr testování**
     Protože $\ln o\rho_0 = 0$ náleží do Waldova 99% empirického oboustranného intervalu spolehlivosti, tj. $\ln o\rho_0 = 0 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

5. **Testování p-hodnotou**

   - **p-hodnota**

$$\begin{aligned}p\text{-hodnota} &= 2\min\{\Pr(Z_W \le z_W), \Pr(Z_W > z_W)\}\\ &= 2\min\{\Pr(Z_W \le 0.4579854), 1 - \Pr(Z_W \le 0.4579854)\}\\ &= 0.646963 \doteq 0.6470\end{aligned}$$

```
649 p.hodnota <- 2 * min(pnorm(zw), 1 - pnorm(zw)) # 0.6469629
```

   - **Závěr testování**
     Protože p-hodnota $= 0.6470$ je větší než $\alpha = 0.01$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

6. **Interpretace výsledků:** Mezi pohlavím novorozence a počtem starších sourozenců existuje zanedbatelný stupeň závislosti. Na základě všech tří způsobů testování nezamítáme hypotézu H_0. Pohlaví novorozence a počet starších sourozenců jsou stochasticky nezávislé náhodné veličiny.

★

10.7 Fisherův exaktní test

Jako poslední z této kapitoly si ukážeme *Fisherův exaktní test* známý také jako *Fisherův faktoriálový test*, pomocí kterého lze opět porovnávat dva náhodné výběry z alternativních rozdělení. Na rozdíl od testu o podílu šancí, který je pouze asymptotický, Fisherův faktoriálový test je testem exaktním. Jeho výsledky jsou tedy spolehlivě správné i pro náhodné výběry malých rozsahů. Analogicky jako v předchozí sekci nám Fisherův exaktní test umožňuje porovnávat dva náhodné výběry z alternativních rozdělení, o kterých předpokládáme, že popisují výskyt stejné sledované události, ovšem za dvou různých okolností (I a II). První náhodný výběr popisuje výskyt sledované události za okolnosti I, druhý náhodný výběr popisuje výskyt tentýž sledované události ovšem za okolnosti II.

Počet nastání a nenastání sledované události za nastání okolnosti I nebo okolnosti II můžeme přehledně zavést do čtyřpolní kontingenční tabulky

Okolnost	Sledovaná událost		$\sum$
	nastala	nenastala	
I	n_{11}	n_{12}	$n_{11}+n_{12}$
II	n_{21}	n_{22}	$n_{21}+n_{22}$
$\sum$	$n_{11}+n_{21}$	$n_{12}+n_{22}$	N

a definovat podíl šancí OR nastání sledované události za okolnosti I ku nastání sledované události za okolnosti II pomocí vztahu

$$OR = \frac{n_{11}/n_{12}}{n_{21}/n_{22}} = \frac{n_{11}n_{22}}{n_{12}n_{21}}. \tag{10.12}$$

Nechť $o\rho$ je teoretická hodnota podílu šancí a $o\rho_0$ je konstanta. Na hladině významnosti α testujeme jednu z následujících tří hypotéz oproti příslušné alternativní hypotéze.

$H_{01} : o\rho = o\rho_0$	oproti	$H_{11} : o\rho \neq o\rho_0$	(oboustranná alt.)
$H_{02} : o\rho \leq o\rho_0$	oproti	$H_{12} : o\rho > o\rho_0$	(pravostranná alt.)
$H_{03} : o\rho \geq o\rho_0$	oproti	$H_{13} : o\rho < o\rho_0$	(levostranná alt.)

Algoritmus Fisherova exaktního testu patří mezi pokročilé statistické metody a proto se jím zde nebudeme detailně zabývat. Bližší informace o algoritmu lze v případě zájmu nalézt v literatuře CITACE. K otestování nulové hypotézy použijeme funkci fisher.test() implementovanou v softwaru . Mezi její výstupy patří hodnota podílu šancí OR, hranice $(1-\alpha)\times 100\%$ intervalu spolehlivosti pro OR a p-hodnota. Testování pomocí Fisherova exaktního testu budeme tedy provádět pouze pomocí intervalu spolehlivosti nebo p-hodnoty.

Poznámka: Hypotézy H_{01}, H_{11}, H_{02}, H_{12}, H_{03} a H_{13} uvedené v této sekci bychom mohli klidně zapsat pomocí logaritmu podílu šancí $\ln OR$. Protože však funkce fisher.test() poskytuje jako výstup podíl šancí OR a $(1-\alpha)\times 100\%$ intervalu spolehlivosti pro podíl šancí, zapisujeme i zde hypotézy pomocí podílu šancí. Z hlediska algoritmu Fisherova testu není podstatné, zda hypotézu definujeme pomocí podílu šancí, nebo logaritmu podílu šancí, protože ani jedna z těchto statistik v algoritmu Fisherova exaktního testu nefiguruje.

