8 Ověřování normality, parametrické úlohy o jednom náhodném výběru z normálního rozložení

8.1 Grafické ověřování normality

$X_1, \ldots, X_n$ je náhodný výběr, o němž se domníváme, že pochází z normálního rozložení. Tuto domněnku můžeme vizuálně ověřit pomocí následujících grafů:

- kvantil-kvantilový graf (dvojice $(u_{\alpha_j}, x_{(j)})$ leží v těsné blízkosti přímky),
- histogram (jeho tvar se blíží Gaussově křivce).

Výhody grafického přístupu: Nelineární trend v kvantil-kvantilovém grafu může upozornit na vhodnost jiného než normálního rozložení. V grafech jsou dobře patrná vybočující pozorování.

Příklad 8.1. Řešený příklad
Načtěte datový soubor 01-one-sample-mean-skull-mf.txt. Pro největší délku mozkovny (skull.L) mužů starověké egyptské populace proveďte grafické ověření normality (a) histogramem superponovaným křivkou jádrového odhadu hustoty a Gaussovou křivkou; (b) kvantil-kvantilovým grafem.

Řešení příkladu 8.1
Histogram vykreslíme pomocí funkce hist(), kde v souladu se Sturgesovým pravidlem (viz kapitola XXX) rozdělíme naměřené hodnoty do devíti třídicích intervalů s hranicemi 164, 168, ..., 198 mm. Jádrový odhad hustoty vypočítáme z dat pomocí funkce density(). Hustotu normálního (Gaussova) rozložení vypočítáme pomocí funkce dnorm().

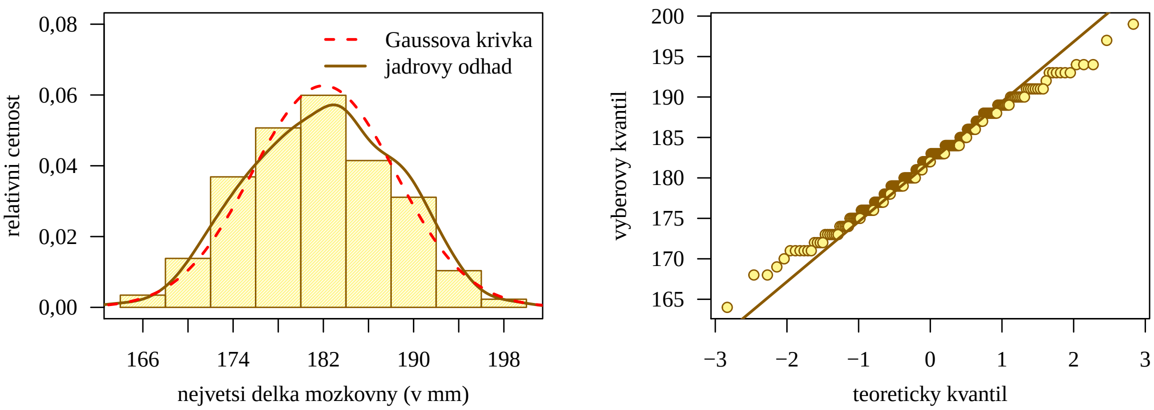
```
data <- read.delim('01-one-sample-mean-skull-mf.txt')
skull.LM <- na.omit(data[data$sex == 'm', 'skull.L'])
b <- seq(from = 164, to = 200, by = 4)
centr <- seq(from = 166, to = 198, by = 4)
hist(skull.LM, prob = T, breaks = b, ylim = c(0, 0.08), las = 1, col = 'khaki1',
     density = 60, border = 'orange4', xlab = '', ylab = 'relativni cetnost', main = '')
box(bty = 'o')
axis(1, centr)
axis(2, las = 1)
lines(density(skull.LM), col = 'orange4', lwd = 2)
xfit <- seq(from = 150, to = 210, length = 512)
yfit <- dnorm(xfit, mean(skull.LM), sd(skull.LM))
lines(xfit, yfit, col = 'red', lwd = 2, lty = 2)
mtext('nejvetsi delka mozkovny (v mm)', side = 1, line = 2.3)
legend('topright', lwd = c(2, 2), lty = c(2, 1), col = c('red', 'orange4'),
       legend = c('Gaussova krivka', 'jadrovy odhad'), bty = 'n')
```

Kvantil-kvantilový graf vykreslíme pomocí funkce qqnorm(), referenční přímku pomocí funkce qqline(). Histogram i kvantil-kvantilový graf jsou zobrazeny na obrázku 8.1.

```
qqnorm(skull.LM, pch = 21, col = 'orange4', bg = 'khaki1',
       main = '', xlab = '', ylab = 'vyberovy kvantil', las = 1)
qqline(skull.LM, col = 'orange4', lwd = 2)
mtext('teoreticky kvantil', side = 1, line = 2.3)
```

Křivka jádrového odhadu hustoty je svým tvarem a polohou blízká Gaussově křivce. Body v kvantil-kvantilovém grafu se realizují poblíž referenční přímky. Na základě grafické vizualizace předpokládáme, že náhodný výběr 217 největších délek mozkovny u mužů starověké egyptské populace pochází z normálního rozložení. ★

Příklad 8.2. Řešený příklad
Načtěte datový soubor 19-more-samples-correlations-skull.txt. Pro šířku nosu (nose.B) mužů peruánské populace proveďte grafické ověření normality (a) histogramem superponovaným křivkou jádrového odhadu hustoty a Gaussovou



Obrázek 8.1: (a) Histogram; (b) kvantil-kvantilový graf největší délky mozkovny mužů starověké egyptské populace

křivkou; (b) kvantil-kvantilovým grafem.

