9 Parametrické úlohy o dvou nezávislých výběrech z normálního rozložení a jednom výběru z alternativního rozložení

9.1 Intervaly spolehlivosti pro rozdíl středních hodnot a podíl rozptylů

$X_{11}, \ldots, X_{1n_1}$ je náhodný výběr z rozložení $N(\mu_1, \sigma_1^2)$ a $X_{21}, \ldots, X_{2n_2}$ je na něm nezávislý náhodný výběr z rozložení $N(\mu_2, \sigma_2^2)$, přičemž $n_1 \geq 2$, $n_2 \geq 2$. Označme M_1, M_2 výběrové průměry, S_1^2, S_2^2 výběrové rozptyly a $S_*^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$ vážený průměr výběrových rozptylů. Při konstrukci intervalů spolehlivosti pro parametrické funkce $\mu_1 - \mu_2$ a $\frac{\sigma_1^2}{\sigma_2^2}$ vycházíme z těchto pivotových statistik:

- $U = \frac{(M_1-M_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \sim N(0,1) \ldots$ slouží k řešení úloh o $\mu_1 - \mu_2$ když σ_1^2 a σ_2^2 známe,
- $K = \frac{(n_1+n_2-2)S_*^2}{\sigma^2} \sim \chi^2(n_1+n_2-2) \ldots$ slouží k řešení úloh o neznámém společném rozptylu σ^2,
- za předpokladu, že $\sigma_1^2 = \sigma_2^2 = \sigma^2$, pak $T = \frac{(M_1-M_2)-(\mu_1-\mu_2)}{S_*\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t(n_1+n_2-2) \ldots$ slouží k řešení úloh o $\mu_1 - \mu_2$, když σ_1^2 a σ_2^2 neznáme, ale víme, že jsou shodné,
- $F = \frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2} \sim F(n_1-1, n_2-1) \ldots$ slouží k řešení úloh o $\frac{\sigma_1^2}{\sigma_2^2}$.

Interval spolehlivosti pro $\mu_1 - \mu_2$, když σ_1^2, σ_2^2 známe (využití pivotové statistiky U):

- oboustranný:

$$(d;h) = \left(m_1 - m_2 - \sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}u_{1-\alpha/2}; m_1 - m_2 + \sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}u_{1-\alpha/2}\right), \tag{9.1}$$

- levostranný:

$$(d;\infty) = \left(m_1 - m_2 - \sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}u_{1-\alpha}; \infty\right), \tag{9.2}$$

- pravostranný:

$$(-\infty;h) = \left(-\infty; m_1 - m_2 + \sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}u_{1-\alpha}\right). \tag{9.3}$$

Interval spolehlivosti pro $\mu_1 - \mu_2$, když σ_1^2, σ_2^2 neznáme, ale víme, že jsou shodné (využití pivotové statistiky T):

- oboustranný:

$$(d;h) = \left(m_1 - m_2 - s_*\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}t_{1-\alpha/2}(n_1+n_2-2); m_1 - m_2 + s_*\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}t_{1-\alpha/2}(n_1+n_2-2)\right), \tag{9.4}$$

- levostranný:

$$(d;\infty) = \left(m_1 - m_2 - s_*\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}t_{1-\alpha}(n_1+n_2-2); \infty\right), \tag{9.5}$$

- pravostranný:

$$(-\infty;h) = \left(-\infty; m_1 - m_2 + s_*\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}t_{1-\alpha}(n_1+n_2-2)\right). \tag{9.6}$$

Upozornění: Není-li splněn předpoklad o shodě rozptylů, lze sestrojit přibližný $100(1-\alpha)\%$ interval spolehlivosti pro $\mu_1 - \mu_2$. Nazývá se Welchův interval spolehlivosti. V tomto případě má statistika T přibližně rozložení $t(df)$, kde počet stupňů volnosti

$$df = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{\frac{(s_1^2/n_1)^2}{n_1-1} + \frac{(s_2^2/n_2)^2}{n_2-1}}. \tag{9.7}$$

Interval spolehlivosti pro $\frac{\sigma_1^2}{\sigma_2^2}$ (využití pivotové statistiky F):

- oboustranný:

$$(d;h) = \left(\frac{s_1^2/s_2^2}{F_{1-\alpha/2}(n_1-1,n_2-1)}; \frac{s_1^2/s_2^2}{F_{\alpha/2}(n_1-1,n_2-1)}\right), \tag{9.8}$$

- levostranný:

$$(d;\infty) = \left(\frac{s_1^2/s_2^2}{F_{1-\alpha}(n_1-1,n_2-1)}; \infty\right), \tag{9.9}$$

- pravostranný:

$$(0;h) = \left(0; \frac{s_1^2/s_2^2}{F_{\alpha}(n_1-1 n_2-1)}\right). \tag{9.10}$$

Příklad 9.1. Řešený příklad
Načtěte datový soubor 13-two-samples-correlations-trunk.txt. Za předpokladu, že délka dolní končetiny (lowex.L) mužů (resp. žen) pochází z normálního rozložení s parametry μ_1 (resp. μ_2) a σ_1^2 (resp. σ_2^2), sestrojte (a) 95% oboustranný interval spolehlivosti pro podíl σ_1^2/σ_2^2; (b) 90% pravostranný interval spolehlivosti pro rozdíl $\mu_1 - \mu_2$, za předpokladu, že rozptyly σ_1^2 a σ_2^2 jsou neznámé ale shodné (ověření tohoto předpokladu viz příklad 9.5).