Poznámka: Pomocí Fisherova exaktního testu můžeme také testovat hypotézu H_{01}: *Náhodné veličiny X a Y jsou o stochasticky nezávislé.* Tato hypotéza odpovídá znění nulové hypotézy $H_{01} : o\rho = 1$. Znění alternativní hypotézy H_{11} : *Náhodné veličiny X a Y nejsou stochasticky nezávislé.* potom odpovídá znění alternativní hypotézy H_{01} : $o\rho \neq 1$. Poznamenejme, že jednostranné hypotézy v tomto případě testovat nemůžeme, protože X a Y jako znaky nominálního typu neumožňují diagnostikovat přímou nebo nepřímou závislost (viz. kapitola **??**).

Příklad 10.25. Fisherův exaktní test (oboustranná alternativa)
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 šance na narození chlapce je stejná u novorozenců s žádným sourozencem a u novorozenců s jedním sourozencem.

Řešení příkladu 10.25
Zadání příkladu je analogické zadání příkladů 10.17, 10.21 a 10.24. Sledovanou událostí je v tomto případě opět narození novorozence mužského pohlaví, přičemž výskyt této události nás zajímá jednak za situace, kdy novorozenec nemá žádného sourozence (okolnost I) a jednak za situace, kdy novorozenec má jednoho staršího sourozence (okolnost II). Pro připomenutí uvádíme kontingenční tabulku absolutních četností (viz tabulka 11).

Tabulka 11: Kontingenční tabulka absolutních četností narození chlapce s žádným nebo jedním starším sourozencem

Počet starších sourozenců	Narození chlapce		$\sum$
	ano	ne	
žádný (okolnost I)	310	285	595
jeden (okolnost II)	277	241	518
$\sum$	587	526	1113

Kontingenční tabulku si vytvoříme také pomocí softwaru , analogicky jako v příkladu 10.21.

```
650 data <- read.delim('00-Data/14-two-samples-probabilities-sexratio.txt')
651 data <- na.omit(data)
652 tab <- table(data$o.sib.N, data$sex)
653 tab <- as.data.frame.matrix(tab)
654 names(tab) <- c('chlapec', 'devce')
655 row.names(tab) <- c('zadny', 'jeden')
656 tab
```

```
      chlapec devce                                                                    657
zadny     285   310                                                                    658
jeden     241   277                                                                    659
```

Nyní již můžeme přistoupit k procesu testování nulové hypotézy. V zadání příkladu máme uvedeno přesné znění nulové hypotézy. Zbývá tedy stanovit alternativní hypotézu tak, aby byla doplňkem k hypotéze nulové. K vytvoření správného matematického zápisu nulové hypotézy si stačí uvědomit, že shodě šancí na výskyt události za okolnosti I a okolnosti II odpovídá hodnota podílu šancí $o\rho = 1$.

1. **Stanovení hypotéz**

   - **slovní formulace** nulové a alternativní hypotézy
     H_0 : *Šance na narození chlapce u novorozenců s žádným sourozencem a u novorozenců s jedním sourozencem jsou stejné.*
     H_1 : *Šance na narození chlapce u novorozenců s žádným sourozencem a u novorozenců s jedním sourozencem nejsou stejné.*
   - **matematická formulace** nulové a alternativní hypotézy
     $H_0 : o\rho = o\rho_0$, kde $o\rho_0 = 1$
     $H_1 : o\rho \neq o\rho_0$, kde $o\rho_0 = 1$ (oboustranná alternativa)

2. **Volba hladiny významnosti**

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

3. **Syntaxe a výstupy funkce** fisher.test()
   Testování nulové hypotézy provedeme pomocí Fisherova exaktního testu implementovaného v softwaru ve funkci fisher.test(). Testování provedeme pomocí intervalu spolehlivosti a p-hodnoty, jelikož k těmto dvěma

způsobům testování poskytuje funkce fisher.test() výsledky. Prvním vstupním argumentem funkce bude kontingenční tabulka absolutních četností tab. Následně zvolíme oboustranný typ alternativní hypotézy nastavením argumentu alternative = 'two.sided' a hladinu významnosti $\alpha = 0.01$ nastavením hodnoty koeficientu spolehlivosti $(1 - \alpha)$ pomocí argumentu conf.level = 0.99.

```
660 fisher.test(tab, alternative = 'two.sided', conf.level = 0.99)
```

```
661
662         Fisher's Exact Test for Count Data
663
664 data:  tab
665 p-value = 0.6737
666 alternative hypothesis: true odds ratio is not equal to 1
667 99 percent confidence interval:
668  0.7695517 1.4511781
669 sample estimates:
670 odds ratio
671   1.056607
```