Řešení příkladu 8.2

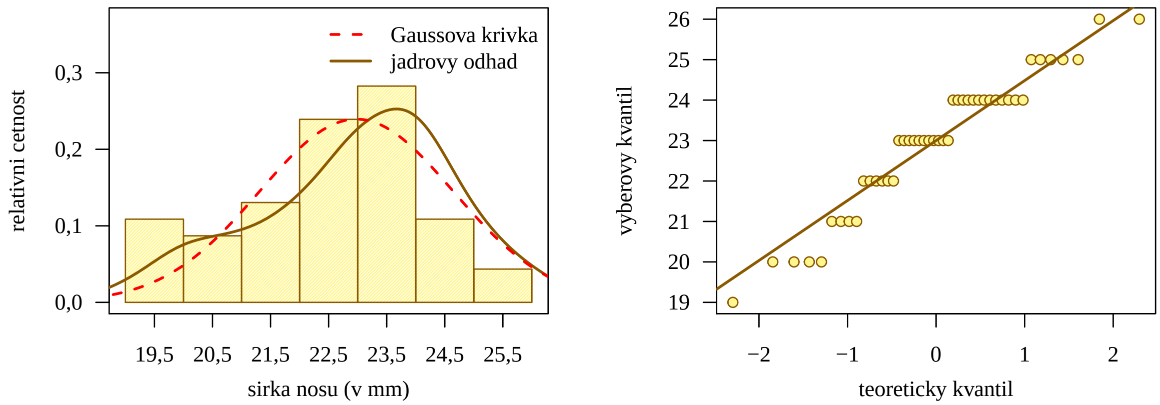
Histogram vykreslíme pomocí funkce hist(), kde naměřené hodnoty rozdělíme do sedmi třídicích intervalů s hranicemi 19, 20, …, 26 mm. Jádrový odhad hustoty vypočítáme z dat pomocí funkce density(). Hustotu Gaussova rozložení vypočítáme pomocí funkce dnorm().

```
data <- read.delim('19-more-samples-correlations-skull.txt')
nose.BP <- na.omit(data[data$pop == 'per', 'nose.B'])
b <- seq(from = 19, to = 26, by = 1)
centr <- seq(from = 19.5, 25.5, by = 1)
hist(nose.BP, prob = T, breaks = b, ylim = c(0, 0.37), axes = F, col = 'khaki1',
     density = 60, border = 'orange4', xlab = '', ylab = 'relativni cetnost', main = '')
box(bty = 'o')
axis(1, centr)
axis(2, las = 1)
lines(density(nose.BP), col = 'orange4', lwd = 2)
xfit <- seq(from = 0, to = 40, length = 512)
yfit <- dnorm(xfit, mean(nose.BP), sd(nose.BP))
lines(xfit, yfit, col = 'red', lwd = 2, lty = 2)
mtext('sirka nosu (v mm)', side = 1, line = 2.3)
legend('topright', lwd = c(2, 2), lty = c(2, 1), col = c('red', 'orange4'),
       legend = c('Gaussova krivka', 'jadrovy odhad'), bty = 'n')
```

Kvantil-kvantilový graf vykreslíme pomocí funkce qqnorm(), referenční přímku pomocí funkce qqline(). Histogram i kvantil-kvantilový graf jsou zobrazeny na obrázku 8.2.

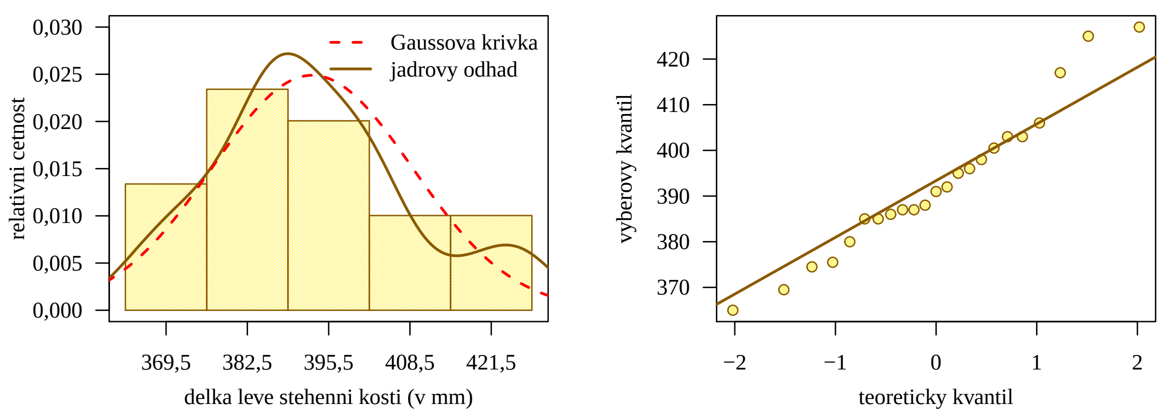
```
qqnorm(nose.BP, las = 1, pch = 21, col = 'orange4', bg = 'khaki1',
       xlab = '', ylab = 'vyberovy kvantil', main = '')
qqline(nose.BP, col = 'orange4', lwd = 2)
mtext('teoreticky kvantil', side = 1, line = 2.3)
```

