

9 Jednovýběrové neparametrické testy

9.1 Znaménkový jednovýběrový exaktní test

Nechť X_1, \dots, X_{1n} , $n \geq 2$ je náhodný výběr ze spojitého rozdělení a necht' \tilde{x}_0 je konstanta. Na hladině významnosti α testujeme jednu z následujících tří hypotéz oproti příslušné alternativní hypotéze.

$$\begin{array}{lll} H_{01} : \tilde{x} = \tilde{x}_0 & \text{oproti} & H_{11} : \tilde{x} \neq \tilde{x}_0 \quad (\text{oboustranná alt.}) \\ H_{02} : \tilde{x} \leq \tilde{x}_0 & \text{oproti} & H_{12} : \tilde{x} > \tilde{x}_0 \quad (\text{pravostranná alt.}) \\ H_{03} : \tilde{x} \geq \tilde{x}_0 & \text{oproti} & H_{13} : \tilde{x} < \tilde{x}_0 \quad (\text{levostranná alt.}) \end{array}$$

Test nazýváme jednovýběrový znaménkový test o mediánu \tilde{x} . Testovací statistika má tvar

$$S_E = \sum_{i=1}^n I(X_i - \tilde{x}_0 > 0), \quad (9.1)$$

$I(X_i - \tilde{x}_0 > 0)$ je indikační funkce, která nabývá hodnoty 1, pokud $X_i - \tilde{x}_0 > 0$, a hodnoty 0, pokud $X_i - \tilde{x}_0 \leq 0$. S_E udává tedy počet kladných rozdílů $X_i - \tilde{x}_0$. Za platnosti nulové hypotézy pochází statistika S_E z binomického rozdělení s parametry $N = m$, kde m je počet nenulových rozdílů $X_i - \tilde{x}_0$, a $p = \frac{1}{2}$, tj.

$$S_E = \sum_{i=1}^n I(X_i - \tilde{x}_0 > 0) \stackrel{H_0}{\sim} \text{Bin} \left(m, \frac{1}{2} \right).$$

Kritický obor podle zvolené alternativní hypotézy má tvar

$$\begin{array}{ll} H_{11} : \tilde{x} \neq \tilde{x}_0 & W = \left(-\infty; b_{m,1/2}(\alpha/2) - 1 \right) \cup \left(b_{m,1/2}(1 - \alpha/2); \infty \right) \\ H_{12} : \tilde{x} > \tilde{x}_0 & W = \left(b_{m,1/2}(1 - \alpha); \infty \right) \\ H_{13} : \tilde{x} < \tilde{x}_0 & W = \left(-\infty; b_{m,1/2}(\alpha) - 1 \right) \end{array}$$

kde m je počet nenulových rozdílů $X_i - \tilde{x}_0$, $b_{m,1/2}(\alpha/2)$, $b_{m,1/2}(1 - \alpha/2)$, $b_{m,1/2}(\alpha)$ a $b_{m,1/2}(1 - \alpha)$ jsou kvantily binomického rozdělení, jejichž hodnoty získáme pomocí \mathcal{R} a implementované funkce `qbinom()`.

Interval spolehlivosti má podle zvolené alternativní hypotézy jeden z následujících tvarů

$$\begin{array}{ll} H_{11} : \tilde{x} \neq \tilde{x}_0 & (d, h) = \left(X^{(b_{n,1/2}(\alpha/2))}; X^{(b_{n,1/2}(1-\alpha/2)+1)} \right) \\ H_{12} : \tilde{x} > \tilde{x}_0 & (d, \infty) = \left(X^{(b_{n,1/2}(\alpha))}; \infty \right) \\ H_{13} : \tilde{x} < \tilde{x}_0 & (-\infty, h) = \left(-\infty; X^{(b_{n,1/2}(1-\alpha)+1)} \right) \end{array}$$

kde n je rozsah náhodného výběru, $X^{(1)} \leq \dots \leq X^{(n)}$ značí vzestupně seřazené hodnoty X_i , $i = 1, \dots, n$, a $X^{(k)}$ značí k -tou hodnotu v seřazené posloupnosti $X^{(1)} \leq \dots \leq X^{(n)}$.

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

$$\begin{array}{ll} H_{11} : \tilde{x} \neq \tilde{x}_0 & p\text{-hodnota} = 2 \min\{\Pr(S_E \leq s_E), \Pr(S_E \geq s_E)\} \\ H_{12} : \tilde{x} > \tilde{x}_0 & p\text{-hodnota} = \Pr(S_E \geq s_E) \\ H_{13} : \tilde{x} < \tilde{x}_0 & p\text{-hodnota} = \Pr(S_E \leq s_E) \end{array}$$

kde S_E je náhodná veličina pocházející z binomického rozdělení, s_E je realizace testovací statistiky S_E (viz vzorec 9.1), tedy konkrétní číslo, $\Pr(S_E \geq s_E) = 1 - \Pr(S_E < s_E) = 1 - \Pr(S_E \leq s_E - 1)$, což vyplývá z faktu, že náhodná veličina S_E pochází z binomického (diskrétního) rozdělení (viz kapitola ??), a $\Pr(S_E \leq s_E)$, resp. $\Pr(S_E \leq s_E - 1)$ je distribuční funkce binomického rozdělení, jejíž hodnotu získáme pomocí \mathcal{R} a implementované funkce `pbinom()`.

Poznámka: Všimněme si, že ve vzorcích intervalu spolehlivosti figuruje rozsah náhodného výběru n , zatímco ve vzorcích hranic kritického oboru a u p -hodnoty pracujeme s počtem nenulových rozdílů m .

Příklad 9.1. Znaménkový jednovýběrový exaktní test

Mějme datový soubor `21-goldman-measures.csv` a proměnnou `humer.DR` popisující největší délku hlavice pažní kosti z pravé strany v mm u skeletů z období raného středověku z oblasti *Staubing* (viz sekce ??). Dále máme k dispozici údaje ze studie (Mall et al.) z roku 1999, v rámci které byla měřena největší délka hlavice pažní kosti u mužů současné německé populace ($m_m = 50.0$ mm, $n_m = 64$). Na hladině významnosti $\alpha = 0.05$ zjistíte, zda existuje rozdíl mezi největší délkou hlavice pažní kosti z pravé strany u mužů z raně středověké německé populace a u mužů současné německé populace.

Řešení příkladu 9.1

Pomocí příkazu `read.delim()` načteme datový soubor a pomocí operátoru `[]` vybereme z datové tabulky údaje o největší délce hlavice pažní kosti z pravé strany (`humer.DR`) u mužů (`sex == 'm'`) z oblasti *Staubing* (`pop == 'Staubing'`). Z vektoru naměřených hodnot odstraníme chybějící údaje (`na.omit()`) zjistíme rozsah náhodného výběru (`length()`).

```
1 data <- read.delim('00-Data//21-goldman-measures.csv', sep = ';')
2 head(data)
3 humer.DRM <- data[data$pop == 'Staubing' & data$sex == 'm', 'humer.DR']
4 humer.DRM <- na.omit(humer.DRM)
5 n <- length(humer.DRM) # 12
```

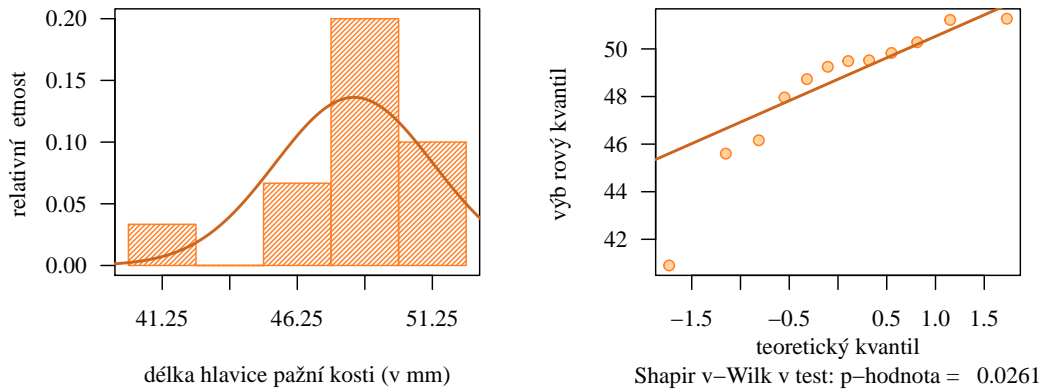
Datový soubor obsahuje údaje o největší délce hlavice pažní kosti z pravé strany u 12 mužů raně středověké německé populace.

Naším úkolem ze zadání je porovnat střední hodnoty dvou německých populací, přičemž u jedné populace (německé raně středověké populace) máme k dispozici naměřené hodnoty. Na základě těchto hodnot můžeme zjistit, zda náhodný výběr pochází z normálního rozdělení, tj. zda náhodná veličina X popisující největší délku hlavice pažní kosti u mužů raně středověké populace pochází z normálního rozdělení, tj. $X \sim N(\mu, \sigma^2)$, kde skutečný rozptyl σ^2 neznáme. Druhá populace (současná německá populace) je reprezentována pouze hodnotou aritmetického průměru ($m_R = 50.00$ mm). Řešení příkladu tedy vede na situaci, kdy střední hodnotu jednoho náhodného výběru (jehož skutečnou hodnotu rozptylu neznáme) porovnáváme s konkrétním číslem, tedy na jednovýběrový test o střední hodnotě μ při neznámém rozptylu σ^2 (viz kapitola ??). Jediným předpokladem k použití tohoto testu je normalita náhodného výběru naměřených délek hlavic pažních kostí. Před použitím testu tedy tento předpoklad ověříme.

Jelikož je rozsah náhodného výběru menší než 30, ověříme předpoklad normality Shapirovým-Wilkovým testem ($\alpha = 0.05$). Grafické ověření provedeme na základě QQ-diagramu a histogramu superponovaného křivkou normálního rozdělení, jejíž parametry odhadneme pomocí výběrového průměru a výběrového rozptylu (viz obrázek 1). Datový soubor rozdělíme do pěti ekvidistatních intervalů s šířkou 3 mm prostřednictvím stanovených hranic 37, 40, ..., 52.

Protože p -hodnota = 0.0261 je menší než 0.05, nulovou hypotézu o normalitě dat zamítáme na hladině významnosti $\alpha = 0.05$. Z histogramu na obrázku 1 vidíme, že naměřené hodnoty jsou vyšikmené doprava s chvostem na levé straně a tvar křivky hustoty normálního rozdělení příliš dobře nekopírují. Na QQ-diagramu je zřejmé odchýlení bodů od referenční křivky zejména na levém chvostu. Náhodný výběr největších délek hlavice pažní kosti mužů z raně středověké německé populace tedy nepochází z normálního rozdělení.

Protože náhodný výběr nepochází z normálního rozdělení, nemůžeme hypotézu ze zadání otestovat pomocí parametrického testu o střední hodnotě μ podle metod uvedených v sekci ??). K testování hypotézy musíme použít neparametrickou alternativu tohoto testu. Zde konkrétně použijeme znaménkový jednovýběrový test. Jelikož rozsah náhodného výběru je menší než 30, použijeme exaktní variantu znaménkového testu. Naším úkolem je zjistit, zda existuje rozdíl mezi největší délkou hlavice pažní kosti z pravé strany u mužů z raně středověké německé populace a u mužů současné německé populace. Tato věta bude součástí alternativní hypotézy, neboť rozdíl implikuje nerovnost a nerovnost je vždy součástí alternativní hypotézy. Nulovou hypotézu potom stanovíme jako doplněk k alternativní hypotéze. Závěrem poznamenejme, že zatímco u parametrických testů (viz kapitola ??) figuruje v hypotézách pojem *střední hodnota*, u neparametrických testů tento pojem nahrazujeme neparametrických ekvivalentem, kterým



Obrázek 1: Histogram a QQ-diagram největší délký hlavičky pažní kosti na pravé straně u skeletů mužů z raně středověké populace z oblasti Staubing

je *medián*. Testování provedeme v posloupnosti sedmi kroků.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : Medián největší délký hlavičky pažní kosti raně středověké německé populace je shodný s mediánem největší délký hlavičky pažní kosti na pravé straně mužů současné německé populace.

H_1 : Medián největší délký hlavičky pažní kosti raně středověké německé populace není shodný s mediánem největší délký hlavičky pažní kosti na pravé straně mužů současné německé populace.

- **matematická formulace** nulové a alternativní hypotézy

H_0 : $\tilde{x} = \tilde{x}_0$, kde $\tilde{x}_0 = 50.00$

H_1 : $\tilde{x} \neq \tilde{x}_0$, kde $\tilde{x}_0 = 50.00$ (oboustranná alternativa)

2. Volba hladiny významnosti

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

3. Testování kritickým oborem

- **Testovací statistika**

Nejprve vypočítáme vektor rozdílů naměřených hodnot X_i a konstanty \tilde{x}_0 , tj. $X_i - \tilde{x}_0$, a stanovíme, které rozdílů jsou kladné. (viz tabulka 1).