Řešení příkladu 9.1
Nejprve je třeba ověřit normální rozložení náhodného výběru délek dolní končetiny mužů (resp. žen). Na hladině významnosti $\alpha = 0,05$ testujeme H_0: *Data pochází z normálního rozložení.* proti H_1: *Data nepochází z normálního rozložení.* K otestování předpokladu normality použijeme vzhledem k rozsahu náhodného výběru ($n_1 = 75$, resp. $n_2 = 100$) Andersonův-Darlingův test (resp. Lillieforsův test).

```
data <- read.delim('13-two-samples-correlations-trunk.txt', sep = '\t')
lowex.LM <- na.omit(data[data$sex == 'm', 'lowex.L'])
lowex.LF <- na.omit(data[data$sex == 'f', 'lowex.L'])
nortest::ad.test(lowex.LM)$p.val # 0,1313783
nortest::lillie.test(lowex.LF)$p.val # 0,8859276
```

Náhodný výběr délek dolních končetin mužů pochází z normálního rozložení (p-hodnota $= 0,1314$). Náhodný výběr délek dolních končetin žen pochází z normálního rozložení (p-hodnota $= 0,8859$).

Hranice 95% oboustranného intervalu spolehlivosti pro podíl σ_1^2/σ_2^2 bychom získali dosazením do vzorce 9.8. V softwaru je získáme jako výstup conf.int funkce var.test(). Hranice 90% pravostranného intervalu spolehlivosti pro rozdíl $\mu_1 - \mu_2$ za předpokladu, že rozptyly σ_1^2 a σ_2^2 jsou neznámé, ale shodné, bychom získali po dosazení do vzorce 9.6. V softwaru je získáme jako výstup conf.int funkce t.test() s nastavením argumentu var.equal = T.

```
IS.mm <- var.test(lowex.LM, lowex.LF, alternative = 'two.sided',
                  conf.level = 0.95)$conf.int[c(1, 2)] # 0,7020597; 1,6553295
```

95% oboustranný interval spolehlivosti pro podíl σ_1^2/σ_2^2 má tvar $(0,70; 1,66)$. S pravděpodobností 95 % se bude podíl rozptylů délek dolních končetin mužů a žen nacházet v rozmezí 0,70–1,66.

```
IS.ss <- t.test(lowex.LM, lowex.LF, alternative = 'less', conf.level = 0.90,
                var.equal = T)$conf.int[c(1, 2)] # -Inf; 77,32987
```

90% pravostranný interval spolehlivosti pro rozdíl $\mu_1 - \mu_2$ má tvar $(-\infty; 77,33)$. S pravděpodobností 90 % bude rozdíl středních hodnot délek dolních končetin mužů a žen menší než $77,33\,\mathrm{mm}$. ★

Příklad 9.2. Řešený příklad
Načtěte datový soubor 31-goldman-alaska.csv. Za předpokladu, že anteroposteriorní průměr diafýzy pažní kosti z levé strany (humer.ADL) žen z aljašské populace z kmene Tigara (resp. z kmene Ipituaq) pochází z normálního rozložení s parametry μ_1 (resp. μ_2) a σ_1^2 (resp. σ_2^2), sestrojte (a) 95% oboustranný interval spolehlivosti pro rozdíl $\mu_1 - \mu_2$, za předpokladu, že rozptyly σ_1^2 a σ_2^2 jsou neznámé a různé (ověření tohoto předpokladu viz příklad 9.6); (b) 99% levostranný interval spolehlivosti pro podíl σ_1^2/σ_2^2.

Řešení příkladu 9.2
Nejprve je třeba ověřit normální rozložení náhodného výběru anteroposteriorních průměrů diafýzy pažní kosti z levé strany žen z kmene Tigara (resp. z kmene Ipituaq). Na hladině významnosti $\alpha = 0,05$ testujeme H_0: *Data pochází z normálního rozložení.* proti H_1: *Data nepochází z normálního rozložení.* K otestování předpokladu normality použijeme vzhledem k rozsahu náhodného výběru ($n_1 = 10$, resp. $n_2 = 25$) Shapirův-Wilkův test.

```
data <- read.delim('31-goldman-alaska.csv', sep = ';')
humer.ADLT <- na.omit(data[data$pop == 'Tigara' & data$sex == 'f', 'humer.ADL'])
humer.ADLI <- na.omit(data[data$pop == 'Ipituaq' & data$sex == 'f', 'humer.ADL'])
shapiro.test(humer.ADLT)$p.val # 0,7238248
shapiro.test(humer.ADLI)$p.val # 0,5130621
```

Náhodný výběr anteroposteriorních průměrů diafýzy pažní kosti z levé strany žen z kmene Tigara pochází z normálního rozložení (p-hodnota $= 0,7238$). Náhodný výběr anteroposteriorních průměrů diafýzy pažní kosti z levé strany žen z kmene Ipituaq pochází z normálního rozložení (p-hodnota $= 0,5131$).