Křivka jádrového odhadu hustoty se tvarem odlišuje od Gaussovy křivky. Body v kvantil-kvantilovém grafu se v pravém dolním rohu vzdalují od referenční přímky. Vzhledem k vyššímu rozsahu náhodného výběru ($n = 46$) by podobnost křivky jádrového dohadu hustoty a Gaussovy křivky měla být vyšší a příchylnost bodů k referenční přímce v kvantil-kvantilovém grafu by měla být těsnější. Na základě grafické vizualizace předpokládáme, že náhodný výběr šířek nosu mužů peruánské populace nepochází z normálního rozložení. ★



Obrázek 8.2: (a) Histogram; (b) kvantil-kvantilový graf šířky nosu pro muže peruánské populace

Příklad 8.3. Neřešený příklad
Načtěte datový soubor 31-goldman-alaska.csv. Pro délku stehenní kosti z levé strany (femur.L) žen aljašské populace z kmene Tigara proveďte grafické ověření normality (a) histogramem superponovaným křivkou jádrového odhadu hustoty a Gaussovou křivkou; (b) kvantil-kvantilovým grafem.



Obrázek 8.3: (a) Histogram; (b) kvantil-kvantilový graf délky stehenní kosti z levé strany žen aljašské populace z kmene Tigara

Výsledky: Histogram a kvantil-kvantilový graf viz obrázek 8.3. Na základě grafické vizualizace předpokládáme, že náhodný výběr 23 délek stehenních kostí z levé strany u žen aljašské populace z kmene Tigara pochází z normálního rozložení. ★

8.2 Testy normality

Na hladině významnosti α testujeme hypotézu, která tvrdí, že náhodný výběr $X_1, \ldots, X_n$ pochází z normálního rozložení s parametry μ a σ^2. Existuje řada testů normality, zde si uvedeme tři. Tyto testy nevyžadují specifikaci parametrů μ a σ^2:

- Lillieforsův test (je založen na porovnání teoretické a empirické distribuční funkce; pro svou dobrou vypovídací schopnost je pro výběry rozsahu $n \geq 100$ předepisován normou ČSN 01 0225 – Aplikovaná statistika. Testy shody empirického rozdělení s teoretickým.),

- Shapirův-Wilkův test (je založen na zjištění, zda body v kvantil-kvantilovém grafu jsou významně odlišné od regresní přímky proložené těmito body; funkce v R je určena pro výběry rozsahu $n = 3$ až $n = 5000$; Pro $n \leq 50$ je předepisován i normou ČSN 01 0225 – Aplikovaná statistika. Testy shody empirického rozdělení s teoretickým.),
- Andersonův-Darlingův test (je založen na analýze empirické distribuční funkce testovaného výběru; pro výběry rozsahu $n \geq 50$ je předepisován normou ČSN 01 0225 – Aplikovaná statistika. Testy shody empirického rozdělení s teoretickým.).

Upozornění: V praxi se doporučuje využívat kombinaci grafického přístupu a testů normality.

Příklad 8.4. Řešený příklad
Načtěte datový soubor 01-one-sample-mean-skull-mf.txt. Pro největší délku mozkovny (skull.L) mužů starověké egyptské populace proveďte ověření normality na základě (a) Shapirova-Wilkova testu; (b) Lillieforsova testu; (c) Andersonova-Darlingova testu. Hladinu významnosti α zvolte $0,05$.

Řešení příkladu 8.4
Na hladině významnosti $\alpha = 0,05$ testujeme H_0: *Data pochází z normálního rozložení.* oproti H_1: *Data nepochází z normálního rozložení.* Shapirův-Wilkův test provedeme pomocí funkce shapiro.test(), Lillieforsův test pomocí funkce lillie.test() z knihovny nortest a Andersonův-Darlingův test pomocí funkce ad.test() z knihovny nortest.

```
data <- read.delim('01-one-sample-mean-skull-mf.txt')
skull.LM <- na.omit(data[data$sex == 'm', 'skull.L'])
shapiro.test(skull.LM)
```

```

        Shapiro-Wilk normality test

data:  skull.LM
W = 0,99229, p-value = 0,3129
```