Tabulka 1: Naměřené hodnoty X_i , rozdíly $X_i - \tilde{x}_0$ a znaménka těchto rozdílů

měření	1	2	3	4	5	6	7	8	9	10	11	12
X_i	49.83	47.96	49.49	48.73	49.25	40.90	51.22	51.27	46.16	49.52	50.28	45.60
$X_i - \tilde{x}_0$	-0.17	-2.04	-0.51	-1.27	-0.75	-9.10	1.22	1.27	-3.84	-0.48	0.28	-4.40
+/-	-	-	-	-	-	-	+	+	-	-	+	-

Z tabulky 1 vidíme, že celkem tři rozdíly $X_i - \tilde{x}_0$, $i = 1, \dots, 12$ jsou kladné. Hodnota testovací statistiky S_E , která je definovaná jako počet kladných rozdílů, bude tedy rovná 3.

$$S_E = \sum_{i=1}^n I(X_i - \tilde{x}_0 > 0) = \sum_{i=1}^{13} I(X_i - 50 > 0) = 3.$$

```

6 x0 <- 50.00
7 I <- (humer.DRM > x0)
8 tab <- data.frame(rbind("Xi" = humer.DRM, "Xi-x0" = humer.DRM - x0, "+/-" = I))
9 names(tab) <- 1 : 12
10 #
11 # Xi      49.83 47.96 49.49 48.73 49.25 40.9 51.22 51.27 46.16 49.52 50.28 45.6
12 # Xi-x0  -0.17 -2.04 -0.51 -1.27 -0.75 -9.1  1.22  1.27 -3.84 -0.48  0.28  -4.4
13 # +/-    0.00  0.00  0.00  0.00  0.00  0.0  1.00  1.00  0.00  0.00  1.00  0.0
14 SE <- sum(I) # 3

```

- **Kritický obor**

Z tabulky 1 dále vidíme, že žádný z celkového počtu $n = 12$ rozdílů není nulový, a tedy počet nenulových rozdílů $m = 12$. Počet nenulových rozdílů využíváme při stanovení hranic kritického oboru a později také u výpočtu p -hodnoty.

$$\begin{aligned}
W &= (-\infty; b_{m,1/2}(\alpha/2) - 1) \cup (b_{m,1/2}(1 - \alpha/2); \infty) \\
&= (-\infty; b_{12,1/2}(0.05/2) - 1) \cup (b_{12,1/2}(1 - 0.05/2); \infty) \\
&= (-\infty; b_{12,1/2}(0.025) - 1) \cup (b_{12,1/2}(0.975); \infty) \\
&= (-\infty; 3 - 1) \cup (9; \infty) \\
&= (-\infty; 2) \cup (9; \infty)
\end{aligned}$$

```

15 m <- sum(humer.DRM - x0 != 0) # 12
16 alpha <- 0.05
17 q1 <- qbinom(alpha / 2, m, 1 / 2) - 1 # 2
18 q2 <- qbinom(1 - alpha / 2, m, 1 / 2) # 9

```

- **Závěr testování**

Protože realizace testovací statistiky $s_E = 3$ nenáleží do kritického oboru, tj. $s_E \notin W$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

4. Testování intervalem spolehlivosti

- **Interval spolehlivosti**

Abychom mohli stanovit hranice 95 % intervalu spolehlivosti, musíme nejprve naměřené hodnoty největší délky hlavice pažní kosti vzestupně seřadit. To provedeme jednoduše příkazem `sort()`. Seřazené hodnoty jsou k dispozici v tabulce 2.

Tabulka 2: Seřazené hodnoty největší délky hlavice pažní kosti

pořadí	1	2	3	4	5	6	7	8	9	10	11	13
seřazené X_i	40.90	45.60	46.16	47.96	48.73	49.25	49.49	49.52	49.83	50.28	51.22	51.27

Hranice intervalu spolehlivosti potom tvoří ty hodnoty, které se v seřazeném vektoru hodnot nachází na $(b_{n,1/2}(\alpha/2))$ -té pozici a na $(b_{n,1/2}(1 - \alpha/2) + 1)$ -té pozici. Hodnoty $b_{n,1/2}(\alpha/2)$ a $b_{n,1/2}(1 - \alpha/2) + 1$ nalezneme opět pomocí funkce `qbinom()`.

$$\begin{aligned}
(d, h) &= (X^{(b_{n,1/2}(\alpha/2))}; X^{(b_{n,1/2}(1-\alpha/2)+1)}) \\
&= (X^{(b_{12,1/2}(0.05/2))}; X^{(b_{12,1/2}(1-0.05/2)+1)}) \\
&= (X^{(3)}; X^{(10)}) \\
&= (46.16; 50.28)
\end{aligned}$$

```

19 qbinom(alpha / 2, n, 1 / 2) # 3
20 qbinom(1 - alpha / 2, n, 1 / 2) + 1 # 10
21 humer.DRMs[3] # 46.16
22 humer.DRMs[10] # 50.28

```

- **Závěr testování**

Protože $\tilde{x}_0 = 50.00$ náleží do 95% empirického oboustranného intervalu spolehlivosti, tj. $\tilde{x}_0 = 50.00 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

5. Testování p -hodnotou

- **p -hodnota**

$$\begin{aligned}
 p\text{-hodnota} &= 2 \min\{\Pr(S_E \leq s_E), \Pr(S_E \geq s_E)\} \\
 &= 2 \min\{\Pr(S_E \leq 3), \Pr(S_E \geq 3)\} \\
 &= 2 \min\{\Pr(S_E \leq 3), 1 - \Pr(S_E \leq 2)\} \\
 &= 2 \min\{0.07299805, 0.9807129\} \\
 &= 2 \times 0.07299805 = 0.1459961 \doteq 0.1460
 \end{aligned}$$

```

23 2 * min (pbinom(SE, m, 1 / 2), 1 - pbinom(SE - 1, m, 1 / 2)) # 0.1459961

```

- **Závěr testování**

Protože p -hodnota = 0.1460 je větší než $\alpha = 0.05$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

6. Interpretace výsledků

Za základě všech tří typů testování nezamítáme nulovou hypotézu na hladině významnosti $\alpha = 0.05$. Mezi největší délkou hlavice pažní kosti na pravé straně u mužů raně středověké a současné populace neexistuje statisticky významný rozdíl.

7. Grafická vizualizace výsledku testování

Porovnání náhodného výběru s konstantou $\tilde{x}_0 = 50.00$ zobrazíme nejlépe pomocí krabicového diagramu (viz obrázek 2).

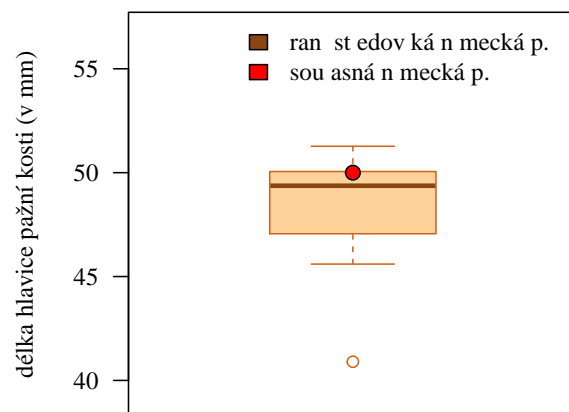
Poznámka: Znaménkový jednovýběrový exaktní test můžeme provést pomocí funkce `SIGN.test()` implementované v knihovně `BSDA`. Vstupními parametry budou vektor reprezentující náhodný výběr (`humer.DRM`), hodnota parametru \tilde{x}_0 z nulové hypotézy zadaná argumentem `md = 50.00`, hodnota hladiny významnosti α zadaná prostřednictvím koeficientu spolehlivosti $1 - \alpha$ nastavením hodnoty argumentu `conf.level = 0.95` a typ zvolené alternativní hypotézy (oboustranná) zadaný pomocí argumentu `alternative = 'two.sided'`. Součástí výstupu je hodnota mediánu náhodného

```

24 BSDA::SIGN.test(humer.DRM, md = 50.0, conf.level = 0.95,
25                 alternative = 'two.sided') # IS interpolovany

```

výběru `median of x = 49.37`, hodnota testovací statistiky `s = 3`, interpolované hranice 95% Waldova empirického oboustranného intervalu spolehlivosti 46.35145 a 50.23214, které jsou mírně přesnější, než námi stanovené hranice intervalu spolehlivosti (zpřesnění hranic bylo provedeno procesem nazývaným *interpolace*), a p -hodnota `p-value = 0.146`. Jediné, co musíme stanovit zvlášť, jsou dolní a horní hranice kritického oboru. ★



Obrázek 2: Krabicový diagram největší délky hlavičky pažní kosti na pravé straně u skeletů mužů z raně středověké populace z oblasti Staubing

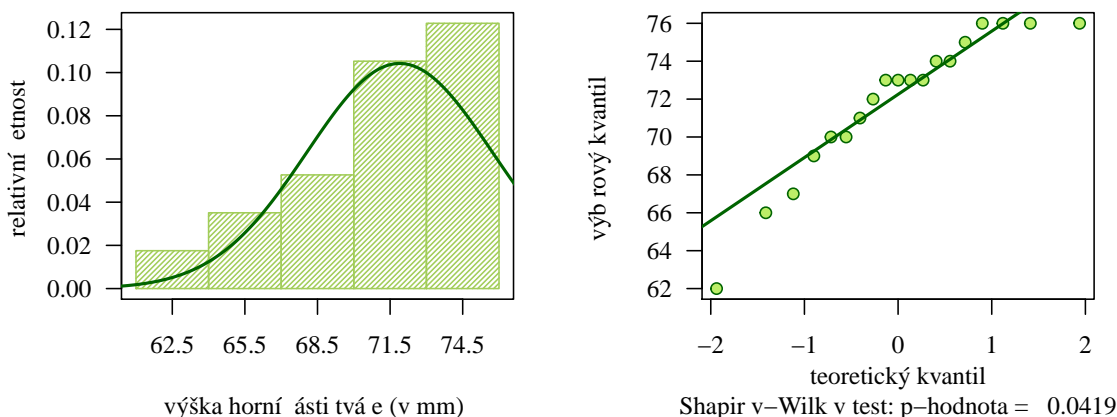
One-sample Sign-Test			
data:	humer.DRM		
s =	3	p-value =	0.146
alternative hypothesis:	true median is not equal to 50		
95 percent confidence interval:	46.35145	50.23214	
sample estimates:	median of x		
	49.37		
Achieved and Interpolated Confidence Intervals:			
	Conf.Level	L.E.pt	U.E.pt
Lower Achieved CI	0.8540	47.9600	49.8300
Interpolated CI	0.9500	46.3515	50.2321
Upper Achieved CI	0.9614	46.1600	50.2800

Příklad 9.2. Znaménkový jednovýběrový exaktní test

Mějme datový soubor 15-anova-means-skull.txt a proměnnou upface.H popisující výšku horní části tváře mužů německé populace (viz sekce ??). Dále máme k dispozici údaje o výšce horní části tváře mužů Černjachovské kultury na území dnešní Ukrajiny ($m_{ck} = 70.00$ mm, $n_{ck} = 99$). Na hladině významnosti $\alpha = 0.10$ testujte hypotézu, že výška horní části tváře německé mužské populace je menší nebo rovna výšce horní části tváře mužské populace Černjachovské kultury.

Řešení příkladu 9.2

Úvodu tohoto příkladu jsme se věnovali v příkladu ?? v rámci sekce ?. Zde jsme za základě testování normality pomocí Shapirova-Wilkova testu (p -hodnota = 0.0419) a grafické vizualizace (kombinace histogramu a QQ-diagramu) dospěli k závěru, že náhodný výběr výšek horní části tváře u 19 mužů německé populace nepochází z normálního rozdělení (pro připomenutí viz obrázek 3).



Obrázek 3: Histogram a QQ-diagram výšky horní části tváře u mužů německé populace

Jelikož data nepochází z normálního rozdělení, není možné otestovat hypotézu ze zadání pomocí parametrického testu o střední hodnotě μ , ale bude potřeba použít neparametrickou alternativu tohoto testu. Vzhledem k nízkému rozsahu náhodného výběru použijeme k otestování hypotézy ze zadání exaktní znaménkový jednovýběrový test. Naším úkolem je testovat (nulovou) hypotézu, že výška horní části tváře německé mužské populace je menší nebo rovna výšce horní části tváře mužské populace Černjachovské kultury. Alternativní hypotézu potom stanovíme jako doplněk k nulové hypotéze. Protože budeme hypotézu ze zadání testovat pomocí neparametrického testu, bude v jejím přesném znění figurovat pojem medián namísto pojmu střední hodnota.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : Medián výšky horní části tváře německé mužské populace je menší nebo roven mediánu výšky horní části tváře mužské populace z Černjachovské kultury.

H_1 : Medián výšky horní části tváře německé mužské populace je větší než medián výšky horní části tváře mužské populace z Černjachovské kultury.

- **matematická formulace** nulové a alternativní hypotézy

H_0 : $\tilde{x} \leq \tilde{x}_0$, kde $\tilde{x}_0 = 70.00$

H_1 : $\tilde{x} > \tilde{x}_0$, kde $\tilde{x}_0 = 70.00$ (pravostranná alternativa)

2. Volba hladiny významnosti

- Hladinu významnosti volíme v souladu se zadáním jako $\alpha = 0.10$.

3. Testování kritickým oborem

- **Testovací statistika**

Nejprve vypočítáme vektor rozdílů naměřených hodnot X_i a konstanty \tilde{x}_0 , tj. $X_i - \tilde{x}_0$, a stanovíme, které rozdíly jsou kladné (viz tabulka 3).