Hranice 95% oboustranného intervalu spolehlivosti pro rozdíl $\mu_1 - \mu_2$ za předpokladu, že rozptyly σ_1^2 a σ_2^2 jsou neznámé a různé, získáme jako výstup conf.int funkce t.test() s nastavením argumentu var.equal = F. Hranice 99% levostranného intervalu spolehlivosti pro podíl σ_1^2/σ_2^2 bychom získali dosazením do vzorce 9.9. V softwaru je získáme jako výstup conf.int funkce var.test().

```
IS.mm <- t.test(humer.ADLT, humer.ADLI, alternative = 'two.sided', conf.level = 0.95,
                var.equal = F)$conf.int[c(1, 2)] # -2,5142107; 0,6482107
```

95% oboustranný interval spolehlivosti pro rozdíl $\mu_1 - \mu_2$ má tvar $(-2,51; 0,65)$. S pravděpodobností 95 % se bude rozdíl středních hodnot anteroposteriorních průměrů diafýzy z levé strany u žen z kmene Tigara a z kmene Ipituaq nacházet v rozmezí $-2,51$–$0,65\,\mathrm{mm}$.

```
IS.ss <- var.test(humer.ADLT, humer.ADLI, alternative = 'greater',
                  conf.level = 0.99)$conf.int[c(1, 2)] # 0,0744114; Inf
```

99% levostranný interval spolehlivosti pro podíl σ_1^2/σ_2^2 má tvar $(0,074; \infty)$. S pravděpodobností 99 % bude podíl rozptylů anteroposteriorních průměrů diafýzy z levé strany u žen z kmene Tigara a z kmene Ipituaq větší než 0,074. ★

Příklad 9.3. Neřešený příklad
Načtěte datový soubor 15-anova-means-skull.txt. Za předpokladu, že výška horní části tváře (upface.H) mužů bantuské (resp. malajské) populace pochází z normálního rozložení s parametry μ_1 (resp. μ_2) a σ_1^2 (resp. σ_2^2), sestrojte (a) 90% levostranný interval spolehlivosti pro rozdíl $\mu_1 - \mu_2$, za předpokladu, že rozptyly σ_1^2 a σ_2^2 jsou neznámé a shodné (ověření tohoto předpokladu viz příklad 9.7); (b) 95% pravostranný interval spolehlivosti pro podíl σ_1^2/σ_2^2.
Výsledky: Shapirův-Wilkův test (bantuská populace): p-hodnota $= 0,4321$, $\alpha = 0,05$; data pochází z normálního rozložení; Andersonův-Darlingův test (malajská populace): p-hodnota $= 0,4964$, $\alpha = 0,05$; data pochází z normálního rozložení; (a) $IS = (-1,40; \infty)$; (b) $IS = (0; 2,81)$. ★

Příklad 9.4. Neřešený příklad

Načtěte datový soubor 31-goldman-alaska.csv. Za předpokladu, že délka stehenní kosti z pravé strany (femur.R) mužů (resp. žen) z aljašské populace z kmene Ipituaq pochází z normálního rozložení s parametry μ_1 (resp. μ_2) a σ_1^2 (resp. σ_2^2), sestrojte (a) 90% oboustranný interval spolehlivosti pro podíl σ_1^2/σ_2^2; (b) 99% pravostranný interval spolehlivosti pro rozdíl $\mu_1 - \mu_2$, za předpokladu, že rozptyly σ_1^2 a σ_2^2 jsou neznámé a různé (ověření tohoto předpokladu viz příklad 9.8).
Výsledky: Shapirův-Wilkův test (muži): p-hodnota $= 0,2078$, $\alpha = 0,05$; data pochází z normálního rozložení; Shapirův-Wilkův test (ženy): p-hodnota $= 0,3491$, $\alpha = 0,05$; data pochází z normálního rozložení; (a) $IS = (1,17; 4,74)$; (b) $IS = (0; 48,84)$. ★

9.2 Testy o rozdílu středních hodnot a podílu rozptylů

$X_{11}, \dots, X_{1n_1}$ je náhodný výběr z rozložení $N(\mu_1, \sigma_1^2)$ a $X_{21}, \dots, X_{2n_2}$ je na něm nezávislý náhodný výběr z rozložení $N(\mu_2, \sigma_2^2)$, $n_1 \geq 2$, $n_2 \geq 2$, c je reálná konstanta a $\alpha \in (0,1)$ je hladina významnosti.

Test o rozdílu středních hodnot při známých rozptylech (dvouvýběrový z-test):
Na hladině významnosti α testujeme H_0: $\mu_1 - \mu_2 = c$ proti H_1: $\mu_1 - \mu_2 \neq c$ resp. proti H_1: $\mu_1 - \mu_2 < c$ resp. proti H_1: $\mu_1 - \mu_2 > c$. Vypočteme realizaci testové statistiky $t_0 = \frac{(m_1 - m_2) - c}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$. Stanovíme kritický obor W. Pokud $t_0 \in W$, H_0 zamítáme na hladině významnosti α a přijímáme H_1.
Pro oboustranný test má kritický obor tvar: $W = (-\infty; -u_{1-\alpha/2}\rangle \cup \langle u_{1-\alpha/2}; \infty)$, pro levostranný test $W = (-\infty; -u_{1-\alpha}\rangle$ a pro pravostranný test $W = \langle u_{1-\alpha}; \infty)$.

Test o rozdílu středních hodnot při neznámých, ale shodných rozptylech (dvouvýběrový t-test):
Formulace H_0 a H_1 je stejná jako u dvouvýběrového z-testu. Testová statistika má realizaci $t_0 = \frac{(m_1 - m_2) - c}{s_* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$. Kritické obory se od předešlé situace liší tím, že místo kvantilů $u_{1-\alpha/2}$ resp. $u_{1-\alpha}$ v nich figurují kvantily $t_{1-\alpha/2}(n_1 + n_2 - 2)$ resp. $t_{1-\alpha}(n_1 + n_2 - 2)$.

Upozornění: Není-li splněn předpoklad o shodě rozptylů, použije se Welchova modifikace dvouvýběrového t-testu.