Protože p-hodnota $= 0,3129$ je větší než $\alpha = 0,05$, H_0 nezamítáme na hladině významnosti $\alpha = 0,05$.

```
nortest::lillie.test(skull.LM)
```

```

        Lilliefors (Kolmogorov-Smirnov) normality test

data:  skull.LM
D = 0,057783, p-value = 0,07545
```

Protože p-hodnota $= 0,0755$ je větší než $\alpha = 0,05$, H_0 nezamítáme na hladině významnosti $\alpha = 0,05$.

```
nortest::ad.test(skull.LM)
```

```

        Anderson-Darling normality test

data:  skull.LM
A = 0,58504, p-value = 0,1262
```

Protože p-hodnota $= 0,1262$ je větší než $\alpha = 0,05$, H_0 nezamítáme na hladině významnosti $\alpha = 0,05$. Náhodný výběr největších délek mozkovny mužů starověké egyptské populace pochází z normálního rozložení. ★

Příklad 8.5. Neřešený příklad
Načtěte datový soubor 19-more-samples-correlations-skull.txt. Pro šířku nosu (nose.B) mužů peruánské populace proveďte ověření normality na základě (a) Shapirova-Wilkova testu; (b) Lillieforsova testu; (c) Andersonova-Darlingova testu. Hladinu významnosti α zvolte $0,05$.

Výsledky: (a) Shapirův-Wilkův test: p-hodnota $= 0,0233$; (b) Lillieforsův test: p-hodnota $= 0,0008$; (c) Andersonův-Darlingův test: p-hodnota $= 0,0043$. Náhodný výběr šířek nosu mužů peruánské populace nepochází z normálního rozložení ($\alpha = 0,05$). ★

8.3 Intervaly spolehlivosti pro parametry μ a σ^2 normálního rozložení

$X_1, \ldots, X_n$ je náhodný výběr z rozložení $N(\mu, \sigma^2)$, $n \geq 2$. Označme M výběrový průměr a S^2 výběrový rozptyl tohoto výběru. Při konstrukci intervalů spolehlivosti pro parametry μ a σ^2 vycházíme z těchto pivotových statistik:

- $U = \frac{M-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$ ... slouží k řešení úloh o μ, když σ^2 známe (tato situace se v praxi vyskytuje málo, nebudeme se jí dále zabývat),
- $T = \frac{M-\mu}{\frac{S}{\sqrt{n}}} \sim t(n-1)$ ... slouží k řešení úloh o μ, když σ^2 neznáme,
- $K = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$ ... slouží k řešení úloh o σ^2, když μ neznáme,
- $\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2} \sim \chi^2(n)$ ... slouží k řešení úloh o σ^2, když μ známe (tato situace se v praxi vyskytuje málo, nebudeme se jí dále zabývat).

Interval spolehlivosti pro μ, když σ^2 známe (využití pivotové statistiky U):

- oboustranný:
$$(d;h) = \left(m - \frac{\sigma}{\sqrt{n}}u_{1-\alpha/2}; m + \frac{\sigma}{\sqrt{n}}u_{1-\alpha/2}\right), \tag{8.1}$$
- levostranný:
$$(d;\infty) = \left(m - \frac{\sigma}{\sqrt{n}}u_{1-\alpha}; \infty\right), \tag{8.2}$$
- pravostranný:
$$(-\infty;h) = \left(-\infty; m + \frac{\sigma}{\sqrt{n}}u_{1-\alpha}\right). \tag{8.3}$$

Interval spolehlivosti pro μ, když σ^2 neznáme (využití pivotové statistiky T):

- oboustranný:
$$(d;h) = \left(m - \frac{s}{\sqrt{n}}t_{1-\alpha/2}(n-1); m + \frac{s}{\sqrt{n}}t_{1-\alpha/2}(n-1)\right), \tag{8.4}$$
- levostranný:
$$(d;\infty) = \left(m - \frac{s}{\sqrt{n}}t_{1-\alpha}(n-1); \infty\right), \tag{8.5}$$
- pravostranný:
$$(-\infty;h) = \left(-\infty; m + \frac{s}{\sqrt{n}}t_{1-\alpha}(n-1)\right). \tag{8.6}$$

Interval spolehlivosti pro σ^2, když μ neznáme (využití pivotové statistiky K):

- oboustranný:
$$(d;h) = \left(\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)}; \frac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)}\right), \tag{8.7}$$
- levostranný:
$$(d;\infty) = \left(\frac{(n-1)s^2}{\chi^2_{1-\alpha}(n-1)}; \infty\right), \tag{8.8}$$

- pravostranný:

$$(0; h) = \left(0; \frac{(n-1)s^2}{\chi^2_{\alpha}(n-1)}\right). \tag{8.9}$$

Upozornění: Interval spolehlivosti pro σ získáme tak, že meze intervalu spolehlivosti pro σ^2 odmocníme.

Příklad 8.6. Řešený příklad

Načtěte datový soubor 11-two-samples-means-skull.txt. Za předpokladu, že výška lebky žen (skull.H) starověké egyptské populace pochází z normálního rozložení s parametry μ a σ^2, sestrojte (a) 95% oboustranný interval spolehlivosti pro parametru μ; (b) 99% levostranný interval spolehlivosti pro parametr σ^2; (c) 90% pravostranný interval spolehlivosti pro parametr σ.