Tabulka 3: Naměřené hodnoty X_i , rozdíly $X_i - \tilde{x}_0$ a znaménka těchto rozdílů

měření	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
X_i	73	73	67	75	70	62	76	73	71	66	76	73	74	74	70	76	72	69	76
$X_i - \tilde{x}_0$	3	3	-3	5	0	-8	6	3	1	-4	6	3	4	4	0	6	2	-1	6
+/-	+	+	-	+	0	-	+	+	+	-	+	+	+	+	0	+	+	-	+

Z tabulky 3 vidíme, že celkem 13 rozdílů $X_i - \tilde{x}_0$, $i = 1, \dots, 19$ je kladných. Hodnota testovací statistiky S_E , která je definovaná jako počet kladných rozdílů, bude tedy rovná 13.

$$S_E = \sum_{i=1}^n I(X_i - \tilde{x}_0 > 0) = \sum_{i=1}^{19} I(X_i - 70.00 > 0) = 13.$$

```

44 x0 <- 70.00
45 I <- (upface.HN > x0)
46 tab <- data.frame(rbind("Xi" = upface.HN, "Xi-x0" = upface.HN - x0, "+/-" = I))
47 names(tab) <- 1 : 19
48 #      1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19
49 # Xi    73 73 67 75 70 62 76 73 71 66 76 73 74 74 70 76 72 69 76
50 # Xi-x0  3  3 -3  5  0 -8  6  3  1 -4  6  3  4  4  0  6  2 -1  6
51 # +/-   1  1  0  1  0  0  1  1  1  0  1  1  1  1  0  1  1  0  1
52 SE <- sum(I) # 13

```

- **Kritický obor**

Z tabulky 3 dále vidíme, že z celkového počtu $n = 19$ rozdílů jsou dva rozdíly nulové, a tedy počet nenulových rozdílů $m = 17$. Počet nenulových rozdílů využijeme při stanovení hranic kritického oboru a u výpočtu p -hodnoty.

$$\begin{aligned}
 W &= \langle b_{m,1/2}(\alpha); \infty \rangle \\
 &= \langle b_{17,1/2}(0.10); \infty \rangle \\
 &= \langle 10; \infty \rangle
 \end{aligned}$$

```

53 m <- sum(upface.HN - x0 != 0) # 17
54 alpha <- 0.10
55 q <- qbinom(1 - alpha, m, 1 / 2) - 1 # 10

```

- **Závěr testování**

Protože realizace testovací statistiky $s_E = 13$ náleží do kritického oboru, tj. $s_E \in W$, H_0 zamítáme na hladině významnosti $\alpha = 0.10$.

4. Testování intervalem spolehlivosti

- **Interval spolehlivosti**

Abychom mohli stanovit hranice 90% intervalu spolehlivosti, musíme nejprve naměřené hodnoty výšky horní části lebky vzestupně seřadit. To provedeme příkazem `sort()`. Seřazené hodnoty jsou k dispozici v tabulce 4.

Tabulka 4: Seřazené hodnoty naměřených výšek horní části tváře

pořadí	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
seřazené X_i	62	66	67	69	70	70	71	72	73	73	73	73	74	74	75	76	76	76	76

Hranice intervalu spolehlivosti potom tvoří hodnota, která se v seřazeném vektoru hodnot nachází na $(b_{n,1/2}(\alpha))$ -té pozici a nekonečno.

$$\begin{aligned}
 (d, \infty) &= \left(X^{(b_{n,1/2}(\alpha))}; \infty \right) \\
 &= \left(X^{(b_{19,1/2}(0.10))}; \infty \right) \\
 &= \left(X^{(7)}; \infty \right) \\
 &= (71; \infty)
 \end{aligned}$$

```
56 qbinom(alpha, n, 1 / 2) # 7
57 upface.HNs [7] # 71
```

- **Závěr testování**

Protože $\tilde{x}_0 = 70.00$ nenáleží do 90% empirického levostranného intervalu spolehlivosti, tj. $\tilde{x}_0 = 70.00 \notin IS$, H_0 zamítáme na hladině významnosti $\alpha = 0.10$.

5. Testování p -hodnotou

- **p -hodnota**

$$p\text{-hodnota} = \Pr(S_E \geq s_E) = \Pr(S_E \geq 13) = 1 - \Pr(S_E \leq 12) = 1 - 0.9754791 = 0.0245209 \doteq 0.02452$$

```
58 1 - pbinom(SE - 1, m, 1 / 2) # 0.02452087
```

- **Závěr testování**

Protože p -hodnota = 0.02452 je menší než $\alpha = 0.10$, H_0 zamítáme na hladině významnosti $\alpha = 0.10$.

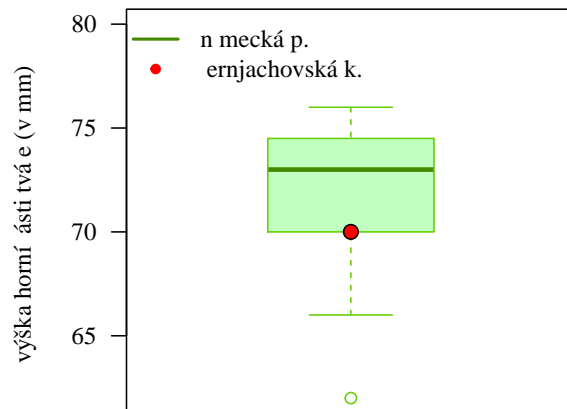
6. Interpretace výsledků

Za základě všech tří typů testování zamítáme nulovou hypotézu na hladině významnosti $\alpha = 0.10$. Výška horní části tváře mužů německé populace je statisticky významně větší než výška horní části tváře mužů Černjachovské populace z oblasti dnešní Ukrajiny.

7. Grafická vizualizace výsledku testování

Významný rozdíl horní části tváře mužů mezi oběma populacemi vizualizujeme krabicovým diagramem (viz obrázek 4).

Poznámka: K provedení znaménkového jednovýběrového exaktního testu můžeme použít funkci `SIGN.test()` z knihovny BSDA. Vstupními parametry budou vektor reprezentující náhodný výběr (`upface.HN`), hodnota parametru \tilde{x}_0 z nulové hypotézy (`md = 70.00`), hodnota hladiny významnosti α zadaná prostřednictvím koeficientu spolehlivosti $1 - \alpha$ (`conf.level = 0.90`) a typ zvolené alternativní hypotézy (pravostranná, `alternative = 'greater'`). Součástí výstupu je hodnota mediánu náhodného výběru `median of x = 73`, hodnota testovací statistiky `s = 13`, interpolované hranice 90% empirického levostranného intervalu spolehlivosti `71.17133` a `lnf` a p -hodnota `p-value = 0.02452`. Jediné, co musíme stanovit zvlášť, je dolní hranice kritického oboru. ★



Obrázek 4: Krabicový diagram výšky horní části tváře u mužů německé populace

```
59 BSDA::SIGN.test(upface.HN, md = 70.00, conf.level = 0.90, alt = 'greater') # IS
    interpolovany
```

```

    One-sample Sign-Test
data: upface.HN
s = 13, p-value = 0.02452
alternative hypothesis: true median is greater than 70
90 percent confidence interval:
 71.17133      Inf
sample estimates:
median of x
      73

Achieved and Interpolated Confidence Intervals:

      Conf.Level  L.E.pt  U.E.pt
Lower Achieved CI  0.8204 72.0000  Inf
Interpolated CI   0.9000 71.1713  Inf
Upper Achieved CI  0.9165 71.0000  Inf
```

60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77

Příklad 9.3. Znaménkový jednovýběrový exaktní test

Mějme datový soubor 21-goldman-measures.csv a proměnnou tibia.LR popisující největší délku lýtkové kosti z pravé strany v mm u skeletů z období neolitu z oblasti *Yoshigo* (viz sekce ??). Dále máme k dispozici údaje ze studie (Hasegawa et al.) z roku 2009, v rámci které byla měřena největší délka lýtkové kosti z pravé strany u žen současné japonské populace ($m_f = 329.40$ mm, $s_f = 17.3$ mm, $n_f = 342$). Na hladině významnosti $\alpha = 0.01$ zjistěte, zda je největší délka lýtkové kosti z pravé strany u žen z neolitické japonské populace významně menší než u žen současné japonské populace.

Řešení příkladu 9.3

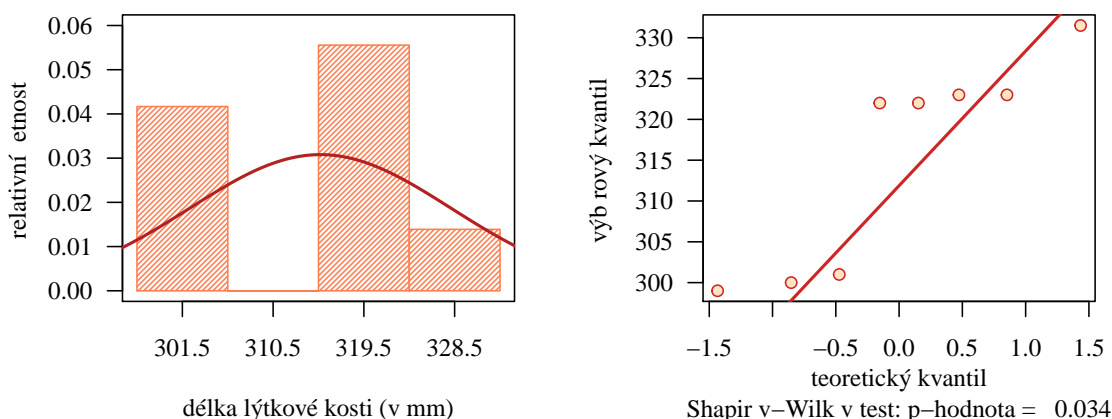
Načteme datový soubor a pomocí operátoru `[]` vybereme z datové tabulky údaje o největší délce lýtkové kosti z pravé strany (tibia.LR) u žen (`sex == 'f'`) z oblasti Yoshigo (`pop == 'Yoshigo Shell Mound'`). Z vektoru naměřených hodnot odstraníme chybějící údaje a zjistíme rozsah náhodného výběru.

```
78 data <- read.delim('00-Data//21-goldman-measures.csv', sep = ';')
79 # head(data)
80 tibia.LRF <- data[data$pop == 'Yoshigo Shell Mound' & data$sex == 'f', 'tibia.LR']
81 tibia.LRF <- na.omit(tibia.LRF)
82 n <- length(tibia.LRF) # 8
```

Datový soubor obsahuje údaje o největší délce lýtkové kosti z pravé strany u 8 žen z neolitické japonské populace.

V příkladu se zaměříme na porovnání délky lýtkové kosti dvou japonských populací, přičemž u jedné populace (neolitická populace) máme k dispozici naměřené hodnoty. Na základě těchto hodnot můžeme zjistit, zda náhodná veličina X popisující největší délku lýtkové kosti u žen neolitické japonské populace pochází z normálního rozdělení, tj. $X \sim N(\mu, \sigma^2)$, kde skutečný rozptyl σ^2 neznáme. Druhá populace (současná japonská populace) je reprezentována pouze hodnotou aritmetického průměru ($m_f = 329.40$ mm) a směrodatné odchylky ($s_f = 17.3$ mm). Řešení příkladu vede na situaci, kdy střední hodnotu jednoho náhodného výběru porovnáváme s konkrétním číslem, tedy na jednovýběrový test o střední hodnotě μ při neznámém rozptylu σ^2 . Před použitím parametrického testu je třeba ověřit normalitu náhodného výběru naměřených délek lýtkových kostí.

Předpoklad normality ověříme Shapirovým-Wilkovým testem ($\alpha = 0.05$), QQ-diagramem a histogramem (viz obrázek 5). Datový soubor rozdělíme do čtyř ekvidistatních intervalů s šířkou 9 mm prostřednictvím stanovených hranic 297, 306, ..., 333.



Obrázek 5: Histogram a QQ-diagram délky lýtkové kosti na pravé straně u skeletů žen neolitické populace z oblasti Yoshigo Shell Mound

Protože p -hodnota = 0.034 je menší než 0.05, nulovou hypotézu o normalitě dat zamítáme na hladině významnosti

$\alpha = 0.05$. Z histogramu na obrázku 5 vidíme, že naměřené hodnoty charakterem normálního rozdělení příliš nedisponují. Navíc body v QQ-diagramu se realizují jve shlucích, kde se střídavě vzdalují a přibližují k referenční přímce. Náhodný výběr největších délek lýtkových kostí žen z neolitické japonské populace nepochází z normálního rozdělení.

Protože náhodný výběr nepochází z normálního rozdělení, nemůžeme k testování použít parametrický test o střední hodnotě μ . Testování tedy provedeme na základě exaktního neparametrického znaménkového jednovýběrového testu. Naším úkolem je zjistit, zda je největší délka lýtkové kosti z pravé strany u žen z neolitické japonské populace významně menší než u žen současné japonské populace. Tato věta bude zněním alternativní hypotézy, neboť v zadání se zaměřujeme na nerovnost, přičemž nikde není zmínka o znění nulové hypotézy. Nulovou hypotézu stanovíme následně jako doplněk k alternativní hypotéze.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : Medián největší délky lýtkové kosti z pravé strany u žen z neolitické japonské populace je větší nebo roven mediánu největší délky lýtkové kosti z pravé strany u žen současné japonské populace.

H_1 : Medián největší délky lýtkové kosti z pravé strany u žen z neolitické japonské populace je menší než medián největší délky lýtkové kosti z pravé strany u žen současné japonské populace.

- **matematická formulace** nulové a alternativní hypotézy

$H_0 : \tilde{x} \geq \tilde{x}_0$, kde $\tilde{x}_0 = 329.40$

$H_1 : \tilde{x} < \tilde{x}_0$, kde $\tilde{x}_0 = 329.40$ (levostranná alternativa)

2. Volba hladiny významnosti

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

3. Testování kritickým oborem

- **Testovací statistika**

Nejprve vypočítáme vektor rozdílů naměřených hodnot X_i a konstanty \tilde{x}_0 , tj. $X_i - \tilde{x}_0$ a stanovíme znaménka těchto rozdílů (viz tabulka 6).