Výstupem funkce fisher.test() je souhrnná tabulka výsledků poskytující hodnotu podílu šancí OR 1.056607, hranice 99% Waldova oboutstranného empirického intervalu spolehlivosti 0.7695517 a 1.4511781 a p-hodnotu 0.6737. Výsledky konkrétních statistik můžeme z výstupu funkce fisher.test() získat prostřednictvím dolarové syntaxe $.

Hodnotu podílu šancí OR získáme z příkazu fisher.test(tab, ...) zavoláním výstupu nazvaného estimate.

```
672 OR <- fisher.test(tab, alternative = 'two.sided', conf.level = 0.99)$estimate # 1.056607
```

Interpretace podílu šancí: Hodnota podílu šancí $OR = 1.05661$. Šance, že novorozenec bude mužského pohlaví, je 1.05661 krát větší u novorozence, který nemá žádného staršího sourozence, než u novorozence, který má jednoho staršího sourozence.

Poznámka: Pro srovnání připomeňme, že hodnota podílu šancí OR vypočítaná v příkladu 10.21 nabývala hodnoty 1.056686. Podíly šancí vypočítané exaktním výpočtem a funkcí fisher.test() se liší na pátém desetinném místě.

4. **Testování kritickým oborem**
   Testování kritickým oborem nelze provést, protože funkce fisher.test() neposkytuje potřebné výsledky pro tento způsob testování.

5. **Testování intervalem spolehlivosti**

   - **Interval spolehlivosti**
     Hranice 99% Waldova empirického oboustranného IS získáme z příkazu fisher.test(tab, ...) zavoláním výstupu nazvaného conf.int. Tento výstup je vektor dvou hodnot. Dolní hranici odpovídá první položka vektoru, horní hranici odpovídá druhá položka vektoru.

```
673 dh <- fisher.test(tab, alternative = 'two.sided', conf.level = 0.99)$conf.int[1] #
        0.7695517
674 hh <- fisher.test(tab, alternative = 'two.sided', conf.level = 0.99)$conf.int[2] #
        1.4511781
```

$$(d, h) = (0.7695517\,;\, 1.4511781)$$

- **Závěr testování**
  Protože $o\rho_0 = 1$ náleží do Waldova 99% empirického oboustranného intervalu spolehlivosti, tj. $o\rho_0 = 1 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

6. **Testování p-hodnotou**

   - **p-hodnota**
     Výslednou p-hodnotu získáme z příkazu fisher.test(tab, ...) zavoláním výstupu nazvaného p.value.

```
675 p.hodnota <- fisher.test(tab, alternative = 'two.sided', conf.level = 0.99)$p.value
```

$$p\text{-hodnota} = 0.6736532 \doteq 0.6737$$

   - **Závěr testování**
     Protože p-hodnota $= 0.6737$ je větší než $\alpha = 0.01$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

7. **Interpretace výsledků:** Na základě testování pomocí IS a p-hodnoty nezamítáme hypotézu H_0. Hodnota podílu šancí OR vyšla 1.05661. Šance, že novorozenec bude mužského pohlaví, je 1.05661 krát větší u novorozence, který nemá žádného staršího sourozence, než u novorozence, který má jednoho staršího sourozence. Tento rozdíl v šancích však nevyšel jako statisticky významný. Ke stejnému závěru testování jsme dospěli taktéž v řešení příkladu 10.17 a v řešení příkladu 10.21.

★

Příklad 10.26. Fisherův exaktní test (pravostranná alternativa)
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 mechské populace a 87 žen z populace Rajbanshi (viz sekce **??**. Na hladině významnosti $\alpha = 0.05$ zjistěte, zda je šance na výskyt nízkého zakončení tří dlaňových linií na levé straně u žen z mechské populace vyšší než šance na výskyt nízkého zakončení tří dlaňových linií na levé straně u žen z populace Rajbanshi.

Řešení příkladu 10.26
Zadání příkladu je analogické zadání příkladů 10.18 a 10.22. Sledovanou událostí je v tomto případě nízké zakončení tří dlaňových linií na levé straně, přičemž výskyt nízkého zakončení dlaňových linií sledujeme jednak u žen z mechské populace (okolnost I) a jednak u žen z populace Rajbanshi (okolonost II). Pro připomenutí zjištěné údaje o výskytu nízkého zakončení palmárních linií u žen mechské populace a žen z populace Rajbanshi shrneme ve čtyřpolní kontingenční tabulce 12.