Test o podílu rozptylů (F-test):
Na hladině významnosti α testujeme H_0: $\sigma_1^2/\sigma_2^2 = 1$ proti H_1: $\sigma_1^2/\sigma_2^2 \neq 1$ resp. proti H_1: $\sigma_1^2/\sigma_2^2 < 1$ resp. H_1: $\sigma_1^2/\sigma_2^2 > 1$. Vypočteme realizaci testové statistiky $t_0 = s_1^2/s_2^2$. Stanovíme kritický obor W. Pokud $t_0 \in W$, H_0 zamítáme na hladině významnosti α a přijímáme H_1.
Pro oboustranný test má kritický obor tvar: $W = (0; F_{\alpha/2}(n_1 - 1, n_2 - 1)\rangle \cup \langle F_{1-\alpha/2}(n_1 - 1, n_2 - 1); \infty)$, pro levostranný test $W = (0; F_{\alpha}(n_1 - 1, n_2 - 1)\rangle$ a pro pravostranný test $W = \langle F_{1-\alpha}(n_1 - 1, n_2 - 1); \infty)$.

Příklad 9.5. Řešený příklad
Načtěte datový soubor 13-two-samples-correlations-trunk.txt. Za předpokladu, že délka dolní končetiny (lowex.L) mužů (resp. žen) pochází z normálního rozložení s parametry μ_1 (resp. μ_2) a σ_1^2 (resp. σ_2^2), testujte hypotézu, že (a) rozptyl délky dolní končetiny mužů a žen je shodný; (b) střední hodnota délky dolní končetiny mužů a žen je shodná. Testování proveďte vždy (1) kritickým oborem; (2) intervalem spolehlivosti; (3) p-hodnotou. Hladinu významnosti zvolte $\alpha = 0,05$.

Řešení příkladu 9.5
Nejprve je třeba ověřit normální rozložení náhodného výběru délek dolních končetin mužů (resp. žen). Tento předpoklad bychom ověřili analogicky jako v příkladu 9.1. Náhodný výběr délek dolních končetin mužů (resp. žen) pochází z normálního rozložení (p-hodnota $= 0,1314$, resp. p-hodnota $= 0,8859$, $\alpha = 0,05$).

```
data <- read.delim('13-two-samples-correlations-trunk.txt', sep = '\t')
lowex.LM <- na.omit(data[data$sex == 'm', 'lowex.L'])
lowex.LF <- na.omit(data[data$sex == 'f', 'lowex.L'])
```

(a) Na hladině významnosti $\alpha = 0,05$ testujeme H_0: $\sigma_1^2/\sigma_2^2 = 1$ proti H_1: $\sigma_1^2/\sigma_2^2 \neq 1$ (oboustranný test). K testování H_0 použijeme test o podílu rozptylů (F-test) implementovaný ve funkci var.test(). Výstupem funkce je realizace testové statistiky t_0, interval spolehlivosti a p-hodnota. Hranice kritického oboru dopočítáme příkazem qf().

```
alpha <- 0.05
n1 <- length(lowex.LM) # 75
n2 <- length(lowex.LF) # 100
qf(alpha / 2, n1 - 1, n2 - 1) # 0,6468828
qf(1 - alpha / 2, n1 - 1, n2 - 1) # 1,525232
var.test(lowex.LM, lowex.LF, alternative = 'two.sided', conf.level = 0.95)
```

```

        F test to compare two variances

data:  lowex.LM and lowex.LF
F = 1,0708, num df = 74, denom df = 99, p-value = 0,7453
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0,7020597 1,6553295
sample estimates:
ratio of variances
          1,070804
```

Realizace testové statistiky $t_0 = 1,0708$, kritický obor $W = (0; 0,6469\rangle \cup \langle 1,5252; \infty)$. Protože $t_0 \notin W$, H_0 nezamítáme na hladině významnosti $\alpha = 0,05$. Interval spolehlivosti $IS = (0,70; 1,66)$. Protože $c = 1 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0,05$. Protože p-hodnota $= 0,7453$ je větší než $\alpha = 0,05$, H_0 nezamítáme na hladině významnosti $\alpha = 0,05$. Mezi rozptylem délky dolní končetiny mužů a žen neexistuje statisticky významný rozdíl.

(b) Na hladině významnosti $\alpha = 0,05$ testujeme H_0: $\mu_1 - \mu_2 = 0$ proti H_1: $\mu_1 - \mu_2 \neq 0$ (oboustranný test). Protože v příkladu (a) jsme zjistili, že mezi rozptyly neexistuje statisticky významný rozdíl, použijeme k otestování H_0 test o rozdílu středních hodnot při neznámých, ale shodných rozptylech. Tento test je implementovaný ve funkci t.test() s nastavením argumentu var.equal = T. Výstupem funkce je realizace testové statistiky t_0, interval spolehlivosti a p-hodnota. Hranice kritického oboru dopočítáme příkazem qt().

```
qt(alpha / 2, n1 + n2 - 2) # -1,973771
qt(1 - alpha / 2, n1 + n2 - 2) # 1,973771
t.test(lowex.LM, lowex.LF, alternative = 'two.sided', conf.level = 0.95, var.equal = T)
```

```

        Two Sample t-test

data:  lowex.LM and lowex.LF
t = 9,682, df = 173, p-value < 2,2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 54,34446 82,17554
sample estimates:
mean of x mean of y
  1008,76    940,50
```

Realizace testové statistiky $t_0 = 9,682$, kritický obor $W = (-\infty; -1,9738\rangle \cup \langle 1,9738; \infty)$. Protože $t_0 \in W$, H_0 zamítáme na hladině významnosti $\alpha = 0,05$. Interval spolehlivosti $IS = (54,34; 82,18)$. Protože $c = 0 \notin IS$, H_0 zamítáme na hladině významnosti $\alpha = 0,05$. Protože p-hodnota $< 0,001$ je menší než $\alpha = 0,05$, H_0 zamítáme na hladině významnosti $\alpha = 0,05$. Mezi střední hodnotou délky dolní končetiny mužů a žen existuje statisticky významný rozdíl.