Tabulka 5: Naměřené hodnoty X_i , rozdíly $X_i - \tilde{x}_0$ a znaménka těchto rozdílů

měření	1	2	3	4	5	6	7	8
X_i	300.0	299.0	323.0	323.0	322.0	301.0	331.5	322.0
$X_i - \tilde{x}_0$	-29.4	-30.4	-6.4	-6.4	-7.4	-28.4	2.1	-7.4
+/-	-	-	-	-	-	-	+	-

Z tabulky 6 vidíme, že pouze jeden rozdíl $X_i - \tilde{x}_0$, $i = 1, \dots, 8$ je kladný. Hodnota testovací statistiky S_E bude tedy rovná 1.

$$S_E = \sum_{i=1}^n I(X_i - \tilde{x}_0 > 0) = \sum_{i=1}^8 I(X_i - 329.40 > 0) = 1.$$

```

83 x0 <- 329.40
84 I <- (tibia.LRF > x0)
85 tab <- data.frame(rbind("Xi" = tibia.LRF, "Xi-x0" = tibia.LRF - x0, "+/-" = I))
86 names(tab) <- 1 : 8
87 #
88 # Xi      300.0 299.0 323.0 323.0 322.0 301.0 331.5 322.0
89 # Xi-x0  -29.4 -30.4 -6.4  -6.4  -7.4 -28.4  2.1  -7.4
90 # +/-    0.0  0.0  0.0  0.0  0.0  0.0  1.0  0.0
91 SE <- sum(I) # 1

```

- **Kritický obor**

Z tabulky 6 dále vidíme, že žádný z celkového počtu $n = 8$ rozdílů není nulový, a tedy počet nenulových rozdílů $m = 8$.

$$\begin{aligned} W &= (-\infty; b_{m,1/2}(\alpha) - 1) \\ &= (-\infty; b_{8,1/2}(0.01) - 1) \\ &= (-\infty; 1 - 1) \\ &= (-\infty; 0) \end{aligned}$$

```
92 m <- sum(tibia.LRF - x0 != 0) # 8
93 alpha <- 0.01
94 q1 <- qbinom(alpha, m, 1 / 2) - 1 # 1
```

- **Závěr testování**

Protože realizace testovací statistiky $s_E = 1$ nenáleží do kritického oboru, tj. $s_E \notin W$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

4. Testování intervalem spolehlivosti

- **Interval spolehlivosti**

Abychom mohli stanovit hranice 95% pravostranného intervalu spolehlivosti, musíme nejprve naměřené hodnoty největší délky lýtkové kosti vzestupně seřadit. Seřazené hodnoty jsou k dispozici v tabulce 6.

Tabulka 6: Seřazené hodnoty největší délky lýtkové kosti

pořadí	1	2	3	4	5	6	7	8
seřazené X_i	300.0	299.0	323.0	323.0	322.0	301.0	331.5	322.0

Hranice intervalu spolehlivosti potom tvoří mínus nekonečno a hodnota, která se v seřazeném vektoru hodnot nachází na $(b_{n,1/2}(1 - \alpha) + 1)$ -té pozici.

$$\begin{aligned} (d, h) &= (-\infty; X^{(b_{n,1/2}(1-\alpha)+1)}) \\ &= (-\infty; X^{(b_{8,1/2}(1-0.01)+1)}) \\ &= (-\infty; X^{(b_{8,1/2}(0.99)+1)}) \\ &= (-\infty; X^{(8)}) \\ &= (-\infty; 331.5) \end{aligned}$$

```
95 qbinom(1 - alpha, n, 1 / 2) + 1 # 8
96 tibia.LRFs[8] # 331.5
```

- **Závěr testování**

Protože $\tilde{x}_0 = 329.4$ náleží do 99% empirického pravostranného intervalu spolehlivosti, tj. $\tilde{x}_0 = 329.4 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

5. Testování p -hodnotou

- **p -hodnota**

$$p\text{-hodnota} = \Pr(S_E \leq s_E) \Pr(S_E \leq 1) = 0.03515625 \doteq 0.03516$$

- **Závěr testování**

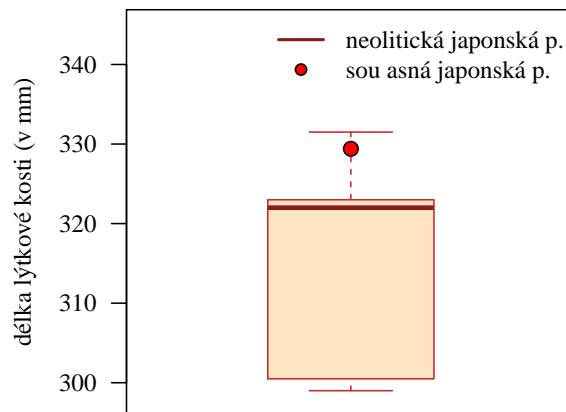
Protože p -hodnota = 0.03515625 je větší než $\alpha = 0.01$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

6. Interpretace výsledků

Za základě všech tří typů testování nezamítáme nulovou hypotézu na hladině významnosti $\alpha = 0.01$. Největší délka lýtkové kosti na pravé straně u žen neolitické japonské populace není statisticky významně menší než u žen současné japonské populace.

7. Grafická vizualizace výsledku testování

Porovnání náhodného výběru s konstantou $\tilde{x}_0 = 329.40$ zobrazíme pomocí krabicového diagramu (viz obrázek 6).



Obrázek 6: Krabicový diagram nejvyšší délky lýtkové kosti na pravé straně u skeletů žen z neolitické japonské populace z oblasti Yoshio Shell Mound

Poznámka: Znaménkový jednovýběrový exaktní test můžeme provést pomocí funkce `SIGN.test()`. Vstupními parametry budou vektor reprezentující náhodný výběr (`tibia.LRF`), hodnota parametru \tilde{x}_0 z nulové hypotézy zadaná argumentem `md = 329.40`, hodnota hladiny významnosti α zadaná prostřednictvím koeficientu spolehlivosti $1 - \alpha$ nastavením hodnoty argumentu `conf.level = 0.99` a typ zvolené alternativní hypotézy (levostranná) zadaný pomocí argumentu `alternative = 'less'`. Součástí výstupu je hodnota mediánu náhodného výběru `median of x = 322`, hod-

```
98 BSDA::SIGN.test(tibia.LRF, md = 329.4, conf.level = 0.99, alt = 'less', exact = F) #
interopolovany IS
```

nota testovací statistiky $s = 1$, interpolované hranice 99% empirického oboustranného intervalu spolehlivosti `-Inf` a `329.8425` a p -hodnota `p-value = 0.03516`. Jediné, co musíme stanovit zvlášť, je horní hranice kritického oboru. ★

One-sample Sign-Test				99
data:	tibia.LRF			100
s = 1, p-value =	0.03516			101
alternative hypothesis:	true median is less than 329.4			102
99 percent confidence interval:				103
	-Inf 329.8425			104
sample estimates:				105
median of x				106
	322			107
Achieved and Interpolated Confidence Intervals:				108
				109
	Conf.Level	L.E.pt	U.E.pt	110
Lower Achieved CI	0.9648	-Inf	323.0000	111
Interpolated CI	0.9900	-Inf	329.8425	112
Upper Achieved CI	0.9961	-Inf	331.5000	113
				114
				115
				116

9.2 Znaménkový jednovýběrový asymptotický test

Pro náhodný výběr o rozsahu $n > 30$ máme možnost použít k otestování nulové hypotézy asymptotickou variantu testu. Tuto variantu nazýváme znaménkovým jednovýběrovým asymptotickým testem. Testovací statistika asymptotického variantu testu má tvar

$$S_A = \frac{S_E - \frac{m}{2}}{\sqrt{\frac{m}{4}}} \quad (9.2)$$

kde S_E je testovací statistika definovaná vztahem 9.1 a m je počet nenulových rozdílů $X_i - \tilde{x}_0$. Za platnosti nulové hypotézy pochází statistika S_A ze standardizovaného normálního rozdělení, tj.

$$S_A = \frac{S_E - \frac{m}{2}}{\sqrt{\frac{m}{4}}} \stackrel{H_0}{\sim} N(0, 1). \quad (9.3)$$

Kritický obor podle zvolené alternativní hypotézy má tvar

$$\begin{array}{ll} H_{11} : \tilde{x} \neq \tilde{x}_0 & W = (-\infty; u_{\alpha/2}) \cup (u_{1-\alpha/2}; \infty) \\ H_{12} : \tilde{x} > \tilde{x}_0 & W = (u_{1-\alpha}; \infty) \\ H_{13} : \tilde{x} < \tilde{x}_0 & W = (-\infty; u_{\alpha}) \end{array}$$

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

Interval spolehlivosti má podle zvolené alternativní hypotézy jeden z následujících tvarů

$$\begin{array}{ll} H_{11} : \tilde{x} \neq \tilde{x}_0 & (d, h) = (X^{(C_{1-\alpha/2})}; X^{(n+1-C_{1-\alpha/2})}) \\ H_{12} : \tilde{x} > \tilde{x}_0 & (d, \infty) = (X^{(C_{1-\alpha})}; \infty) \\ H_{13} : \tilde{x} < \tilde{x}_0 & (-\infty, h) = (-\infty; X^{(n+1-C_{1-\alpha})}) \end{array}$$

kde n je rozsah náhodného výběru, $C_{1-\alpha/2} = \frac{n}{2} - u_{1-\alpha/2}\sqrt{\frac{n}{4}}$, $C_{1-\alpha} = \frac{n}{2} - u_{1-\alpha}\sqrt{\frac{n}{4}}$, $X^{(1)} \leq \dots \leq X^{(n)}$ značí vzestupně seřazené hodnoty X_i , $i = 1, \dots, n$, a $X^{(k)}$ značí k -tou hodnotu v seřazené posloupnosti $X^{(1)} \leq \dots \leq X^{(n)}$.

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

$$\begin{array}{ll} H_{11} : \tilde{x} \neq \tilde{x}_0 & p\text{-hodnota} = 2 \min\{\Pr(S_A \leq s_A), \Pr(S_A \geq s_A)\} \\ H_{12} : \tilde{x} > \tilde{x}_0 & p\text{-hodnota} = \Pr(S_A \geq s_A) \\ H_{13} : \tilde{x} < \tilde{x}_0 & p\text{-hodnota} = \Pr(S_A \leq s_A) \end{array}$$

kde S_A je náhodná veličina, s_A je realizace testovací statistiky S_A (viz vzorec 9.2), tedy konkrétní číslo, $\Pr(S_A \geq s_A) = 1 - \Pr(S_A < s_A) = 1 - \Pr(S_A \leq s_A)$, což vyplývá z faktu, že náhodná veličina S_A pochází z normálního (spojitého) rozdělení (viz kapitola ??), a $\Pr(S_A \leq s_A)$ je distribuční funkce standardizovaného normálního rozdělení, jejíž hodnotu získáme pomocí \mathbb{R} a implementované funkce `pnorm()`.

Poznámka: Všimněme si, že ve vzorcích intervalu spolehlivosti figuruje rozsah náhodného výběru n , zatímco ve vzorcích testovací statistiky a hranic kritického oboru pracujeme s počtem nenulových rozdílů m .

Příklad 9.4. Znaménkový jednovýběrový asymptotický test

Mějme datový soubor 19-more-samples-correlations-skull.txt a proměnnou nose.B popisující šířku nosu v mm (viz sekce ??). Dále máme k dispozici údaje o šířce nosu mužů současné malajské populace ($m_m = 26.90$ mm, $n_m = 45$) uveřejněné ve studii (Ibrahim, 2017). Na hladině významnosti $\alpha = 0.01$ testujte hypotézu o shodě šířky nosu starověké malajské populace a současné malajské populace.

Řešení příkladu 9.4

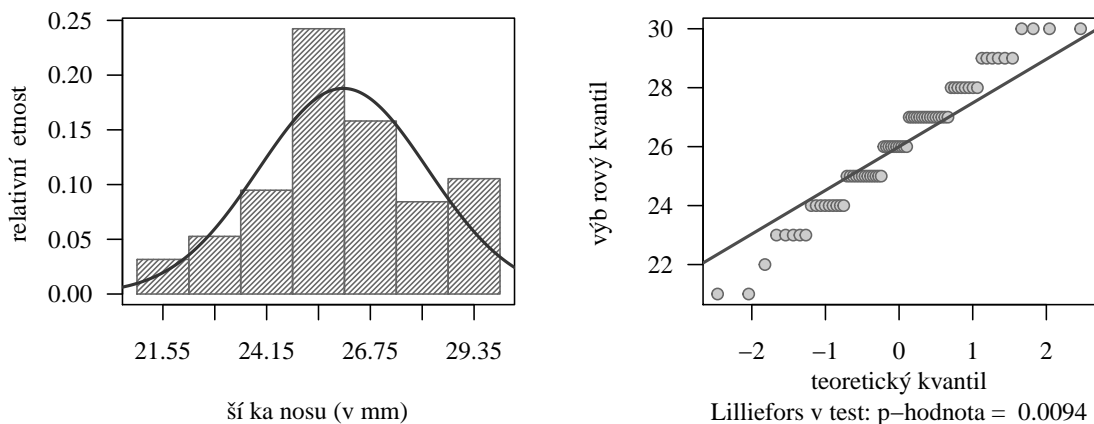
Načteme datový soubor a vybereme z datové tabulky naměřené šířky nosu (nose.B) mužů malajské populace (pop == 'mal'). Nakonec z vektoru naměřených hodnot odstraníme chybějící hodnoty a zjistíme rozsah náhodného výběru.

```
117 data <- read.delim('00-Data//19-more-samples-correlations-skull.txt')
118 # head(data)
119 nose.BM <- data[data$pop == 'mal', 'nose.B']
120 nose.BM <- na.omit(nose.BM)
121 n <- length(nose.BM) # 73
```

Datový soubor obsahuje údaje o šířce nosu u 73 mužů starověké malajské populace.

V příkladu se zaměříme na porovnání šířek nosu dvou malajských populací, přičemž u jedné populace (starověká) máme k dispozici naměřené hodnoty, na základě kterých můžeme zjistit, zda náhodná veličina X popisující šířku nosu u mužů této populace pochází z normálního rozdělení, tj. $X \sim N(\mu, \sigma^2)$, kde skutečný rozptyl σ^2 neznáme. Druhá populace (současné) je reprezentována pouze hodnotou aritmetického průměru ($m_m = 26.90$ mm). Řešení příkladu tedy vede na situaci, kdy střední hodnotu jednoho náhodného výběru porovnáváme s konkrétním číslem, tedy na jednovýběrový test o střední hodnotě μ při neznámém rozptylu σ^2 . Nejprve je však potřeba ověřit, zda náhodný výběr šířek nosu u mužů starověké malajské populace pochází z normálního rozdělení.