Tabulka 12: Kontingenční tabulka absolutních četností výskytu nízkého zakončení dlaňových linií u mechské ženské populace a u žen z populace Rajbanshi

Populace žen	Nízké zakončení dlaňových linií		$\sum$
	ano	ne	
mechská (okolnost I)	19	31	50
Rajbanshi (okolnost II)	21	66	87
$\sum$	40	97	137

Čtyřpolní kontingenční tabulku vytvoříme, analogicky jako v příkladu 10.22 pomocí příkazu data.frame()

```
(tab <- data.frame(ano = c(19, 21), ne = c(31, 66),
                   row.names = c('mechska', 'Rajbanshi')))
```

```
          ano ne
mechska    19 31
Rajbanshi  21 66
```

Naším úkolem je zjistit, zda je šance na výskyt nízkého zakončení tří dlaňových linií na levé straně u žen z mechské populace vyšší než u žen z populace Rajbanshi. Stejně jako v příkladech 10.18 a 10.26 je i zde uvedené tvrzení zněním alternativní hypotézy. Matematický ekvivalent ke slovní formulaci alternativní hypotézy potom odpovídá výroku, že podíl šancí je větší než jedna. Zbývá dodefinovat slovní a matematickou formulaci nulové hypotézy tak, aby byla doplňkem hypotézy alternativní.

1. **Stanovení hypotéz**
   - **slovní formulace** nulové a alternativní hypotézy
     H_0 : *Šance na výskyt nízkého zakončení tří dlaňových linií na levé straně u žen z mechské populace je nižší nebo rovna šanci na výskyt nízkého zakončení tří dlaňových linií na levé straně u žen z populace Rajbanshi.*
     H_1 : *Šance na výskyt nízkého zakončení tří dlaňových linií na levé straně u žen z mechské populace je vyšší než šance na výskyt nízkého zakončení tří dlaňových linií na levé straně u žen z populace Rajbanshi.*
   - **matematická formulace** nulové a alternativní hypotézy
     $H_0 : o\rho \leq o\rho_0$, kde $o\rho_0 = 1$
     $H_1 : o\rho > o\rho_0$, kde $o\rho_0 = 1$ (pravostranná alternativa)
2. **Volba hladiny významnosti**
   - Hladinu významnosti volíme v souladu se zadáním $\alpha = 0.05$.
3. **Syntaxe a výstupy funkce** fisher.test()
   Testování nulové hypotézy provedeme pomocí Fisherova exaktního testu implementovaného v softwaru ve funkci fisher.test(). Testování opět provedeme pouze pomocí intervalu spolehlivosti a p-hodnoty. Prvním

vstupním argumentem funkce bude kontingenční tabulka absolutních četností tab. Následně zvolíme pravostranný typ alternativní hypotézy nastavením argumentu alternative = 'greater' a hladinu významnosti $\alpha = 0.05$ nastavením hodnoty koeficientu spolehlivosti prostřednictvím argumentu conf.level = 0.95.

```
fisher.test(tab, alternative = 'greater', conf.level = 0.95)
```

```
        Fisher's Exact Test for Count Data

data:  tab
p-value = 0.06477
alternative hypothesis: true odds ratio is greater than 1
95 percent confidence interval:
 0.9517978       Inf
sample estimates:
odds ratio
  1.916682
```

Výstupem funkce fisher.test() je souhrnná tabulka výsledků poskytující hodnotu podílu šancí OR 1.916682, hranice 95% Waldova levostranného empirického intervalu spolehlivosti 0.9517978 a Inf a p-hodnotu 0.06477. Výsledky konkrétních statistik získáme z výstupu funkce fisher.test() prostřednictvím dolarové syntaxe $.

Hodnotu podílu šancí OR získáme z příkazu fisher.test(tab, ...) zavoláním výstupu nazvaného estimate.

```
OR <- fisher.test(tab, alternative = 'greater', conf.level = 0.95)$estimate # 1.916682
```

Interpretace podílu šancí: Hodnota podílu šancí $OR = 1.9167$. Šance na výskyt nízkého zakončení dlaňových linií u žen mechské populace je 1.9167 krát větší než u žen z populace Rajbanshi.

Poznámka: Pro srovnání připomeňme, že hodnota podílu šancí OR vypočítaná v příkladu **??** nabývala hodnoty 1.926267. Podíly šancí vypočítané exaktním výpočtem a funkcí fisher.test() se liší na druhém desetinném místě.