Řešení příkladu 8.6

Nejprve je třeba ověřit normální rozložení náhodného výběru výšek lebky žen. Na hladině významnosti $\alpha = 0,05$ testujeme H_0: *Data pochází z normálního rozložení.* oproti H_1: *Data nepochází z normálního rozložení.* K otestování předpokladu normality použijeme vzhledem k rozsahu náhodného výběru ($n = 107$) Lillieforsův test.

```
data <- read.delim('11-two-samples-means-skull.txt')
skull.HF <- as.numeric(na.omit(data[data$sex == 'f', 'skull.H']))
nortest::lillie.test(skull.HF)$p.val # 0,09888442
```

Na hladině významnosti $\alpha = 0,05$ nelze zamítnout hypotézu, že náhodný výběr výšek lebky žen starověké egyptské populace pochází z normálního rozložení (p-hodnota $= 0,0989$).

Hranice 95% oboustranného intervalu spolehlivosti pro parametru μ bychom získali dosazením do vzorce 8.4. V softwaru je získáme jako výstup conf.int funkce t.test(). Hranice 99% levostranného intervalu spolehlivosti pro parametr σ^2 bychom získali dosazením do vzorce 8.8. V softwaru je získáme jako výstup conf.int funkce varTest() z knihovny EnvStats. Hranice 90% pravostranného intervalu spolehlivosti pro parametr σ bychom získali po dosazení do vzorce 8.9 a následným odmocněním výsledných hranic. V softwaru je získáme jako výstup conf.int funkce varTest() z knihovny EnvStats, který je nakonec třeba odmocnit.

```
IS.mu <- t.test(skull.HF, alternative = 'two.sided',
                conf.level = 0.95)$conf.int[1:2] # 124,7904; 126,5741
```

95% oboustranný interval spolehlivosti pro parametr μ má tvar $(124,8; 126,6)$. S pravděpodobností 95 % se bude střední hodnota výšky lebky žen nacházet v rozmezí 124,8–126,6 mm.

```
IS.sigma2 <- EnvStats::varTest(skull.HF, alternative = 'greater',
                               conf.level = 0.99)$conf.int[1:2] # 16,075; Inf
```

99% levostranný interval spolehlivosti pro parametr σ^2 má tvar $(16,1; \infty)$. S pravděpodobností 99 % bude rozptyl výšky lebky žen větší než $16,1\,\mathrm{mm}^2$.

```
IS.sigma <- EnvStats::varTest(skull.HF, alternative = 'less',
                              conf.level = 0.90)$conf.int[1:2]
sqrt(IS.sigma) # 0; 5,11224
```

90% pravostranný interval spolehlivosti pro parametr σ má tvar $(0; 5,1)$. S pravděpodobností 90 % bude směrodatná odchylka σ větší než $5,1$ mm. ★

Příklad 8.7. Neřešený příklad

Načtěte datový soubor 31-goldman-alaska.csv. Za předpokladu, že délka pažní kosti z levé strany (humer.L) mužů aljašské populace z kmene Ipituaq pochází z normálního rozložení s parametry μ a σ^2, sestrojte (a) 99% levostranný interval spolehlivosti pro parametru μ; (b) 95% pravostranný interval spolehlivosti pro parametr σ^2; (c) 90% oboustranný interval spolehlivosti pro parametr σ.

Výsledky: Shapirův-Wilkův test: p-hodnota $= 0,2933$, $\alpha = 0,05$; data pochází z normálního rozložení; (a) $IS = (-\infty; 315.5)$; (b) $IS = (239.9; \infty)$; (c) $IS = (15.5; 29.4)$. ★

8.4 Testy o parametrech μ a σ^2 normálního rozložení

$X_1, \ldots, X_n$ je náhodný výběr z rozložení $N(\mu, \sigma^2)$, $n \geq 2$, c reálná konstanta, $\alpha \in (0,1)$ je hladina významnosti.

Test o střední hodnotě při známém rozptylu (jednovýběrový z-test):
Na hladině významnosti α testujeme H_0: $\mu = c$ oproti H_1: $\mu \neq c$, resp. proti H_1: $\mu < c$, resp. H_1: $\mu > c$. Vypočteme realizaci testové statistiky $t_0 = \frac{m-c}{\frac{\sigma}{\sqrt{n}}}$. 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 střední hodnotě při neznámém rozptylu (jednovýběrový t-test):
Formulace H_0 a H_1 je stejná jako u z-testu. Testová statistika má realizaci $t_0 = \frac{m-c}{\frac{s}{\sqrt{n}}}$. Kritické obory se od předešlé situace liší tím, že místo kvantilů $u_{1-\alpha/2}$, resp. $u_{1-\alpha}$ v nich figurují kvantily $t_{1-\alpha/2}(n-1)$, resp. $t_{1-\alpha}(n-1)$.