★

Příklad 9.6. Řešený příklad

Načtěte datový soubor 31-goldman-alaska.csv. Za předpokladu, že anteroposteriorní průměr diafýzy pažní kosti z levé strany (humer.ADL) žen z aljašské populace z kmene Tigara (resp. z kmene Ipituaq) pochází z normálního rozložení s parametry μ_1 (resp. μ_2) a σ_1^2 (resp. σ_2^2), testujte hypotézu, že (a) střední hodnota anteroposteriorního průměru diafýzy pažní kosti z levé strany u žen z kmene Tigara je menší nebo rovna střední hodnotě u žen z kmene Ipituaq; (b) rozptyl anteroposteriorního průměru diafýzy pažní kosti z levé strany u žen z kmene Tigara je větší nebo roven rozptylu u žen z kmene Ipituaq. Testování proveďte vždy (1) kritickým oborem; (2) intervalem spolehlivosti; (3) p-hodnotou. Hladinu významnosti zvolte $\alpha = 0,10$.

Řešení příkladu 9.6

Nejprve je třeba ověřit normální rozložení náhodného výběru anteroposteriorních průměrů diafýzy pažní kosti z levé strany žen z kmene Tigara (resp. z kmene Ipituaq). Tento předpoklad bychom ověřili analogicky jako příkladu 9.2. Náhodný výběr anteroposteriorních průměrů diafýzy pažní kosti z levé strany žen z kmene Tigara (resp. z kmene Ipituaq) pochází z normálního rozložení (p-hodnota $= 0,7238$, resp. p-hodnota $= 0,5131$, $\alpha = 0,10$).

(a) Na hladině významnosti $\alpha = 0,10$ testujeme H_0: $\mu_1 - \mu_2 \leq 0$ proti H_1: $\mu_1 - \mu_2 > 0$ (pravostranný test). Před zvolením vhodného testu musíme nejprve ověřit, zda jsou rozptyly obou výběrů shodné nebo různé. Na hladině významnosti $\alpha = 0,10$ testujeme tedy nejprve H_0: $\sigma_1^2/\sigma_2^2 = 1$ oproti H_1: $\sigma_1^2/\sigma_2^2 \neq 1$ (oboustranný test). Závěr stanovíme na základě F-testu (funkce var.test()), a to pomocí p-hodnoty.

```
data <- read.delim('31-goldman-alaska.csv', sep = ';')
humer.ADLT <- na.omit(data[data$pop == 'Tigara' & data$sex == 'f', 'humer.ADL'])
humer.ADLI <- na.omit(data[data$pop == 'Ipituaq' & data$sex == 'f', 'humer.ADL'])
var.test(humer.ADLT, humer.ADLI, alternative = 'two.sided',
         conf.level = 0.90)$p.val # 0,03954399
```

Protože p-hodnota $= 0,0395$ je menší než $\alpha = 0,10$, H_0 zamítáme na hladině významnosti $\alpha = 0,10$. Mezi rozptylem anteroposteriorních průměrů diafýzy pažní kosti z levé strany žen z kmene Tigara a žen z kmene Ipituaq existuje statisticky významný rozdíl.

Protože rozptyly obou výběrů nejsou shodné, použijeme na otestování H_0: $\mu_1 - \mu_2 \leq 0$ proti H_1: $\mu_1 - \mu_2 > 0$ test o rozdílu středních hodnot s Welchovou modifikací stupňů volnosti. Tento test je implementovaný ve funkci t.test() s nastavením argumentu var.equal = F. Výstupem funkce je realizace testové statistiky t_0, počet stupňů volnosti df, interval spolehlivosti a p-hodnota. Dolní hranici kritického oboru dopočítáme příkazem qt().

```
alpha <- 0.10
df <- t.test(humer.ADLT, humer.ADLI, var.equal = F)$parameter # 11,62556
qt(1 - alpha, df) # 1,358764
t.test(humer.ADLT, humer.ADLI, alternative = 'greater', conf.level = 0.90,
       var.equal = F)
```

```

        Welch Two Sample t-test

data:  humer.ADLT and humer.ADLI
t = -1,2902, df = 11,626, p-value = 0,889
alternative hypothesis: true difference in means is greater than 0
90 percent confidence interval:
 -1,915563       Inf
sample estimates:
mean of x mean of y
   19,036    19,969
```

Realizace testové statistiky $t_0 = -1,2902$, kritický obor $W = \langle 1,3588; \infty)$. Protože $t_0 \notin W$, H_0 nezamítáme na hladině významnosti $\alpha = 0,10$. Interval spolehlivosti $IS = (-1,92; \infty)$. Protože $c = 0 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0,10$. Protože p-hodnota $= 0,8890$ je větší než $\alpha = 0,10$, H_0 nezamítáme na hladině významnosti $\alpha = 0,10$. Střední hodnota anteroposteriorního průměru diafýzy pažní kosti z levé strany u žen z kmene Tigara není statisticky významně větší než u žen z kmene Ipituaq.