4. **Testování kritickým oborem**
   Testování kritickým oborem nelze provést, protože funkce fisher.test() neposkytuje potřebné výsledky pro tento způsob testování.

5. **Testování intervalem spolehlivosti**

   - **Interval spolehlivosti**
     Hranice 95% Waldova empirického oboustranného IS získáme z příkazu fisher.test(tab, ...) zavoláním výstupu nazvaného conf.int.

```
dh <- fisher.test(tab, alternative = 'greater', conf.level = 0.95)$conf.int[1] # 0.9517978
hh <- fisher.test(tab, alternative = 'greater', conf.level = 0.95)$conf.int[2] # Inf
```

$$(d, h) = (0.9517978\,;\, \infty)$$

   - **Závěr testování**
     Protože $o\rho_0 = 1$ náleží do Waldova 95% empirického levostranného intervalu spolehlivosti, tj. $o\rho_0 = 1 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

6. **Testování p-hodnotou**

```
696 p.hodnota <- fisher.test(tab, alternative = 'greater', conf.level = 0.95)$p.value
```

- **p-hodnota**
  Výslednou p-hodnotu získáme z příkazu fisher.test(tab, ...) zavoláním výstupu nazvaného p.value.

$$p\text{-hodnota} = 0.06477234 \doteq 0.06477$$

- **Závěr testování**
  Protože p-hodnota $= 0.06477234$ je větší než $\alpha = 0.05$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

7. **Interpretace výsledků:** Na základě testování pomocí IS a p-hodnoty nezamítáme hypotézu H_0. Hodnota podílu šancí OR vyšla 1.9167. Šance na výskyt nízkého zakončení palmárních linií u žen mechské populace je 1.9167 krát větší než u žen z populace Rajbanshi. Tento rozdíl v šancích však nevyšel jako statisticky významný.

Poznámka: Výsledek testování pomocí Fisherova exaktního testu se liší od výsledku získaného v rámci příkladu 10.18 pomocí testu o rozdílu pravděpodobností $p_1 - p_2$ i od výsledku získaného v rámci příkladu 10.22 pomocí testu o logaritmu podílu šancí $\ln OR$. Oba testy na rozdíl od Fisherova testu vyhodnotily rozdíl mezi šancemi jako statisticky významný. S obdobnou situací se můžeme kdykoli setkat v praxi při analýze našich vlastních dat. Všimněme si, že p-hodnota testu o rozdílu pravděpodobností $p_1 - p_2$ (p-hodnota $= 0.04658$), stejně jako p-hodnota testu o logaritmu podílu šancí (p-hodnota $= 0.0440$), je jen o málo nižší než hladina významnosti α. Naopak p-hodnota Fisherova exaktního testu (p-hodnota $= 0.06477$) je jen o málo vyšší než hladina významnosti α. Jde tedy o hraniční situaci, při které se dva asymptotické testy přiklonily k zamítnutí H_0, zatímco exaktní test se kloní k závěru H_0 nezamítat. Poznamenejme, že p-hodnota Fisherova exaktního testu bývá zpravidla vyšší než p-hodnota asymptotických testů. Je to důsledek diskrétního charakteru algoritmu Fisherova testu, který je současně zárukou, že test je spolehlivý i v situaci, kdy máme malé rozsahy náhodných výběrů.

Situaci je tedy třeba posoudit individuálně. Jelikož rozsahy náhodných výběrů jsou dostatečně velké (o čemž svědčí například splnění podmínky dobré aproximace u testu o rozdílu pravděpodobností $p_1 - p_2$, stejně jako fakt, že celkový počet všech pozorování N je větší než 100) a jelikož počet pozorování v obou skupinách (mechská populace i populace Rajbanshi) je celkem vyrovnaný (50 pozorování a 87 pozorování), kloníme se k závěru testu o rozdílu pravděpodobností a testu o logaritmu podílu šancí a konstatujeme, že šance na výskyt nízkého zakončení palmárních linií u žen mechské populace je statisticky významně větší než u žen z populace Rajbanshi. Pamatujme však na to, že finální rozhodnutí jsme podložili pádnými argumenty. Například při nižším počtu pozorování (např. $N < 50$), nebo při výrazně odlišném počtu pozorování v obou skupinách (např. 15 pozorování ku 122 pozorováním) bychom se museli přiklonit k závěru Fisherova exaktního testu. ★

Příklad 10.27. Fisherův exaktní test (levostranná alternativa)
Mějme k dispozici údaje o frekvenci výskytu epigenetického znaku *sutura metopica* (binomické proměnná) na lebkách jedinců z Anatolské populace (viz tabulka 13; (Eroglu, 2008)).