Test o rozptylu při neznámé střední hodnotě (test o rozptylu)
Na hladině významnosti α testujeme H_0: $\sigma^2 = c$ proti H_1: $\sigma^2 \neq c$, resp. proti H_1: $\sigma^2 < c$, resp. H_1: $\sigma^2 > c$. Vypočteme realizaci testové statistiky $t_0 = \frac{(n-1)s^2}{c}$. 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: $W = \langle 0; \chi^2_{\alpha/2}(n-1)\rangle \cup \langle \chi^2_{1-\alpha/2}(n-1); \infty)$, pro levostranný test $W = \langle 0; \chi^2_{\alpha}(n-1)\rangle$ a pro pravostranný test $W = \langle \chi^2_{1-\alpha}(n-1); \infty)$.

Párový t-test
Je-li $(X_1, Y_1)^T, \ldots, (X_n, Y_n)^T$ náhodný výběr z dvourozměrného rozložení, $n \geq 2$, pak zavedeme rozdílový náhodný výběr $Z_1 = X_1 - Y_1, \ldots, Z_n = X_n - Y_n$, o němž předpokládáme, že se řídí normálním rozložením. Tím jsme od dvourozměrného rozložení přešli k jednorozměrnému rozložení. Test H_0: $\mu_1 - \mu_2 = c$ proti H_1: $\mu_1 - \mu_2 \neq c$ se nazývá párový t-test. Označíme-li $\mu = \mu_1 - \mu_2$, pak párový t-test provedeme stejně jako jednovýběrový t-test.

Upozornění: Uvedené testy lze provést nejenom pomocí kritického oboru, ale též pomocí $100(1-\alpha)\%$ intervalu spolehlivosti nebo pomocí p-hodnoty.

Příklad 8.8. Řešený příklad
Načtěte datový soubor 11-two-samples-means-skull.txt obsahující údaje o výšce lebky (skull.H) mužů starověké egyptské populace. Dále mějme k dispozici údaje o výšce lebky mužů novověké egyptské populace ($\bar{x}_m = 133,977$ mm, $s_m = 5,171$ mm, $n_m = 87$). Za předpokladu, že výška lebky mužů starověké egyptské populace pochází z normálního rozložení, testujte hypotézu, že (a) střední hodnota výšky lebky mužů starověké a novověké egyptské populace jsou shodné; (b) rozptyl výšky lebky mužů u starověké populace je vyšší než u novověké populace; (c) směrodatná odchylka výšky lebky mužů u starověké populace je nižší než u novověké populace. Testování proveďte vždy (1) kritickým oborem; (2) intervalem spolehlivosti; (3) p-hodnotou. Hladinu významnosti zvolte $\alpha = 0,05$.

Řešení příkladu 8.8
Nejprve je třeba ověřit normální rozložení náhodného výběru výšek lebky mužů ze starověké populace. Na hladině významnosti $\alpha = 0,05$ testujeme H_0: *Data pochází z normálního rozložení.* oproti H_1: *Data nepochází z normálního rozložení.* K otestování normality použijeme vzhledem k rozsahu náhodného výběru ($n = 215$) Lillieforsův test.

```
data <- read.delim('11-two-samples-means-skull.txt')
skull.HM <- as.numeric(na.omit(data[data$sex == 'm', 'skull.H']))
nortest::lillie.test(skull.HM)$p.val # 0,1262537
```

Náhodný výběr výšek lebky mužů starověké egyptské populace pochází z normálního rozložení (p-hodnota $= 0,1263$).

(a) Na hladině významnosti $\alpha = 0,05$ testujeme H_0: $\mu = 133,977$ oproti H_1: $\mu \neq 133,977$ (oboustranný test). K testování použijeme jednovýběrový t-test implementovaný ve funkci t.test(). Výstupem funkce je realizace testové statistiky t_0, interval spolehlivosti a p-hodnota. Hranice kritického oboru dopočítáme příkazem qt().

```
alpha <- 0.05
n <- length(skull.HM) # 215
- qt(1 - alpha / 2, n - 1) # -1,971111
qt(1 - alpha / 2, n - 1) # 1,971111
t.test(skull.HM, mu = 133.977, alternative = 'two.sided', conf.level = 1 - alpha)
```

```

        One Sample t-test

data:  skull.HM
t = -6,3052, df = 214, p-value = 1,623e-09
alternative hypothesis: true mean is not equal to 133,977
95 percent confidence interval:
 131,2476 132,5477
sample estimates:
mean of x
 131,8977
```

Realizace testové statistiky $t_0 = -6,3052$, kritický obor $W = (-\infty; -1,9711\rangle \cup \langle 1,9711; \infty)$. Protože $t_0 \in W$, H_0 zamítáme na hladině významnosti $\alpha = 0,05$. Interval spolehlivosti $IS = (131,2; 132,5)$. Protože $c = 133,977 \notin IS$, 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 výškou lebky mužů starověké a novověké egyptské populace existuje statisticky významný rozdíl.