Vzhledem k velkému rozsahu náhodného výběru otestujeme normalitu Lillieforsovým testem ($\alpha = 0.05$) v kombinaci s QQ-diagramem a histogramem. Naměřené hodnoty rozdělíme do sedmi ekvidistantních intervalů se šířkou 1.3 mm prostřednictvím stanovených hranic 20.9, 22.2, ..., 30.0 (viz obrázek 7).



Obrázek 7: Histogram a QQ-diagram šířky nosu mužů malajské populace

Protože p -hodnota = 0.0094 je menší než 0.05, nulovou hypotézu o normalitě náhodného výběru šířek nosu mužů starověké malajské populace zamítáme na hladině významnosti $\alpha = 0.05$. Z pohledu na histogram by se mohlo zdát, že naměřené hodnoty kopírují křivku hustoty normálního rozdělení dostatečně. Je však potřeba si uvědomit, že rozsah náhodného výběru je již celkem vysoký a při takovém počtu hodnot by podobnost histogramu s křivkou hustoty měla být mnohem vyšší. Z QQ-diagramu je potom jasně patrné, že vykreslené body se v těsném okolí

referenční křivky příliš nepochybují. Náhodný výběr naměřených šířek nosu u mužů starověké malajské populace tedy nepochází z normálního rozdělení.

Protože se náhodný výběr neřídí normálním rozdělením, nemůžeme hypotézu ze zadání otestovat pomocí parametrického testu o střední hodnotě μ . K testování hypotézy použijeme neparametrický znaménkový jednovýběrový test, zde konkrétně jeho asymptotickou variantu, jelikož rozsah náhodného výběru je větší než 30. Připomeňme, že při použití neparametrických testů pracujeme s mediány namísto se středními hodnotami. Naším úkolem je otestovat hypotézu o shodě šířky nosu starověké malajské populace a současné malajské populace. Tato věta je zněním nulové hypotézy, jednak proto, že v zadání přímo o nulové hypotéze mluvíme a jednak proto, že shoda implikuje rovnost a rovnost je vždy součástí nulové hypotézy. Zbývá dodefinovat alternativní hypotézu tak, aby byla doplňkem k nulové hypotéze. Testování provedeme v posloupnosti sedmi kroků.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : Medián šířky nosu mužů malajské populace je shodný s mediánem šířky nosu mužů současné malajské populace.

H_1 : Medián šířky nosu mužů malajské populace není shodný s mediánem šířky nosu mužů současné malajské populace.

- **matematická formulace** nulové a alternativní hypotézy

H_0 : $\tilde{x} = \tilde{x}_0$, kde $\tilde{x}_0 = 26.90$

H_1 : $\tilde{x} \neq \tilde{x}_0$, kde $\tilde{x}_0 = 26.90$ (oboustranná alternativa)

2. Volba hladiny významnosti

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

3. Testování kritickým oborem

- **Testovací statistika**

Nejprve vypočítáme vektor rozdílů naměřených hodnot X_i a konstanty \tilde{x}_0 , tj. $X_i - \tilde{x}_0$, následně stanovíme počet kladných rozdílů S_E a počet nenulových rozdílů m .

$$S_E = \sum_{i=1}^n I(X_i - \tilde{x}_0 > 0) = \sum_{i=1}^{73} I(X_i - 26.90 > 0) = 33.$$

```
122 x0 <- 26.90
123 I <- (nose.BM > x0)
124 SE <- sum(I) # 33
125 m <- sum(nose.BM - x0 != 0) # 73
```

Všech 73 rozdílů $X_i - \tilde{x}_0$, $i = 1, \dots, 73$, je nenulových, tedy $m = 73$. Počet kladných rozdílů $S_E = 33$. Nyní vypočítáme testovací statistiku asymptotické varianty znaménkového testu podle vzorce 9.2.

$$S_A = \frac{S_E - \frac{m}{2}}{\sqrt{\frac{m}{4}}} = \frac{33 - \frac{73}{2}}{\sqrt{\frac{73}{4}}} = \frac{0}{33 - 36.5} \sqrt{18.25} = \frac{-3.5}{4.272002} = -0.819288 \doteq -0.8193$$

```
126 SA <- (SE - m / 2) / sqrt(m / 4) # -0.819288
```

- Kritický obor

$$\begin{aligned}
 W &= (-\infty; u_{\alpha/2}) \cup (u_{1-\alpha/2}; \infty) \\
 &= (-\infty; u_{0.01/2}) \cup (u_{1-0.01/2}; \infty) \\
 &= (-\infty; u_{0.005}) \cup (u_{0.99}; \infty) \\
 &= (-\infty; -2.575829) \cup (2.575829; \infty)
 \end{aligned}$$

```

127 alpha <- 0.01
128 qnorm(alpha / 2) # -2.575829
129 qnorm(1 - alpha / 2) # 2.575829

```

- Závěr testování

Protože realizace testovací statistiky $s_A = -0.819288$ nenáleží do kritického oboru, tj. $s_A \notin W$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

4. Testování intervalem spolehlivosti

- Interval spolehlivosti

Proti oboustranné alternativě postavíme oboustranný interval spolehlivosti. Pomocí příkazu `sort()` nejprve vzestupně seřadíme naměřené hodnoty X_i , $i = 1, \dots, 73$. Hranice 99% intervalu spolehlivosti budou potom $(C_{1-\alpha/2})$ -tá a $(n + 1 - C_{1-\alpha/2})$ -tá hodnota v seřazeném vektoru naměřených hodnot, kde

$$\begin{aligned}
 C_{1-\alpha/2} &= \frac{n}{2} - u_{1-\alpha/2} \sqrt{\frac{n}{4}} \\
 &= \frac{73}{2} - u_{1-0.01/2} \sqrt{\frac{73}{4}} \\
 &= 36.5 - u_{0.995} \sqrt{18.25} \\
 &= 36.5 - 2.575829 \times 4.272002 \\
 &= 25.49605 \doteq 25
 \end{aligned}$$

Interval spolehlivosti má potom následující tvar

$$\begin{aligned}
 (d, h) &= (X^{(C_{1-\alpha/2})}; X^{(n+1-C_{1-\alpha/2})}) \\
 &= (X^{(25)}; X^{(73+1-25)}) \\
 &= (X^{(25)}; X^{(49)}) \\
 &= (25; 27)
 \end{aligned}$$

```

130 nose.BMs <- sort(nose.BM)
131 C1 <- round(n / 2 - qnorm(1 - alpha / 2) * sqrt(n / 4)) # 29
132 C2 <- round(n + 1 - C1) # 53
133 nose.BMs[C1] # 25
134 nose.BMs[C2] # 27

```

- Závěr testování

Protože $\tilde{x}_0 = 26.90$ náleží do 99% empirického pravostranného intervalu spolehlivosti, tj. $\tilde{x}_0 = 26.90 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

5. Testování p -hodnotou

Ze vztahu 9.3 víme, že testovací statistika S_A pochází ze standardizovaného normálního rozdělení, tj. $S_A \sim N(0, 1)$, které náleží do třídy spojitých rozdělení. Proto $\Pr(S_A \geq s_A) = 1 - \Pr(S_A < s_A) = 1 - \Pr(S_A \leq s_A)$, jak je uvedeno v kapitole XXX.

- p -hodnota

$$\begin{aligned} p\text{-hodnota} &= 2 \min\{\Pr(S_A \leq s_A), \Pr(S_A \geq s_A)\} \\ &= 2 \min\{\Pr(S_A \leq -0.819288), \Pr(S_A \geq -0.819288)\} \\ &= 2 \min\{\Pr(S_A \leq -0.819288), 1 - \Pr(S_A \leq -0.819288)\} \\ &= 2 \min\{0.206311, 0.793689\} \\ &= 2 \times 0.206311 = 0.4126221 \doteq 0.4126 \end{aligned}$$

```
135 p.hodnota <- 2 * min (pnorm(SA), 1 - pnorm(SA)) # 0.4126221
```

- Závěr testování

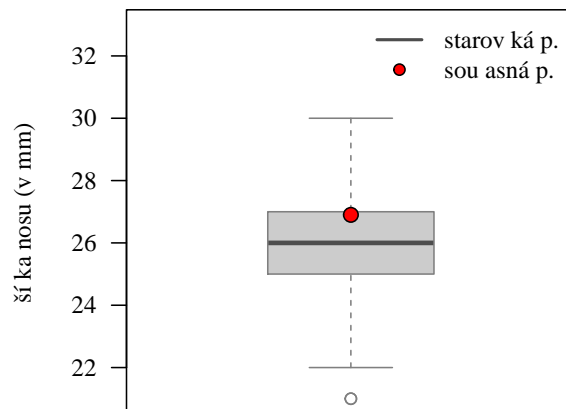
Protože p -hodnota = 0.4126 je větší než $\alpha = 0.01$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

6. Interpretace výsledků

Za základě všech tří typů testování nezamítáme nulovou hypotézu na hladině významnosti $\alpha = 0.01$. Mezi šířkou nosu mužů starověké a současné malajské populace není statisticky významný rozdíl.

7. Grafická vizualizace výsledku testování

Grafické porovnání šířky nosu obou malajských populací provedeme krabicovým diagramem (viz obrázek 8).



Obrázek 8: Krabicový diagram šířky nosu mužů malajské populace



Příklad 9.5. Znaménkový jednovýběrový asymptotický test

Mějme datový soubor 28-one-sample-mean-skull-m5.txt a proměnné skull.H popisující basio-bregmatickou výšku lebky v mm a skull.B popisující největší šířku lebky v mm u skeletů mužského pohlaví z egyptské populace (viz sekce ??). Dále máme k dispozici údaje o basio-bregmatické výšce lebky ($m_H = 134.0$), o největší šířce lebky ($m_B = 131.4$) a o podílu největší šířky lebky a basio-bregmatické výšky lebky ($m_{BH} = 0.981$) u skeletů mužské egyptské populace z předdynastické doby. Na hladině významnosti $\alpha = 0.05$ zjistíte, zda je podíl největší šířky lebky a basio-bregmatické výšky lebky u mužů žijících v době okolo 150 let po našem letopočtu větší než u mužů žijících v předdynastické době.

Řešení příkladu 9.5

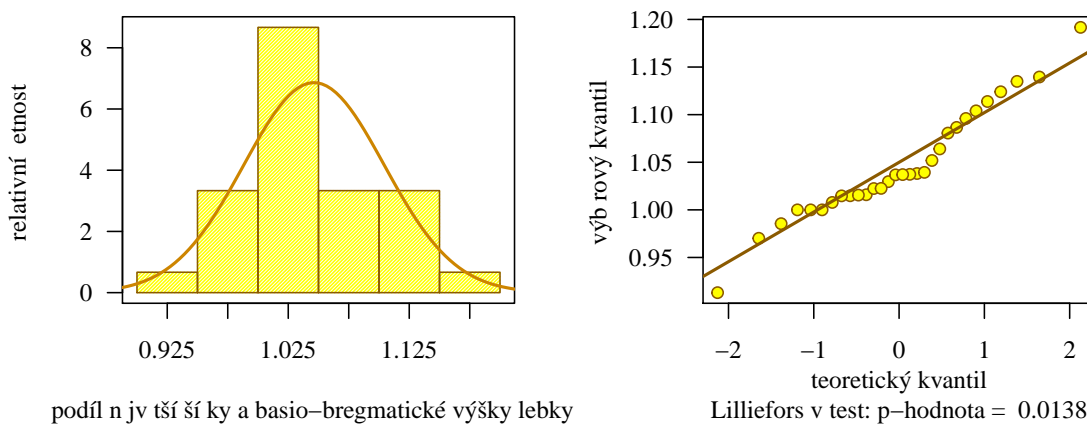
Načteme datový soubor a vybereme z datové tabulky naměřené hodnoty největší šířky lebky (skull.B) a basio-bregmatické výšky lebky skull.H u mužů starověké egyptské populace z období 150 let n.l. (year == 150). Z vektoru naměřených hodnot odstraníme chybějící pozorování a zjistíme rozsah náhodného výběru. Nakonec vypočítáme podíl největší šířky a basio-bregmatické výšky. Hodnoty podílů vložíme do proměnné skull.BH.

```
136 data <- read.delim('00-Data//28-one-sample-mean-skull-m5.txt')
137 # head(data)
138 data.BH <- data[data$year == 150, c('skull.B', 'skull.H')]
139 data.BH <- na.omit(data.BH)
140 skull.BH <- data.BH$skull.B / data.BH$skull.H
141 n <- length(skull.BH)
```

Datový soubor obsahuje údaje o basio-bregmatické výšce a největší šířce lebky u 30 mužů starověké egyptské populace z období 150 let n.l.

V příkladu se zaměříme na porovnání podílu největší šířky lebky a basio-bregmatické výšky lebky dvou egyptských populací, přičemž u jedné populace (150 let n.l.) máme k dispozici naměřené hodnoty, na základě kterých můžeme zjistit, zda náhodná veličina X popisující podíl největší šířky a basio-bregmatické výšky této populace pochází z normálního rozdělení, tj. $X \sim N(\mu, \sigma^2)$, kde skutečný rozptyl σ^2 neznáme. Podíl největší šířky a basio-bregmatické výšky je u druhé populace (z předdynastické doby) je reprezentován pouze hodnotou jeho aritmetického průměru ($m_{BH} = 0.981$ mm). Řešení příkladu tedy vede na situaci, kdy střední hodnotu jednoho náhodného výběru porovnáváme s konkrétním číslem, tedy na jednovýběrový test o střední hodnotě μ při neznámém rozptylu σ^2 . Nejprve je však potřeba ověřit, zda náhodný výběr podílů u populace z období 150 let n.l. pochází z normálního rozdělení.