Tabulka 13: Výskyt epigenetického znaku *sutura metopica* u jedinců Anatolské populace

	sutura metopica	$\sum$
muži	26	334
ženy	15	153

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

Řešení příkladu 10.27
Zadání příkladu je analogické zadání příkladů 10.19 a 10.27. Sledovanou událostí je zde výskyt epigenetického znaku *sutura metopika*, a to jednak u mužů Anatolské populace (okolnost I) a jednak u žen Anatolské populace (okolnost II). Pro připomenutí uvádíme tvar čtyřpolní kontingenční tabulky (viz tabulka 14) absolutních četností, kterou jsme sestavili v rámci příkladu 10.27.

Tabulka 14: Kontingenční tabulka absolutních četností výskytu epigenetického znaku *sutura metopica* u mužů a žen z Anatolské populace

Pohlaví	Epigenetický znak *sutura metopica* ano	ne	$\sum$
muži (okolnost I)	26	308	334
ženy (okolnost II)	15	138	153
$\sum$	41	446	487

Čtyřpolní kontingenční tabulku vytvoříme pomocí softwaru , analogicky jako v příkladu 13.

```
(tab <- data.frame(ano = c(26, 15), ne = c(308, 138),
                   row.names = c('muzi', 'zeny')))
```

```
     ano  ne
muzi  26 308
zeny  15 138
```

Nyní se zaměříme na testování nulové hypotézy. V zadání příkladu je uvedena slovní formulace alternativní hypotézy. Matematický ekvivalent k slovní formulaci alternativní hypotézy odpovídá výroku, že podíl šancí OR je menší než jedna. Slovní a matematickou formulaci nulové hypotézy definujeme jako doplněk k alternativní hypotéze.

1. **Stanovení hypotéz**

   - **slovní formulace** nulové a alternativní hypotézy
     H_0 : *Šance na výskyt epigenetického znaku sutura metopica u mužů Anatolské populace je vyšší nebo rovna šanci na výsky epigenetického znaku sutura metopica u žen Anatolské populace.*
     H_1 : *Šance na výskyt epigenetického znaku sutura metopica u mužů Anatolské populace je nižší než šance na výsky epigenetického znaku sutura metopica u žen Anatolské populace.*
   - **matematická formulace** nulové a alternativní hypotézy
     $H_0 : o\rho \geq o\rho_0$, kde $o\rho_0 = 1$
     $H_1 : o\rho < o\rho_0$, kde $o\rho_0 = 1$
     (levostranná alternativa)

2. **Volba hladiny významnosti**

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

3. **Syntaxe a výstupy funkce** fisher.test()
   Testování nulové hypotézy provedeme pomocí Fisherova exaktního testu implementovaného v softwaru R ve funkci fisher.test(), a to opět pouze pomocí intervalu spolehlivosti a p-hodnoty. Prvním vstupním argumentem funkce bude kontingenční tabulka absolutních četností tab. Následně zvolíme levostranný typ alternativní hypotézy (argument alternative = 'less') a hladinu významnosti $\alpha = 0.01$ (argument conf.level = 0.99).

```
fisher.test(tab, alternative = 'less', conf.level = 0.99)
```

```
        Fisher's Exact Test for Count Data

data:  tab
p-value = 0.2806
alternative hypothesis: true odds ratio is less than 1
99 percent confidence interval:
 0.000000 1.866671
sample estimates:
odds ratio
 0.7770709
```

   Výstupem funkce fisher.test() je souhrnná tabulka výsledků poskytující hodnotu podílu šancí OR 0.7770709, hranice 99% Waldova pravostranného empirického intervalu spolehlivosti 0 a 1.866671 a p-hodnotu 0.2806.

   Hodnotu podílu šancí OR získáme z příkazu fisher.test(tab, ...) zavoláním výstupu nazvaného estimate.

```
OR <- fisher.test(tab, alternative = 'less', conf.level = 0.99)$estimate # 0.7770709
```

   Interpretace podílu šancí: Hodnota podílu šancí $OR = 0.7771$. Šance na výskyt epigenetického znaku *sutura metopica* u mužů Anatolské populace je 0.7771 krát větší než u žen Anatolské populace.

   Poznámka: Pro srovnání připomeňme, že hodnota podílu šancí OR vypočítaná v příkladu 10.27 nabývala hodnoty 0.7766234. Podíly šancí OR vypočítané exaktním výpočtem a funkcí fisher.test() se liší na třetím desetinném místě.