(b) Na hladině významnosti $\alpha = 0,05$ testujeme H_0: $\sigma^2 \geq 5,171^2$ oproti H_1: $\sigma^2 < 5,171^2$ (levostranný test). K testování použijeme test o rozptylu implementovaný ve funkci varTest() z knihovny EnvStats. Výstupem funkce je realizace testové statistiky t_0, interval spolehlivosti a p-hodnota. Horní hranici kritického oboru dopočítáme příkazem qchisq().

```
qchisq(alpha, n - 1) # 181,1454
EnvStats::varTest(skull.HM, sigma.squared = 5.171 ^ 2, alternative = 'less', conf.level
    = 0.95)
```

```

        Chi-Squared Test on Variance

data:  skull.HM
Chi-Squared = 187,13, df = 214, p-value = 0,0926
alternative hypothesis: true variance is less than 26,73924
95 percent confidence interval:
  0,00000 27,62284
sample estimates:
variance
  23,382
```

Realizace testové statistiky $t_0 = 187,1300$, kritický obor $W = \langle 0; 181,1454\rangle$. Protože $t_0 \notin W$, H_0 nezamítáme na hladině významnosti $\alpha = 0,05$. Interval spolehlivosti $IS = (0; 27,6228)$. Protože $c = 5,171^2 = 26,7392 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0,05$. Protože p-hodnota $= 0,0926$ je větší než $\alpha = 0,05$, H_0 nezamítáme na hladině významnosti $\alpha = 0,05$. Rozptyl výšky lebky u mužů starověké egyptské populace není statisticky významně nižší než u mužů novověké egyptské populace.

(c) Na hladině významnosti $\alpha = 0,05$ testujeme H_0: $\sigma \leq 5,171$ oproti H_1: $\sigma > 5,171$ (pravostranný test). K testování použijeme test o rozptylu implementovaný ve funkci varTest() z knihovny EnvStats. Výstupem funkce je realizace testové statistiky t_0, interval spolehlivosti a p-hodnota pro test o rozptylu. Dolní hranici kritického oboru dopočítáme příkazem qchisq().

```
qchisq(1 - alpha, n - 1) # 249,1275
EnvStats::varTest(skull.HM, sigma.squared = 5.171 ^ 2, alternative = 'greater', conf.
    level = 0.95)
```

```

        Chi-Squared Test on Variance

data:  skull.HM
Chi-Squared = 187,13, df = 214, p-value = 0,9074
alternative hypothesis: true variance is greater than 26,73924
95 percent confidence interval:
 20,08509      Inf
sample estimates:
variance
  23,382
```

Realizace testové statistiky $t_0 = 187,130$, kritický obor $W = \langle 249,1275; \infty)$. Protože $t_0 \notin W$, H_0 nezamítáme na hladině významnosti $\alpha = 0,05$. Interval spolehlivosti má po odmocnění tvar $IS = (4,4816; \infty)$. Protože $c = 5,171 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0,05$. Protože p-hodnota $= 0,9074$ je větší než $\alpha = 0,05$, H_0 nezamítáme na hladině významnosti $\alpha = 0,05$. Směrodatná odchylka výšky lebky u mužů starověké egyptské populace není statisticky významně vyšší než u mužů novověké egyptské populace.

★

Příklad 8.9. Řešený příklad (párový t-test)
Načtěte datový soubor 31-goldman-alaska.csv obsahující údaje o délce holenní kosti z pravé strany (tibia.R) a z levé strany (tibia.L) mužů aljašské populace z kmene Tigara. Za předpokladu, že rozdíl délek holenních kostí z pravé a z levé strany u těchto mužů pochází z normálního rozložení, testujte hypotézu, že mezi délkou holenní kosti z pravé a z levé strany neexistuje statisticky významný rozdíl. Testování proveďte vždy (1) kritickým oborem; (2) intervalem spolehlivosti; (3) p-hodnotou. Hladinu významnosti zvolte $\alpha = 0,01$.

Řešení příkladu 8.9
Nejprve je třeba ověřit normální rozložení náhodného výběru rozdílů délek holenních kostí u mužů z pravé a z levé strany. Na hladině významnosti $\alpha = 0,01$ testujeme H_0: *Rozdíly délek holenních kostí z pravé a z levé strany u mužů pochází z normálního rozložení.* oproti H_1: *Rozdíly délek holenních kostí z pravé a z levé strany u mužů nepochází z normálního rozložení.* K otestování normality použijeme vzhledem k rozsahu náhodného výběru ($n = 23$) Shapirův-Wilkův test.

```
data <- read.delim('31-goldman-alaska.csv', sep = ';', dec = '.')
data.MT <- na.omit(data[data$pop == 'Tigara' & data$sex == 'm',
                        c('tibia.R', 'tibia.L')])
tibia.RMT <- data.MT$tibia.R
tibia.LMT <- data.MT$tibia.L
diff <- tibia.RMT - tibia.LMT
shapiro.test(diff)$p.val # 0,7244359
```

Náhodný výběr rozdílů délek holenních kostí z pravé a z levé strany mužů aljašské populace z kmene Tigara pochází z normálního rozložení (p-hodnota $= 0,7244$).