Vzhledem k rozsahu ($n = 30$) náhodného výběru otestujeme normalitu podílů Lillieforsovým testem ($\alpha = 0.05$) v kombinaci s QQ-diagramem a histogramem. Naměřené hodnoty rozdělíme do sedmi ekvidistantních intervalů se šířkou 1.3 mm prostřednictvím stanovených hranic 20.9, 22.2, ..., 30.0 (viz obrázek 9).



Obrázek 9: Histogram a QQ-diagram největší šířky a basio-bregmatické výšky lebky u mužů egyptské populace

Protože p -hodnota = 0.0138 je menší než 0.05, nulovou hypotézu o normalitě náhodného výběru podílu největší šířky a basio-bregmatické zamítáme na hladině významnosti $\alpha = 0.05$. Z pohledu na histogram vidíme, že normalitu porušuje vysoký počet hodnot v intervalu se středem 1.025 a naopak nízký počet hodnot v intervalu se středem 1.075. Náhodný výběr podílů nepochází z normálního rozdělení.

Protože se náhodný výběr podílů neřídí normálním rozdělením, nemůžeme hypotézu ze zadání otestovat pomocí parametrického testu o střední hodnotě μ . K testování hypotézy použijeme neparametrický znaménkový jednovýběrový test, zde konkrétně jeho asymptotickou variantu, neboť rozsah náhodného výběru je rovný 30. Naším úkolem je zjistit, zda je podíl největší šířky a basio-bregmatické výšky u mužů žijících v době okolo 150 let po našem letopočtu větší než u mužů žijících v předdynastické době. Tato věta je zněním alternativní hypotézy. Zbývá dodefinovat nulovou hypotézu.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : Medián podílu největší šířky a basio-bregmatické výšky lebky u mužů žijících v době okolo 150 let n.l. je menší nebo roven mediánu podílu největší šířky a basio-bregmatické výšky lebky u mužů žijících v předdynastické době.

H_1 : Medián podílu největší šířky a basio-bregmatické výšky lebky u mužů žijících v době okolo 150 let n.l. je větší než medián podílu největší šířky a basio-bregmatické výšky lebky u mužů žijících v předdynastické době.

- **matematická formulace** nulové a alternativní hypotézy

$H_0 : \tilde{x} \leq \tilde{x}_0$, kde $\tilde{x}_0 = 0.981$

$H_1 : \tilde{x} > \tilde{x}_0$, kde $\tilde{x}_0 = 0.981$ (pravostranná alternativa)

2. Volba hladiny významnosti

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

3. Testování kritickým oborem

- **Testovací statistika**

Nejprve vypočítáme vektor rozdílů naměřených hodnot X_i a konstanty \tilde{x}_0 , tj. $X_i - \tilde{x}_0 = X_i - 0.981$, následně stanovíme počet kladných rozdílů S_E a počet nenulových rozdílů m .

$$S_E = \sum_{i=1}^n I(X_i - \tilde{x}_0 > 0) = \sum_{i=1}^{30} I(X_i - 0.981 > 0) = 28.$$

```
142 x0 <- 0.981
143 I <- (skull.BH > x0)
144 SE <- sum(I) # 28
145 m <- sum(skull.BH - x0 != 0) # 30
```

Všech 30 rozdílů $X_i - \tilde{x}_0$, $i = 1, \dots, 30$, je nenulových, tedy $m = 30$. Počet kladných rozdílů $S_E = 28$. Nyní vypočítáme testovací statistiku asymptotické varianty znaménkového testu podle vzorce 9.2.

$$S_A = \frac{S_E - \frac{m}{2}}{\sqrt{\frac{m}{4}}} = \frac{28 - \frac{30}{2}}{\sqrt{\frac{30}{4}}} = \frac{0}{28 - 15} \sqrt{7.5} = \frac{13}{2.738613} = 4.751236 \doteq 4.7512$$

```
146 SA <- (SE - m / 2) / sqrt(m / 4) # 4.746929
```

- Kritický obor

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

```

147 alpha <- 0.05
148 qnorm(1 - alpha) # 1.644854

```

- Závěr testování

Protože realizace testovací statistiky $s_A = 4.7469$ náleží do kritického oboru, tj. $s_A \in W$, H_0 zamítáme na hladině významnosti $\alpha = 0.05$.

4. Testování intervalem spolehlivosti

- Interval spolehlivosti

Pomocí příkazu `sort()` nejprve vzestupně seřadíme naměřené hodnoty X_i , $i = 1, \dots, 73$. Hranice 95% intervalu spolehlivosti budou potom nekonečno a $(C_{1-\alpha})$ -tá hodnota v seřazeném vektoru naměřených hodnot, kde

$$\begin{aligned}
 C_{1-\alpha} &= \frac{n}{2} - u_{1-\alpha} \sqrt{\frac{n}{4}} \\
 &= \frac{30}{2} - u_{1-0.05} \sqrt{\frac{30}{4}} \\
 &= 15 - u_{0.95} \sqrt{7.5} \\
 &= 15 - 1.644854 \times 2.738613 \\
 &= 10.49538 \doteq 10
 \end{aligned}$$

Interval spolehlivosti má potom následující tvar

$$\begin{aligned}
 (d, h) &= \left(X^{(C_{1-\alpha})}; \infty \right) \\
 &= \left(X^{(10)}; \infty \right) \\
 &= \left(X^{(10)}; \infty \right) \\
 &= (1.015504; \infty)
 \end{aligned}$$

```

149 skull.BHs <- sort(skull.BH)
150 C1 <- round(n / 2 - qnorm(1 - alpha) * sqrt(n / 4)) # 10
151 skull.BHs[C1] # 1.015504

```

- Závěr testování

Protože $\tilde{x}_0 = 0.981$ nenáleží do 95% empirického jednostranného intervalu spolehlivosti, tj. $\tilde{x}_0 = 0.981 \notin IS$, H_0 zamítáme na hladině významnosti $\alpha = 0.05$.

5. Testování p -hodnotou

- p -hodnota

$$\begin{aligned} p\text{-hodnota} &= \Pr(S_A \geq s_A) \\ &= \Pr(S_A \geq 4.746929) \\ &= 1 - \Pr(S_A \leq 4.746929) \\ &= 1.032643 \times 10^{-06} \doteq 1.03264 \times 10^{-6} \end{aligned}$$

```
152 p.hodnota <- 1 - pnorm(SA) # 1.032643e-06
```

- **Závěr testování**

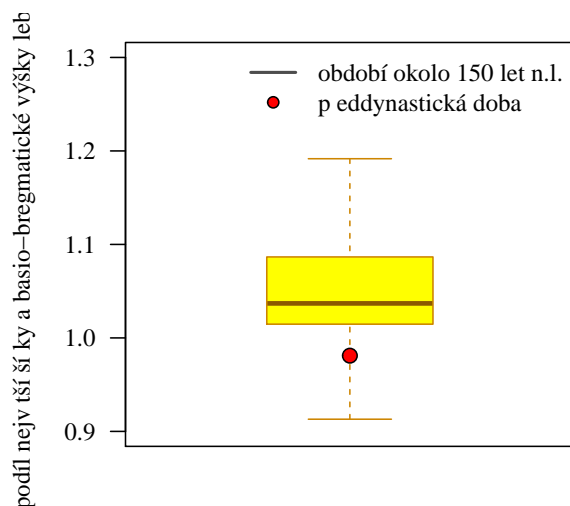
Protože p -hodnota = 1.03264×10^{-6} je menší než $\alpha = 0.05$, H_0 zamítáme na hladině významnosti $\alpha = 0.01$.

6. Interpretace výsledků

Za základě všech tří typů testování zamítáme nulovou hypotézu na hladině významnosti $\alpha = 0.05$. Podíl největší šířky a basio-bregmatické výšky lebky u mužů z egyptské populace z období okolo 150 let n.l. je statisticky významně větší než u mužů z předdynastické doby.

7. Grafická vizualizace výsledku testování

Grafické porovnání podílu největší šířky a basio-bregmatické výšky lebky obou egyptských populací provedeme krabicovým diagramem (viz obrázek 10).



Obrázek 10: Krabicový diagram největší šířky a basio-bregmatické výšky lebky u mužů egyptské populace



Příklad 9.6. Znaménkový jednovýběrový asymptotický test

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 ??). Dále máme k dispozici údaje o největší délce klíční kosti z pravé strany mužů z populace severní Indie ($m_I = 148.0$ mm, $s_I = 8.60$ mm, $n_I = 260$). Na hladině významnosti $\alpha = 0.05$ zjistíte, zda je délka klíční kosti z pravé strany u mužů indické populace z Varanasi menší než u mužů ze severní Indie.

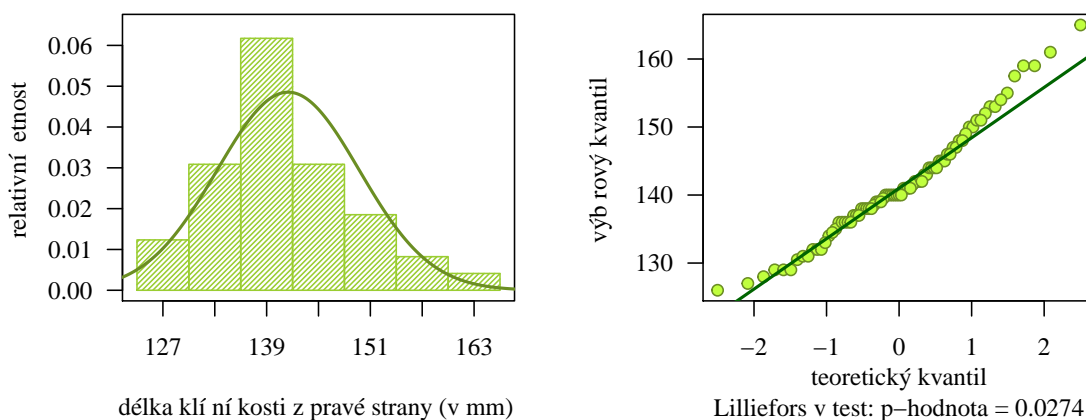
Řešení příkladu 9.6

Načteme datový soubor a vybereme z datové tabulky naměřené délky klíční kosti (cla.L) mužů populace z Varanasi (pop == 'ind2'). Nakonec z vektoru naměřených hodnot odstraníme chybějící údaje a zjistíme rozsah náhodného výběru.

```
153 data <- read.delim('00-Data//18-more-samples-variances-clavicle.txt')
154 # head(data)
155 cla.LV <- data[data$pop == 'ind2', 'cla.L']
156 cla.LV <- na.omit(cla.LV)
157 n <- length(cla.LV) # 81
```

Datový soubor obsahuje údaje o největší délce klíční kosti z pravé strany u 81 mužů indické populace z Varanasi.

V příkladu se zaměříme na porovnání délek klíčních kostí dvou populací, přičemž u jedné populace (indická populace z Varanasi) máme k dispozici naměřené hodnoty. Na základě těchto hodnot můžeme zjistit, zda náhodná veličina X popisující největší délku klíční kosti u mužů indické populace z Varanasi pochází z normálního rozdělení, tj. $X \sim N(\mu, \sigma^2)$, kde skutečný rozptyl σ^2 neznáme. Druhá populace (populace ze severní Indie) je reprezentována pouze hodnotou aritmetického průměru ($m_I = 148.00$ mm) a směrodatnou odchylkou ($s_I = 8.60$ mm). Řešení příkladu tedy vede na situaci, kdy střední hodnotu jednoho náhodného výběru porovnáváme s konkrétním číslem, tedy na jednovýběrový test o střední hodnotě μ při neznámém rozptylu σ^2 (viz kapitola ??). Nejprve je však potřeba ověřit, zda náhodný výběr délek klíčních kostí u mužů indické populace z Varanasi pochází z normálního rozdělení. Tomuto předpokladu jsme se však už věnovali v rámci příkladu ?? v sekci ??, kde jsme zjistili, že náhodný výběr délek klíčních kostí u mužů indické populace z Varanasi z normálního rozdělení nepochází (pro připomenutí viz obrázek 11).



Obrázek 11: Histogram a QQ-diagram největší délky klíční kosti z pravé strany u mužů z populace z Varanasi

Protože náhodný výběr nepochází z normálního rozdělení, nemůžeme hypotézu ze zadání otestovat pomocí parametrického testu o střední hodnotě μ . K testování hypotézy použijeme neparametrický znaménkový jednovýběrový asymptotický test. Naším úkolem je zjistit, zda je délka klíční kosti z pravé strany u mužů indické populace z Varanasi menší než u mužů ze severní Indie. Tato věta je zněním alternativní hypotézy, neboť v jejím znění není zmínka o (nu-

lové) hypotéze, ani o rovnosti. Zbývá dodefinovat nulovou hypotézu tak, aby byla doplňkem k hypotéze alternativní.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : Medián největší délky klíční kosti z pravé strany u mužů indické populace z Varanasi je větší nebo rovný mediánu největší délky klíční kosti z pravé strany u mužů ze severní Indie.

H_1 : Medián největší délky klíční kosti z pravé strany u mužů indické populace z Varanasi je menší než medián největší délky klíční kosti z pravé strany u mužů ze severní Indie.