4. **Testování kritickým oborem**
   Testování kritickým oborem nelze provést, protože funkce fisher.test() neposkytuje potřebné výsledky pro tento způsob testování.

5. **Testování intervalem spolehlivosti**

   - **Interval spolehlivosti**
     Hranice 99% Waldova empirického pravostranného IS získáme z příkazu fisher.test(tab, ...) zavoláním výstupu nazvaného conf.int.

```
dh <- fisher.test(tab, alternative = 'less', conf.level = 0.99)$conf.int[1] # 0
hh <- fisher.test(tab, alternative = 'less', conf.level = 0.99)$conf.int[2] # 1.866671
```

$$(d, h) = (0\,;\, 1.866671)$$

   Poznámka: Protože podíl šancí OR nabývá pouze hodnot z intervalu $(0\,;\,\infty)$ bude i dolní hranice Waldova empirického pravostranného intervalu spolehlivosti nabývat hodnoty 0 namísto mínus nekonečna.

   - **Závěr testování**
     Protože $o\rho_0 = 1$ náleží do Waldova 99% empirického pravostranného intervalu spolehlivosti, tj. $o\rho_0 = 1 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

6. **Testování p-hodnotou**

   - **p-hodnota**
     Výslednou p-hodnotu získáme z příkazu fisher.test(tab, ...) zavoláním výstupu nazvaného p.value.

```
717 p.hodnota <- fisher.test(tab, alternative = 'less', conf.level = 0.99)$p.value #
    0.2806414
```

$$p\text{-hodnota} = 0.2806414 \doteq 0.28064$$

   - **Závěr testování**
     Protože p-hodnota $= 0.28064$ je větší než $\alpha = 0.01$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

7. **Interpretace výsledků:** Šance na výskyt epigenetického znaku *sutura metopica* u mužů Anatolské populace je 0.7771 krát větší než u žen Anatolské populace. Na základě testování pomocí IS a p-hodnoty nezamítáme hypotézu H_0. Šance na výskyt epigenetického znaku *sutura metopica* u mužů není statisticky významně menší než u žen Anatolské populace. Ke stejnému závěru jsme dospěli také v příkladech 10.19 a 10.27.

Příklad 10.28. Testování nezávislosti pomocí Fisherova exaktního testu
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 o nezávislosti pohlaví a počtu starších sourozenců.

Řešení příkladu 10.24
Zadání příkladu je analogické zadání příkladů 10.17, 10.21 a 10.25. Hypotéza uvedená v zadání byla akorát modifikovaná tak, aby odpovídala testu o nezávislosti dvou náhodných veličin. Náhodná veličina X, popisující pohlaví novorozence, pochází z alternativního rozdělení s parametrem p_1, který popisuje pravděpodobnost narození chlapce, tj. $X \sim \text{Alt}(p_1)$. Náhodná veličina Y, popisující počet starších sourozenců, pochází z alternativního rozdělení s parametrem p_2, který popisuje pravděpodobnost, že novorozenec nemá žádného staršího sourozence, tj. $Y \sim \text{Alt}(p_2)$. Zadání příkladu ukazuje na test o nezávislosti dvou náhodných veličin X a Y, který je analogický testu o rovnosti podílu šancí a hodnoty jedna, tj. $\ln OR = 1$.

V příkladu 10.24 jsme vytvořili čtyřpolní kontingenční tabulku 15 upravenou pro test o nezávislosti.

Tabulka 15: Kontingenční tabulka absolutních četností narození chlapce a dívky s žádným nebo jedním starším sourozencem

Počet starších sourozenců	Pohlaví novorozence		$\sum$
	muž	žena	
žádný	310	285	595
jeden	277	241	518
$\sum$	587	526	1113

Kontingenční tabulku vytvoříme elegantně pomocí softwaru analogicky jako v příkladu 10.24.

```
718 data <- read.delim('00-Data/14-two-samples-probabilities-sexratio.txt')
719 data <- na.omit(data)
720 data$sex <- factor(data$sex, levels = c('m', 'f'))
721 tab <- table(data$o.sib.N, data$sex)
722 tab <- as.data.frame.matrix(tab)
723 names(tab) <- c('chlapec', 'devce')
724 row.names(tab) <- c('zadny', 'jeden')
725 tab
```

```
      chlapec devce                                                                    726
zadny     310   285                                                                    727
jeden     277   241                                                                    728
```

Protože zadání příkladu vede na test o nezávislosti dvou náhodných veličin z alternativních rozdělení (znaky nominálního typu), vyjádříme si míru závislosti mezi těmito náhodnými veličinami Cramérovým koeficientem. Hodnotu Cramérova koeficientu vypočítáme pomocí funkce cramersV() z knihovny lsr.

```
729 lsr::cramersV(tab) # 0.02579428
```

Interpretace Cramérova koeficientu: Hodnota Cramérova koeficientu $V = 0.02579$. Mezi pohlavím novorozence a počtem starších sourozenců existuje zanedbatelný stupeň závislosti.