(b) Na hladině významnosti $\alpha = 0,10$ testujeme H_0: $\sigma_1^2/\sigma_2^2 \geq 1$ proti H_1: $\sigma_1^2/\sigma_2^2 < 1$ (levostranný test). K testování použijeme test o podílu rozptylů (F-test) implementovaný ve funkci var.test(). Výstupem funkce je realizace testové statistiky t_0, interval spolehlivosti a p-hodnota. Horní hranici kritického oboru dopočítáme příkazem qf().

```
n1 <- length(humer.ADLT) # 25
n2 <- length(humer.ADLI) # 10
qf(alpha, n1 - 1, n2 - 1) # 0,5245889
var.test(humer.ADLT, humer.ADLI, alternative = 'less', conf.level = 0.90)
```

```

       F test to compare two variances

data:  humer.ADLT and humer.ADLI
F = 0,35189, num df = 24, denom df = 9, p-value = 0,01977
alternative hypothesis: true ratio of variances is less than 1
90 percent confidence interval:
 0,0000000 0,6707945
sample estimates:
ratio of variances
         0,3518913
```

Realizace testové statistiky $t_0 = 0,3519$, kritický obor $W = (0; 0,5246\rangle$. Protože $t_0 \in W$, H_0 zamítáme na hladině významnosti $\alpha = 0,10$. Interval spolehlivosti $IS = (0; 0,67)$. Protože $c = 1 \notin IS$, H_0 zamítáme na hladině významnosti $\alpha = 0,10$. Protože p-hodnota $= 0,01977$ je menší než $\alpha = 0,10$, H_0 zamítáme na hladině významnosti $\alpha = 0,10$. Rozptyl anteroposteriorního průměru diafýzy pažní kosti z levé strany u žen z kmene Tigara je statisticky významně menší než u žen z kmene Ipituaq.

★

Příklad 9.7. Neřešený příklad

Načtěte datový soubor 15-anova-means-skull.txt. Za předpokladu, že výška horní části tváře (upface.H) mužů bantuské (resp. malajské) populace pochází z normálního rozložení s parametry μ_1 (resp. μ_2) a σ_1^2 (resp. σ_2^2), testujte hypotézu, že (a) střední hodnota výšky horní části tváře u mužů bantuské je menší nebo rovná střední hodnotě u mužů malajské populace; (b) rozptyl výšky horní části tváře u mužů bantuské populace je větší nebo rovný rozptylu u mužů malajské populace. Testování proveďte vždy (1) kritickým oborem; (2) intervalem spolehlivosti; (3) p-hodnotou. Hladinu významnosti zvolte $\alpha = 0,01$.

Výsledky: Shapirův-Wilkův test (bantuská populace): p-hodnota $= 0,4321$, $\alpha = 0,01$; data pochází z normálního rozložení; Andersonův-Darlingův test (malajská populace): p-hodnota $= 0,4964$, $\alpha = 0,01$; data pochází z normálního rozložení; (a) test předpokladu o shodě rozptylů: p-hodnota $= 0,6232$, $\alpha = 0,01$; rozptyly obou výběrů jsou shodné; test o rozdílu středních hodnot při neznámých ale shodných rozptylech: $t_0 = 0,3701$, $W = \langle 2,3739; \infty)$; $IS = (-3,04; \infty)$, $c = 0$; p-hodnota $= 0,3561$, $\alpha = 0,01$; H_0 nezamítáme na hladině významnosti $\alpha = 0,01$; (b) test o podílu rozptylů: $t_0 = 1,1850$, $W = (0; 0,2845\rangle$; $IS = (0; 4,1658)$, $c = 1$; p-hodnota $= 0,6884$, $\alpha = 0,01$; H_0 nezamítáme na hladině významnosti $\alpha = 0,01$. ★

Příklad 9.8. Neřešený příklad

Načtěte datový soubor 31-goldman-alaska.csv. Za předpokladu, že délka stehenní kosti z pravé strany (femur.R) mužů (resp. žen) z aljašské populace z kmene Ipituaq pochází z normálního rozložení s parametry μ_1 (resp. μ_2) a σ_1^2 (resp. σ_2^2), testujte hypotézu, že (a) rozptyl délky stehenní kosti z pravé strany u mužů je menší nebo roven rozptylu u žen z kmene Ipituaq; (b) střední hodnota délky stehenní kosti z pravé strany u mužů a žen z kmene Ipituaq je shodná. Testování proveďte vždy (1) kritickým oborem; (2) intervalem spolehlivosti; (3) p-hodnotou. Hladinu významnosti zvolte $\alpha = 0,05$.

Výsledky: Shapirův-Wilkův test (muži): p-hodnota $= 0,2078$, $\alpha = 0,05$; data pochází z normálního rozložení; Shapirův-Wilkův test (ženy): p-hodnota $= 0,3491$, $\alpha = 0,05$; data pochází z normálního rozložení; (a) test o podílu rozptylů: $t_0 = 2,3545$, $W = \langle 2,0144; \infty)$; $IS = (1,1688; \infty)$, $c = 1$; p-hodnota $= 0,02266$, $\alpha = 0,05$; H_0 zamítáme na hladině významnosti $\alpha = 0,05$; (b) test předpokladu o shodě rozptylů: p-hodnota $= 0,0453$, $\alpha = 0,05$; rozptyly obou výběrů nejsou shodné; test o rozdílu středních hodnot při neznámých a různých rozptylech: $t_0 = 6,2086$, $W = (-\infty; -2,0129\rangle \cup \langle 2,0129; \infty)$; IS = (23,69; 46,56), $c = 0$; p-hodnota $< 0,001$, $\alpha = 0,05$; H_0 zamítáme na hladině významnosti $\alpha = 0,05$. ★

9.3 Interval spolehlivosti pro pravděpodobnost

$X_1, \ldots, X_n$ je náhodný výběr z rozložení $A(\vartheta)$. (Realizací tohoto náhodného výběru je posloupnost 0 a 1, veličina X_i nabývá hodnoty 1, když v i-tém pokusu nastane úspěch s pravděpodobností ϑ a hodnoty 0, když nastane neúspěch, $i = 1, \ldots, n$. Realizace m výběrového průměru M je tedy relativní četností úspěchu.) Při konstrukci intervalu spolehlivosti pro parametr ϑ použijeme pivotovou statistiku $U = \frac{M-\vartheta}{\sqrt{\frac{M(1-M)}{n}}} \approx N(0,1)$.