- **matematická formulace** nulové a alternativní hypotézy

H_0 : $\tilde{x} \geq \tilde{x}_0$, kde $\tilde{x}_0 = 148.00$

H_1 : $\tilde{x} < \tilde{x}_0$, kde $\tilde{x}_0 = 148.00$ (oboustranná alternativa)

2. Volba hladiny významnosti

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

3. Testování kritickým oborem

- **Testovací statistika**

Nejprve vypočítáme vektor rozdílů naměřených hodnot X_i a konstanty \tilde{x}_0 , tj. $X_i - \tilde{x}_0$, následně stanovíme počet kladných rozdílů S_E a počet nenulových rozdílů m .

$$S_E = \sum_{i=1}^n I(X_i - \tilde{x}_0 > 0) = \sum_{i=1}^{81} I(X_i - 148 > 0) = 15.$$

```
158 x0 <- 148.00
159 I <- (cla.LV > x0)
160 SE <- sum(I) # 15
161 m <- sum(cla.LV - x0 != 0) # 79
```

Z celkového počtu 81 rozdílů $X_i - \tilde{x}_0$, $i = 1, \dots, 81$, je 79 rozdílů nenulových, tedy $m = 79$. Počet kladných rozdílů $S_E = 15$. Nyní již můžeme vypočítat testovací statistiku asymptotické varianty znaménkového testu podle vzorce 9.2.

$$S_A = \frac{S_E - \frac{m}{2}}{\sqrt{\frac{m}{4}}} = \frac{15 - \frac{79}{2}}{\sqrt{\frac{79}{4}}} = \frac{0}{15 - 39.5} \sqrt{19.75} = \frac{-24.5}{4.444097} = -5.512931 \doteq -5.5129$$

```
162 SA <- (SE - m / 2) / sqrt(m / 4) # -5.512931
```

- **Kritický obor**

$$\begin{aligned} W &= (-\infty; u_\alpha) \\ &= (-\infty; u_{0.05}) \\ &= (-\infty; -1.644854) \end{aligned}$$

```
163 alpha <- 0.05
164 qnorm(alpha) # -1.644854
```

- **Závěr testování**

Protože realizace testovací statistiky $s_A = -5.512931$ náleží do kritického oboru, tj. $s_A \in W$, H_0 zamítáme na hladině významnosti $\alpha = 0.05$.

4. Testování intervalem spolehlivosti

- **Interval spolehlivosti**

Pomocí příkazu `sort()` nejprve vzestupně seřadíme naměřené hodnoty X_i , $i = 1, \dots, 79$. Hranice 95% intervalu spolehlivosti budou potom mínus nekonečno a $(n + 1 - C_{1-\alpha})$ -tá hodnota v seřazeném vektoru naměřených hodnot, kde

$$\begin{aligned} C_{1-\alpha} &= \frac{n}{2} - u_{1-\alpha} \sqrt{\frac{n}{4}} \\ &= \frac{81}{2} - u_{1-0.05} \sqrt{\frac{81}{4}} \\ &= 40.5 - u_{0.95} \sqrt{20.25} \\ &= 40.5 - 1.644854 \times 4.5 \\ &= 33.09816 \doteq 33 \end{aligned}$$

Interval spolehlivosti má potom následující tvar

$$\begin{aligned} (d, h) &= \left(-\infty; X^{(n+1-C_{1-\alpha})}\right) \\ &= \left(-\infty; X^{(81+1-33)}\right) \\ &= \left(-\infty; X^{(49)}\right) \\ &= \left(-\infty; 27\right) \end{aligned}$$

```
165 cla.LVs <- sort(cla.LV)
166 C1 <- round(n / 2 - qnorm(1 - alpha) * sqrt(n / 4)) # 33
167 C2 <- round(n + 1 - C1) # 49
168 cla.LVs[49] # 142
```

- **Závěr testování**

Protože $\tilde{x}_0 = 148.00$ nenáleží do 95% empirického oboustranného intervalu spolehlivosti, tj. $\tilde{x}_0 = 148.00 \notin IS$, H_0 zamítáme na hladině významnosti $\alpha = 0.05$.

5. Testování p -hodnotou

- **p -hodnota**

$$p\text{-hodnota} = \Pr(S_A \leq s_A) = \Pr(S_A \leq -5.512931) = 1.764536^{-8} \doteq 1.7645^{-8}$$

```
169 p.hodnota <- pnorm(SA) # 1.764536e-08
```

- **Závěr testování**

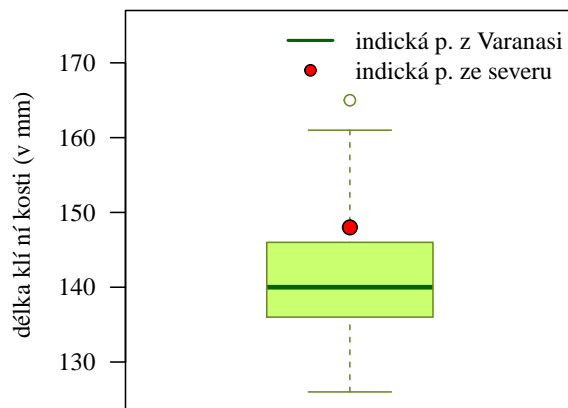
Protože p -hodnota = $1.7645 \cdot 10^{-8}$ je menší než $\alpha = 0.05$, H_0 zamítáme na hladině významnosti $\alpha = 0.05$.

6. **Interpretace výsledků**

Za základě všech tří typů testování zamítáme nulovou hypotézu na hladině významnosti $\alpha = 0.05$. Největší délka klíční kosti na pravé straně u mužů indické populace z Varanasi je statisticky významně menší než největší délka klíční kosti na pravé straně u mužů populace ze severní Indie.

7. **Grafická vizualizace výsledku testování**

Rozdíl v délce klíčních kostí zaznamenaný mezi oběma populacemi vizualizujeme pomocí krabicového diagramu (viz obrázek 12).



Obrázek 12: Krabicový diagram délky klíční kosti z pravé strany u mužů indické populace z Varanasi



9.3 Znaménkový párový test

Nechť $(X_1, Y_1)^T \dots (X_n, Y_n)^T$ je náhodný výběr z libovolného (ne nutně normálního) dvourozměrného rozdělení. Nechť dále Z_1, \dots, Z_n , $n \geq 2$ je náhodný výběr rozdílů $X - Y$, tj. $Z = (Z_1, \dots, Z_n)^T$, kde $Z_i = X_i - Y_i$, $i = 1, \dots, n$, a nechť tento náhodný výběr pochází z libovolného spojitého rozdělení. Konečně, nechť \tilde{z}_0 je konstanta. Na hladině významnosti α testujeme jednu z následujících tří hypotéz oproti příslušné alternativní hypotéze.

$$\begin{array}{lll} H_{01} : \tilde{z} = \tilde{z}_0 & \text{oproti} & H_{11} : \tilde{z} \neq \tilde{z}_0 \quad (\text{oboustranná alt.}) \\ H_{02} : \tilde{z} \leq \tilde{z}_0 & \text{oproti} & H_{12} : \tilde{z} > \tilde{z}_0 \quad (\text{pravostranná alt.}) \\ H_{03} : \tilde{z} \geq \tilde{z}_0 & \text{oproti} & H_{13} : \tilde{z} < \tilde{z}_0 \quad (\text{levostranná alt.}) \end{array}$$

kde \tilde{z} je medián rozdílů Z_1, \dots, Z_n a \tilde{z}_0 je konstanta, jejíž hodnotu nejčastěji volíme jako $\tilde{z}_0 = 0$. Tato volba odpovídá hypotéze, že rozdíl mezi mediány náhodných veličin X a Y neexistuje (resp. hypotéze, že medián náhodné veličiny X je menší, resp. větší, než medián náhodné veličiny Y). Vzhledem k tomu, že jde finálně o situaci, kdy medián \tilde{z} porovnáváme s konstantou \tilde{z}_0 , testujeme hypotézy o rozdílu mediánů $X - Y$ pomocí exaktní nebo asymptotické varianty znaménkového jednovýběrového testu, analogicky jako je uvedeno v sekcích 9.1 a 9.2.

Výše popsaný test, v rámci kterého převádíme problém porovnávání mediánů dvou náhodných veličin X a Y na problém srovnávání mediánu jejich rozdílů Z s konstantou $\tilde{z}_0 = 0$ a následně jej řešíme pomocí exaktní resp. asymptotické varianty znaménkového jednovýběrového testu, nazýváme znaménkový párový test.

Příklad 9.7. Znaménkový párový exaktní test

Mějme datový soubor 21-goldman-measures.csv obsahující údaje o největší výšce kloubní jamky na pravé straně (acetab.HR) a na levé straně (acetab.HL) u skeletů ze starověké egyptské populace (detaily viz sekce ??). Na hladině významnosti $\alpha = 0.05$ zjistěte, zda existuje rozdíl mezi výškou kloubní jamky z pravé a levé strany u skeletů mužského pohlaví.

Řešení příkladu 9.7

Pomocí příkazu read.delim() načteme datový soubor a pomocí operátoru [] vybereme z datové tabulky údaje o výšce kloubní jamky z pravé strany (acetab.HR), resp. z levé strany (acetab.HL) u mužů (sex == 'm') ze starověké egyptské populace (pop == 'Dynastic Egyptian, El Hesa'). Z vektoru naměřených hodnot odstraníme pomocí funkce na.omit() chybějící údaje a zjistíme rozsah náhodného výběru (dim()).

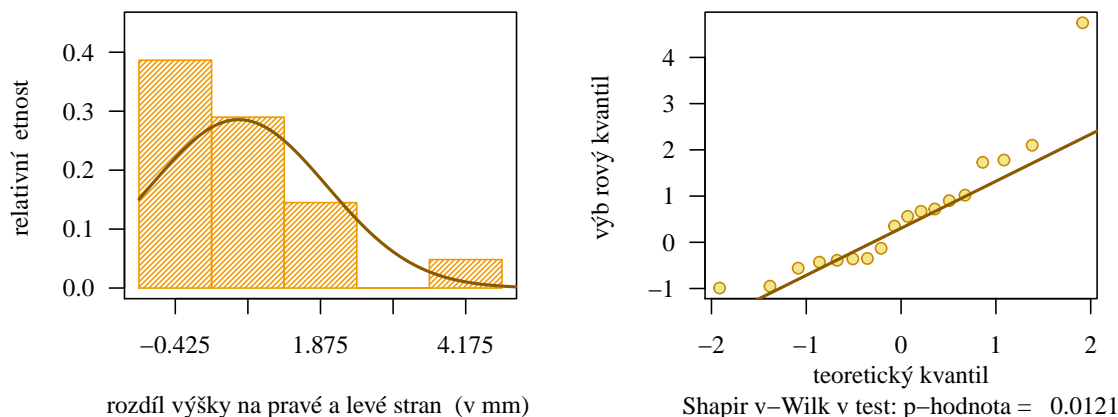
```
170 data <- read.delim('00-Data//21-goldman-measures.csv', sep = ';')
171 # head(data)
172 data.HRL <- data[data$pop == 'Dynastic Egyptian, El Hesa' & data$sex == 'm',
173               c('acetab.HL', 'acetab.HR')]
174 data.HRL <- na.omit(data.HRL)
175 dim(data.HRL) # 18x2
```

Datový soubor obsahuje kompletní údaje o výšce kloubní jamky z pravé i levé strany u 18 mužů ze starověké egyptské populace.

Naším úkolem ze zadání je porovnat naměřené hodnoty na pravé a levé straně. Jde tedy o měření stejného znaku (výška kloubní jamky) sledovaného na stejných subjektech (muži), proto použijeme pro tuto situaci párový test. Prvním krokem k použití tohoto testu je vytvoření rozdílů hodnot naměřených na pravé a levé straně.

```
176 acetab.HR <- data.HRL$acetab.HR
177 acetab.HL <- data.HRL$acetab.HL
178 acetab.HRL <- acetab.HR - acetab.HL
```

Ve druhém kroku je potřeba ověřit normální rozdělení těchto rozdílů. Vzhledem k rozsahu náhodného výběru ($n = 18 < 30$) otestujeme normalitu rozdílů Shapirovým-Wilkovým testem ($\alpha = 0.05$). Pro potřeby vykreslení histogramu rozdělíme naměřené hodnoty do pěti ekvidistatních intervalů s šířkou 1.15 mm prostřednictvím stanovených hranic $-1, 0.15, \dots, 05.75$ (viz obrázek 13).



Obrázek 13: Histogram a diagram rozdílů mezi největší výškou kloubní jamky u mužů na pravé a levé straně

Protože p -hodnota = 0.0121 je menší než 0.05, nulovou hypotézu o normalitě rozdílů zamítáme na hladině významnosti $\alpha = 0.05$. Z histogramu na obrázku 13 vidíme, že naměřené hodnoty jsou vyšikmené doleva s odlehlým pozorováním na pravém chvostu. Histogram naměřených hodnot navíc nevykazuje symetrii ani postupné snižování počtu hodnot na levé straně, jak bychom u normálního rozdělení očekávali. Náhodný výběr rozdílů mezi výškou kloubní jamky na pravé a levé straně nepochází z normálního rozdělení.

Protože náhodný výběr rozdílů nepochází z normálního rozdělení, nemůžeme hypotézu ze zadání otestovat pomocí parametrického párového testu uvedeného v sekci ???. K testování hypotézy musíme použít neparametrickou alternativu párového testu. S ohledem na nízký rozsah náhodného výběru použijeme exaktní variantu znaménkového párového testu. Naším úkolem je zjistit, existuje rozdíl mezi výškou kloubní jamky z pravé a levé strany u skeletů mužského pohlaví. Tato věta bude součástí alternativní hypotézy, neboť rozdíl implikuje nerovnost a nerovnost je vždy součástí alternativní hypotézy. Nulovou hypotézu stanovíme jako doplněk k tomuto tvrzení. Testování provedeme v posloupnosti sedmi kroků.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : Medián rozdílů výšky kloubní jamky na pravé a levé straně u mužů je rovný nule.

H_1 : Medián rozdílů výšky kloubní jamky na pravé a levé straně u mužů není rovný nule.