Nyní již můžeme přistoupit k procesu testování nulové hypotézy. V zadání příkladu máme uvedeno přesné znění nulové hypotézy. Zbývá tedy stanovit alternativní hypotézu tak, aby byla doplňkem k hypotéze nulové. K vytvoření správného matematického zápisu nulové hypotézy si stačí uvědomit, že hypotéze o nezávislosti dvou náhodných veličin X a Y odpovídá hodnota podílu šancí $o\rho$ rovná jedné, tj. $o\rho = 1$.

1. **Stanovení hypotéz**

- **slovní formulace** nulové a alternativní hypotézy
  H_0 : *Pohlaví novorozence a počet starších sourozenců jsou stochasticky nezávislé.*
  H_1 : *Pohlaví novorozence a počet starších sourozenců nejsou stochasticky nezávislé. (Tj. jsou stochasicky závislé.)*
- **matematická formulace** nulové a alternativní hypotézy
  $H_0 : o\rho = o\rho_0$, kde $o\rho_0 = 1$
  $H_1 : o\rho \neq o\rho_0$, kde $o\rho_0 = 1$ (oboustranná alternativa)

2. **Volba hladiny významnosti**

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

3. **Syntaxe a výstupy funkce** fisher.test()
   Testování nulové hypotézy provedeme pomocí Fisherova exaktního testu implementovaného v softwaru R ve funkci fisher.test(), stejnou syntaxí jakou jsme použili v příkladu **??**.

```
730 fisher.test(tab, alternative = 'two.sided', conf.level = 0.99)
```

```
731
732        Fisher's Exact Test for Count Data
733
734 data:  tab
735 p-value = 0.6737
736 alternative hypothesis: true odds ratio is not equal to 1
737 99 percent confidence interval:
738  0.6890953 1.2994578
739 sample estimates:
740 odds ratio
741  0.9464253
```

Výstupem funkce fisher.test() je souhrnná tabulka výsledků poskytující hodnotu podílu šancí OR 1.056607, hranice 99% Waldova oboutstranného empirického intervalu spolehlivosti 0.7695517 a 1.4511781 a p-hodnotu 0.6737.

4. **Testování kritickým oborem**
   Testování kritickým oborem nelze provést, protože funkce fisher.test() neposkytuje potřebné výsledky pro tento způsob testování.

5. **Testování intervalem spolehlivosti**

   - **Interval spolehlivosti**
     Hranice 99% Waldova empirického oboustranného IS získáme z příkazu fisher.test(tab, ...) zavoláním výstupu nazvaného conf.int. Tento výstup je vektor dvou hodnot. Dolní hranici odpovídá první položka vektoru, horní hranici odpovídá druhá položka vektoru.

```
742 dh <- fisher.test(tab, alternative = 'two.sided', conf.level = 0.99)$conf.int[1] #
        0.7695517
743 hh <- fisher.test(tab, alternative = 'two.sided', conf.level = 0.99)$conf.int[2] #
        1.4511781
```

$$(d, h) = (0.7695517\,;\, 1.4511781)$$

   - **Závěr testování**
     Protože $o\rho_0 = 1$ náleží do Waldova 99% empirického oboustranného intervalu spolehlivosti, tj. $o\rho_0 = 1 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

6. **Testování p-hodnotou**

```
744 p.hodnota <- fisher.test(tab, alternative = 'two.sided', conf.level = 0.99)$p.value
```

- **p-hodnota**
  Výslednou p-hodnotu získáme z příkazu fisher.test(tab, ...) zavoláním výstupu nazvaného p.value.

$$p\text{-hodnota} = 0.6736532 \doteq 0.6737$$

- **Závěr testování**
  Protože p-hodnota $= 0.6737$ je větší než $\alpha = 0.01$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

7. **Interpretace výsledků:** Mezi pohlavím novorozence a počtem starších sourozenců existuje zanedbatelný stupeň závislosti. Na základě testování pomocí IS a p-hodnoty nezamítáme hypotézu H_0. Pohlaví novorozence a počet starších sourozenců jsou stochasticky nezávislé náhodné veličiny. Ke stejnému závěru jsme dospěli také v příkladu 10.24.

★

★