(Aproximace rozložením $N(0,1)$ se dá korektně použít, je-li splněna podmínka $n\vartheta(1-\vartheta) > 9$. V praxi se při ověřování této podmínky parametr ϑ nahradí m.)

$100(1-\alpha)\%$ asymptotický interval spolehlivosti pro ϑ:

- oboustranný:
$$(d;h) = \left(m - \sqrt{\frac{m(1-m)}{n}}u_{1-\alpha/2}; m + \sqrt{\frac{m(1-m)}{n}}u_{1-\alpha/2}\right), \tag{9.11}$$

- levostranný:
$$(d;1) = \left(m - \sqrt{\frac{m(1-m)}{n}}u_{1-\alpha}; 1\right), \tag{9.12}$$

- pravostranný:
$$(0;h) = \left(0; m + \sqrt{\frac{m(1-m)}{n}}u_{1-\alpha}\right). \tag{9.13}$$

Příklad 9.9. Řešený příklad
Načtěte datový soubor 25-one-sample-probability-dermatoglyphs.txt obsahující údaje o frekvenci výskytu dermatoglyfických vzorů *vír*, *smyčka* a *oblouček* na deseti prstech 470 jedinců (235 mužů a 235 žen) bagathské populace z Araku Valley. Za předpokladu, že náhodný výběr výskytů dermatoglyfického vzoru *vír* na prstech žen bagathské populace pochází z alternativního rozložení s parametrem ϑ, sestrojte (a) 90% pravostranný interval spolehlivosti pro parametr ϑ; (b) 95% oboustranný interval spolehlivosti pro parametr ϑ.

Řešení příkladu 9.9
Nejprve je třeba ověřit podmínku dobré aproximace $n\vartheta(1-\vartheta) > 9$. Protože hodnotu ϑ neznáme, nahradíme ji výběrovým průměrem m.

```
data <- read.delim('25-one-sample-probability-dermatoglyphs.txt')
presence <- data[data$pattern == 'whorl', 'f']
absence<- sum(data[data$pattern %in% c('loop', 'arch'), 'f'])
n <- presence + absence # 2350
m <- presence / n # 0,3744681
n * m * (1 - m) # 550,4681 > 9 -> splneno
```

Protože $nm(1-m) = 550,47$ je větší než 9, je podmínka dobré aproximace splněna a intervaly spolehlivosti můžeme sestrojit.

Hranice 90% pravostranného intervalu spolehlivosti pro parametr ϑ bychom získali dosazením do vzorce 9.13. Hranice 95% oboustranného intervalu spolehlivosti pro parametr ϑ bychom získali dosazením do vzorce 9.11. V softwaru R získáme hranice obou intervalů spolehlivosti jako výstup conf.int funkce prop.test().

```
PIS <- prop.test(x = presence, n = n, alternative = 'less',
                 conf.level = 0.90)$conf.int[c(1, 2)] # 0; 0,3873464
```

90% pravostranný interval spolehlivosti pro parametr ϑ má tvar $(0; 0,3873)$. S pravděpodobností 90 % bude pravděpodobnost výskytu dermatoglyfického vzoru *vír* na prstech žen bagathské populace z Araku Valley menší než 38,73 %.

```
OIS <- prop.test(x = presence, n = n, alternative = 'two.sided',
                 conf.level = 0.95)$conf.int[c(1, 2)] # 0,3513459; 0,3981724
```

95% oboustranný interval spolehlivosti pro parametr ϑ má tvar $(0,3513; 0,3982)$. S pravděpodobností 95 % bude pravděpodobnost výskytu dermatoglyfického vzoru *vír* na prstech žen bagathské populace z Araku Valley v rozmezí $35,13\%$–$39,82\,\%$. ★

Příklad 9.10. Neřešený příklad
Načtěte datový soubor 27-sutura-metopica-anatol.csv obsahující údaje o výskytu epigenetického znaku *sutura metopica* u mužů a žen anatolské populace z období Byzantské říše. Za předpokladu, že náhodný výběr výskytů epigenetického znaku *sutura metopica* u mužů a žen anatolské populace z období Byzantské říše pochází z alternativního rozložení s parametrem ϑ, sestrojte (a) 99% pravostranný interval spolehlivosti pro parametr ϑ; (b) 90% levostranný interval spolehlivosti pro parametr ϑ.
Výsledky: podmínka dobré aproximace je splněna ($nm(1-m) = 14.83 > 9$); (a) IS = (0; 0,1284); (b) IS = (0,05185; 1). ★

9.4 Test o pravděpodobnosti úspěchu

$X_1, \ldots, X_n$ je náhodný výběr z rozložení $A(\vartheta)$, $n\vartheta(1-\vartheta) > 9$, $c \in (0,1)$ je reálná konstanta, $\alpha \in (0,1)$ je hladina významnosti.