Na hladině významnosti $\alpha = 0,01$ testujeme H_0: $\mu_1 - \mu_2 = 0$ oproti H_1: $\mu_1 - \mu_2 \neq 0$ (oboutranný test). K testování použijeme párový t-test implementovaný ve funkci t.test() s argumentem paired = T. Výstupem funkce je realizace testové statistiky t_0, interval spolehlivosti a p-hodnota. Hranice kritického oboru dopočítáme příkazem qt().

```
alpha <- 0.01
n <- length(diff) # 23
```

```
- qt(1 - alpha / 2, n - 1) # -2,818756
qt(1 - alpha / 2, n - 1) # 2,818756
t.test(tibia.MTR, tibia.MTL, paired = T, conf.level = 0.99)
```

```

        Paired t-test

data:  tibia.MTR and tibia.MTL
t = 0,66823, df = 22, p-value = 0,5109
alternative hypothesis: true difference in means is not equal to 0
99 percent confidence interval:
 -1,329275  2,155362
sample estimates:
mean of the differences
              0,4130435
```

Realizace testové statistiky $t_0 = 0,6682$, kritický obor $W = (-\infty; -2,8188\rangle \cup \langle 2,8188; \infty)$. Protože $t_0 \notin W$, H_0 nezamítáme na hladině významnosti $\alpha = 0,01$. Interval spolehlivosti $IS = (-1,3293; 2,1554)$. Protože $c = 0 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0,01$. Protože p-hodnota $= 0,5109$ je větší než $\alpha = 0,01$, H_0 nezamítáme na hladině významnosti $\alpha = 0,01$. Mezi mezi délkou holenní kosti z pravé a z levé strany mužů aljašské populace z kmene Tigara neexistuje statisticky významný rozdíl.

★

Příklad 8.10. Neřešený příklad
Načtěte datový soubor 01-one-sample-mean-skull-mf.txt obsahující údaje o největší délce mozkovny (skull.L) žen starověké egyptské populace. Dále mějme k dispozici údaje o největší délce mozkovny žen novověké egyptské populace ($\bar{x}_f = 171,962\,\mathrm{mm}$, $s_f = 7.052\,\mathrm{mm}$, $n_f = 52$). Za předpokladu, že největší délka mozkovny žen starověké egyptské populace pochází z normálního rozložení, testujte hypotézu, že (a) střední hodnota největší délky mozkovny u žen starověké populace je menší než u žen novověké populace; (b) rozptyl největší délky mozkovny žen starověké a novověké populace jsou shodné; (c) směrodatná odchylka největší délky mozkovny u žen starověké populace je vyšší než u žen novověké populace. Testování proveďte vždy (1) kritickým oborem; (2) intervalem spolehlivosti; (3) p-hodnotou. Hladinu významnosti zvolte $\alpha = 0,10$.
Výsledky: Lillieforsův test: p-hodnota $= 0,2625$, $\alpha = 0,10$; data pochází z normálního rozložení. (a) $t_0 = 4,3146$, $W = \langle 1,2894; \infty)$; IS $= (173,764; \infty)$, $c = 171,962$; p-hodnota $< 0,001$, $\alpha = 0,10$. H_0 zamítáme na hladině významnosti $\alpha = 0,10$. (b) $t_0 = 83,995$, $W = (-\infty; 85,0149\rangle \cup \langle 133,2569; \infty)$; IS $= (31,3465; 49,1342)$, $c = 7,052^2 = 49,7307$; p-hodnota $= 0,0842$, $\alpha = 0,10$. H_0 zamítáme na hladině významnosti $\alpha = 0,10$. (c) $t_0 = 83,995$, $W = (-\infty; 89,6451\rangle$; IS $= (0; 6,8262)$, $c = 7,052$; p-hodnota $= 0,0421$, $\alpha = 0,10$. H_0 zamítáme na hladině významnosti $\alpha = 0,10$. ★

Příklad 8.11. Neřešený příklad (párový t-test)
Načtěte datový soubor 31-goldman-alaska.csv obsahující údaje o délce pažní kosti z pravé strany (humer.R) a z levé strany (humer.L) žen aljašské populace z kmene Ipituaq. Za předpokladu, že rozdíl délek pažních kostí z levé a z pravé strany u těchto žen pochází z normálního rozložení, testujte hypotézu, že u žen aljašské populace z kmene Ipituaq je délka pažní kosti z levé strany větší než z pravé strany. Testování proveďte vždy (1) kritickým oborem; (2) intervalem spolehlivosti; (3) p-hodnotou. Hladinu významnosti zvolte $\alpha = 0,05$.
Výsledky: Shapirův-Wilkův test: p-hodnota $= 0,7244$, $\alpha = 0,05$; rozdíly pochází z normálního rozložení. $t_0 = -2,9498$, $W = (-\infty; -1,8331\rangle$; $IS = (-\infty; -1.4007)$, $c = 0$; p-hodnota $= 0,0081$, $\alpha = 0,05$. H_0 zamítáme na hladině významnosti $\alpha = 0,05$.

★