Na asymptotické hladině významnosti α testujeme hypotézu H_0: $\vartheta = c$ proti H_1: $\vartheta \neq c$ resp. proti H_1: $\vartheta < c$ resp. proti H_1: $\vartheta > c$. Vypočteme realizaci testové statistiky $t_0 = \frac{m-c}{\sqrt{\frac{c(1-c)}{n}}}$. Stanovíme kritický obor W. Pokud $t_0 \in W$, H_0 zamítáme na asymptotické hladině významnosti α a přijímáme H_1.
Pro oboustranný test má kritický obor tvar: $W = (-\infty; -u_{1-\alpha/2}\rangle \cup \langle u_{1-\alpha/2}; \infty)$, pro levostranný test $W = (-\infty; -u_{1-\alpha}\rangle$ a pro pravostranný test $W = \langle u_{1-\alpha}; \infty)$.

Příklad 9.11. Řešený příklad
Načtěte datový soubor 25-one-sample-probability-dermatoglyphs.txt obsahující údaje o frekvenci výskytu dermatoglyfických vzorů *vír*, *smyčka* a *oblouček* na deseti prstech 470 jedinců (235 mužů a 235 žen) bagathské populace z Araku Valley. Současně máme k dispozici hodnotu pravděpodobnosti výskytu dermatoglyfického vzoru *vír* u jedinců z populace Valmikis ($p_m = 0,4780$ (Rajendra, 1972), $p_f = 0,3500$ (Mrunalini, 1972)). Za předpokladu, že náhodný výběr výskytů dermatoglyfického vzoru *vír* na prstech žen z Araku Valley pochází z alternativního rozložení, testujte hypotézu, že pravděpodobnost výskytu dermatoglyfického vzoru *vír* u žen bagathské populace z Araku Valley je menší nebo rovna pravděpodobnosti výskytu téhož vzoru u žen z populace Valmikis. Testování proveďte vždy (1) kritickým oborem; (2) intervalem spolehlivosti; (3) p-hodnotou. Hladinu významnosti zvolte $\alpha = 0,01$.

Řešení příkladu 9.11
Nejprve je třeba ověřit podmínku dobré aproximace $n\vartheta(1-\vartheta) > 9$.

```
data <- read.delim('25-one-sample-probability-dermatoglyphs.txt')
presence <- data[data$pattern == 'whorl', 'f']
absence <- sum(data[data$pattern %in% c('loop', 'arch'), 'f'])
n <- presence + absence # 2350
theta0 <- 0.35
n * theta0 * (1 - theta0) # 534,625 > 9 -> splneno
```

Protože $n\vartheta(1-\vartheta) = 534,63$ je větší než 9, je podmínka dobré aproximace splněna a test o pravděpodobnosti úspěchu můžeme provést.

Na asymptotické hladině významnosti $\alpha = 0,01$ testujeme H_0: $\vartheta \leq 0,35$ proti H_1: $\vartheta > 0,35$ (pravostranný test). K testování použijeme test o pravděpodobnosti úspěchu implementovaný ve funkci prop.test(). Výstupem funkce je realizace testové statistiky t_0, interval spolehlivosti a p-hodnota. Dolní hranici kritického oboru dopočítáme příkazem qnorm().

```
alpha <- 0.01
qnorm(1 - alpha) # 1,644854
prop.test(x = presence, n = n, p = theta0, alternative = 'greater', conf.level = 0.99)
```

```

        1-sample proportions test with continuity correction

data:  presence out of n, null probability theta0
X-squared = 6,0772, df = 1, p-value = 0,006847
alternative hypothesis: true p is greater than 0,35
99 percent confidence interval:
 0,3513459 1,0000000
sample estimates:
        p
0,3744681
```

Realizace testové statistiky $t_0 = 6,0772$, kritický obor $W = \langle 1,6449; \infty)$. Protože $t_0 \in W$, H_0 zamítáme na asymptotické hladině významnosti $\alpha = 0,01$. Interval spolehlivosti $IS = (0,3513; 1)$. Protože $c = 0,35 \notin IS$, H_0 zamítáme na asymptotické hladině významnosti $\alpha = 0,01$. Protože p-hodnota $= 0,006847$ je menší než $\alpha = 0,01$, H_0 zamítáme na asymptotické hladině významnosti $\alpha = 0,01$. Pravděpodobnost výskytu dermatoglyfického vzoru *vír* na prstech žen bagathské populace z Araku Valley je statisticky významně větší než na prstech žen z populace Valmikis. ★

Příklad 9.12. Neřešený příklad

Načtěte datový soubor 27-sutura-metopica-anatol.csv obsahující údaje o výskytu epigenetického znaku *sutura metopica* u mužů a žen anatolské populace z období Byzantské říše. Dále mějme k dispozici údaj o výskytu epigenetické znaku *sutura metopica* ($p_{67} = 0,0550$) získaný jako výsledek studie zaměřené na výskyt tohoto znaku v anatolské populaci z 6.–7. století (rané období Byzantské říše; Gülec, 1988). Za předpokladu, že náhodný výběr výskytů epigenetického znaku *sutura metopica* u mužů a žen anatolské populace z období Byzantské říše pochází z alternativního rozložení, testujte hypotézu o shodě pravděpodobností výskytu epigenetického znaku *sutura metopica* v obou studiích. Testování proveďte vždy (1) kritickým oborem; (2) intervalem spolehlivosti; (3) p-hodnotou. Hladinu významnosti zvolte $\alpha = 0,05$.

Výsledky: podmínka dobré aproximace je splněna ($n\vartheta(1-\vartheta) = 11,33 > 9$); $t_0 = 1,0873$, $W = (-\infty; -1,96\rangle \cup \langle 1,96; \infty)$; $IS = (0,04390; 0,1186)$, $c = 0,0550$; p-hodnota $= 0,2971$, $\alpha = 0,05$; H_0 nezamítáme na asymptotické hladině významnosti $\alpha = 0,05$. ★