- **matematická formulace** nulové a alternativní hypotézy

H_0 : $\tilde{z} = \tilde{z}_0$, kde $\tilde{z}_0 = 0$

H_1 : $\tilde{z} \neq \tilde{z}_0$, kde $\tilde{z}_0 = 0$ (oboustranná alternativa)

2. Volba hladiny významnosti

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

3. Testování kritickým oborem

- **Testovací statistika**

V úvodu příkladu jsme vytvořili vektor Z_i jako vektor rozdílů výšky kloubní jamky na pravé straně X_i a výšky kloubní jamky na levé straně Y_i , tj. $Z_i = X_i - Y_i$, $i = 1, \dots, 18$. Nyní je potřeba stanovit rozdíl $Z_i - \tilde{z}_0$. Protože ale $\tilde{z}_0 = 0$, jsou rozdíl $Z_i - \tilde{z}_0$ rovny vektoru rozdílů Z_i , tj. $Z_i - \tilde{z}_0 = Z_i - 0 = Z_i$. Původní hodnoty výšky kloubní jamky na pravé straně, resp. na levé straně, rozdíl mezi pravou a levou stranou, rozdíl $Z_i - \tilde{z}_0$ a znaménko posledních uvedených rozdílů jsou pro názornost uvedeny v tabulce 7.

Tabulka 7: Naměřené hodnoty X_i , Y_i , rozdíl $Z_i = X_i - Y_i$, rozdíl $Z_i - \tilde{z}_0$ a znaménka těchto rozdílů

měření	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
X_i	50.84	52.36	47.80	51.75	46.06	46.64	50.54	47.03	49.02	47.96	48.25	47.78	49.86	49.56	55.41	52.82	52.47	46.02
Y_i	50.97	51.80	48.75	51.40	43.96	44.86	49.64	48.02	49.45	48.52	48.64	48.14	48.13	48.54	50.66	52.10	52.82	45.35
Z_i	-0.13	0.56	-0.95	0.35	2.10	1.78	0.90	-0.99	-0.43	-0.56	-0.39	-0.36	1.73	1.02	4.75	0.72	-0.35	0.67
$Z_i - \tilde{z}_0$	-0.13	0.56	-0.95	0.35	2.10	1.78	0.90	-0.99	-0.43	-0.56	-0.39	-0.36	1.73	1.02	4.75	0.72	-0.35	0.61
+/-	-	+	-	+	+	+	+	-	-	-	-	-	+	+	+	+	-	+

Z tabulky 7 vidíme, že celkem deset rozdílů $Z_i - \tilde{z}_0$, $i = 1, \dots, 18$ je kladných. Hodnota testovací statistiky S_E , která je definovaná jako počet kladných rozdílů, bude tedy rovná 10.

$$S_E = \sum_{i=1}^n I(Z_i - \tilde{z}_0 > 0) = \sum_{i=1}^{18} I(Z_i - 0 > 0) = 10.$$

- **Kritický obor**

Z tabulky 7 dále vidíme, že žádný z celkového počtu $n = 18$ rozdílů není nulový, a tedy počet nenulových

```

179 z0 <- 0
180 I <- (acetab.HRL > z0)
181 tab <- data.frame(rbind("Xi" = acetab.HR, "Yi" = acetab.HL,
182 "Zi" = acetab.HRL, "Zi-z0" = acetab.HRL - z0, "+/-" = I))
183 SE <- sum(I) # 10

```

rozdílů $m = 18$. Kritický obor má potom tvar

$$\begin{aligned}
W &= (-\infty; b_{m,1/2}(\alpha/2) - 1) \cup (b_{m,1/2}(1 - \alpha/2); \infty) \\
&= (-\infty; b_{18,1/2}(0.05/2) - 1) \cup (b_{18,1/2}(1 - 0.05/2); \infty) \\
&= (-\infty; b_{18,1/2}(0.025) - 1) \cup (b_{18,1/2}(0.975); \infty) \\
&= (-\infty; 5 - 1) \cup (13; \infty) \\
&= (-\infty; 4) \cup (13; \infty)
\end{aligned}$$

```

184 m <- sum(acetab.HRL - z0 != 0) # 18
185 alpha <- 0.05
186 q1 <- qbinom(alpha / 2, m, 1 / 2) - 1 # 4
187 q2 <- qbinom(1 - alpha / 2, m, 1 / 2) # 13

```

- **Závěr testování**

Protože realizace testovací statistiky $s_E = 10$ nenáleží do kritického oboru, tj. $s_E \notin W$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

4. Testování intervalem spolehlivosti

- **Interval spolehlivosti**

Abychom mohli stanovit hranice 95% intervalu spolehlivosti, musíme nejprve naměřené rozdíly mezi výškami kloubní jamky na pravé a levé straně vzestupně seřadit. To provedeme příkazem `sort()`. Seřazené hodnoty jsou k dispozici v tabulce 8.

Tabulka 8: Seřazené rozdíly mezi výškami kloubní jamky na pravé a levé straně

pořadí	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
seřazené Z_i	-0.99	-0.95	-0.56	-0.43	-0.39	-0.36	-0.35	-0.13	0.35	0.56	0.67	0.72	0.90	1.02	1.73	1.78	2.10	4.75

Hranice intervalu spolehlivosti potom tvoří ty hodnoty, které se v seřazeném vektoru hodnot nachází na $(b_{n,1/2}(\alpha/2))$ -té pozici a na $(b_{n,1/2}(1 - \alpha/2) + 1)$ -té pozici.

$$\begin{aligned}
(d, h) &= \left(Z^{(b_{n,1/2}(\alpha/2))}; Z^{(b_{n,1/2}(1 - \alpha/2) + 1)} \right) \\
&= \left(Z^{(b_{18,1/2}(0.05/2))}; Z^{(b_{18,1/2}(1 - 0.05/2) + 1)} \right) \\
&= \left(Z^{(5)}; Z^{(14)} \right) \\
&= (-0.39; 1.02)
\end{aligned}$$

- **Závěr testování**

Protože $\tilde{z}_0 = 0$ náleží do 95% empirického oboustranného intervalu spolehlivosti, tj. $\tilde{z}_0 = 0 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

5. Testování p -hodnotou


```

188 n <-length(acetab.HRL)
189 qbinom(alpha / 2, n, 1 / 2) # 5
190 qbinom(1 - alpha / 2, n, 1 / 2) + 1 # 14
191 acetab.HRLs [5] # -0.39
192 acetab.HRLs [14] # 1.02

```

- p -hodnota

$$\begin{aligned}
 p\text{-hodnota} &= 2 \min\{\Pr(S_E \leq s_E), \Pr(S_E \geq s_E)\} \\
 &= 2 \min\{\Pr(S_E \leq 10), \Pr(S_E \geq 10)\} \\
 &= 2 \min\{\Pr(S_E \leq 10), 1 - \Pr(S_E \leq 9)\} \\
 &= 2 \min\{0.7596588, 0.4072647\} \\
 &= 2 \times 0.4072647 = 0.8145294 \doteq 0.8145
 \end{aligned}$$

```

193 2 * min (pbinom(SE, m, 1 / 2), 1 - pbinom(SE - 1, m, 1 / 2)) # 0.8145294

```

- Závěr testování

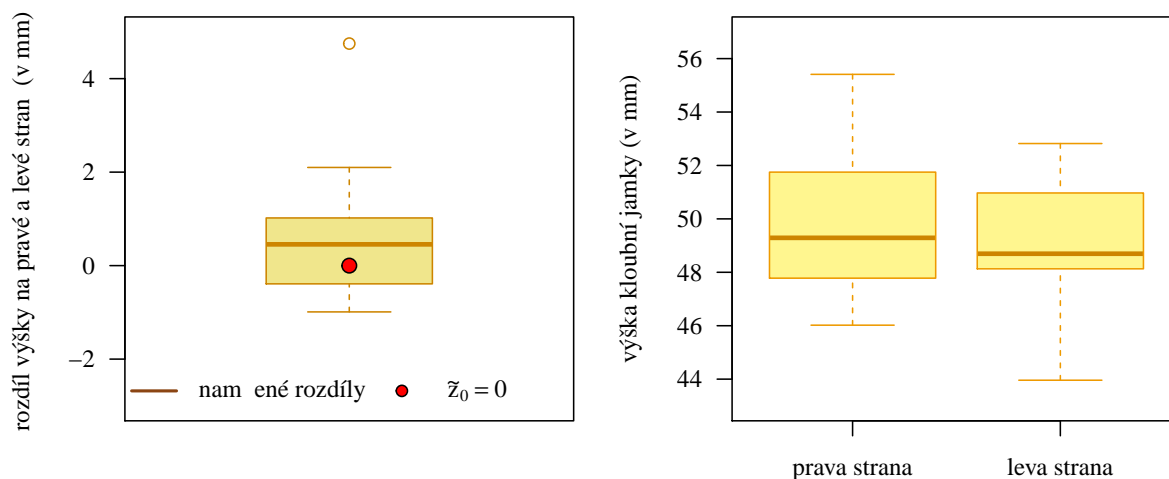
Protože p -hodnota = 0.8145 je větší než $\alpha = 0.05$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

6. Interpretace výsledků

Za základě všech tří typů testování nezamítáme nulovou hypotézu na hladině významnosti $\alpha = 0.05$. Mezi největší výškou kloubní jamky na pravé a levé straně u mužů starověké egyptské populace neexistuje statisticky významný rozdíl.

7. Grafická vizualizace výsledku testování

Porovnání měření na pravé a levé straně zobrazíme nejlépe pomocí krabicového diagramu. Analogicky jako v sekci ?? můžeme buď sestrojít krabicový diagram rozdílů mezi výškami kloubní jamky na pravé a levé straně a porovnat je s konstantou $\tilde{z}_0 = 0$ (viz obrázek 14 vlevo) nebo sestrojít krabicový diagram zvlášť pro výšky kloubní jamky na pravé straně a zvlášť pro výšky kloubní jamky na levé straně a porovnat tyto diagramy navzájem (viz obrázek 14 vpravo).



Obrázek 14: Krabicový diagram rozdílů mezi největší výškou kloubní jamky u mužů na pravé a levé straně

Poznámka: Znaménkový párový test můžeme provést pomocí funkce `SIGN.test()` implementované v knihovně `BSDA`. Vstupními parametry budou vektor naměřených hodnot výšek kloubní jamky na pravé straně (`acetab.HR`), vektor naměřených hodnot výšek kloubní jamky na levé straně (`acetab.HL`), argument `paired = T` určující, že oba vektory považujeme za párová pozorování, hodnota hladiny významnosti α zadaná prostřednictvím koeficientu spolehlivosti $1 - \alpha$ nastavením hodnoty argumentu `conf.level = 0.95` a typ zvolené alternativní hypotézy (oboustranná) zadaný pomocí argumentu `alternative = 'two.sided'`. Součástí výstupu je hodnota mediánu rozdílů naměřených výšek na

```
194 BSDA::SIGN.test(acetab.HR, acetab.HL, paired = T,
195                 conf.level = 0.95, alternative = 'two.sided')
```

```

      Dependent-samples Sign-Test
data:  acetab.HR and acetab.HL
S = 10, p-value = 0.8145
alternative hypothesis: true median difference is not equal to 0
95 percent confidence interval:
 -0.3812269  0.9849076
sample estimates:
median of x-y
      0.455

Achieved and Interpolated Confidence Intervals:

      Conf.Level  L.E.pt U.E.pt
Lower Achieved CI   0.9037 -0.3600 0.9000
Interpolated CI    0.9500 -0.3812 0.9849
Upper Achieved CI   0.9691 -0.3900 1.0200
```

196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213

pravé a levé straně `median of x-y = 0.455`, hodnota testovací statistiky `S = 10`, interpolované hranice 95% empirického oboustranného intervalu spolehlivosti `-0.3812269` a `0.9849076` pro rozdíl naměřených výšek na pravé a levé straně, které jsou mírně přesnější, než námi stanovené hranice intervalu spolehlivosti (zpřesnění hranic bylo provedeno procesem nazývaným *interpolace*), a p -hodnota `p-value = 0.8145`. Jediné, co musíme stanovit zvlášť, jsou dolní a horní hranice kritického oboru.

Druhou možností provedení párového testu je opět pomocí funkce `SIGN.test()`, kde vstupními parametry budou vektor rozdílů naměřených hodnot na pravé a levé straně (`acetab.HRL`), argument `md = 0` určující, že rozdíly porovnáváme s konstantou $\tilde{z}_0 = 0$, hodnota hladiny významnosti α zadaná prostřednictvím koeficientu spolehlivosti $1 - \alpha$ (`conf.level = 0.95`) a typ zvolené alternativní hypotézy (`alternative = 'two.sided'`). Výstup tohoto příkazu je

```
214 BSDA::SIGN.test(acetab.HRL, md = 0, conf.level = 0.95, alternative = 'two.sided')
```

totožný s výše uvedeným výstupem. Záleží tedy na nás, jakou syntaxi k zadání exaktního znaménkového párového testu použijeme. ★

One-sample Sign-Test				215
data: acetab.HRL				216
s = 10, p-value = 0.8145				217
alternative hypothesis: true median is not equal to 0				218
95 percent confidence interval:				219
-0.3812269 0.9849076				220
sample estimates:				221
median of x				222
0.455				223
Achieved and Interpolated Confidence Intervals:				224
	Conf.Level	L.E.pt	U.E.pt	225
Lower Achieved CI	0.9037	-0.3600	0.9000	226
Interpolated CI	0.9500	-0.3812	0.9849	227
Upper Achieved CI	0.9691	-0.3900	1.0200	228
				229
				230
				231
				232

Příklad 9.8. Znaménkový párový exaktní test

Mějme datový soubor 21-goldman-measures.csv obsahující údaje o délkách lýtkových kostí v mm na pravé straně (tibia.LR) a na levé straně (tibia.LL) u skeletů obyvatel útesových obydlí (Cliff Dwellings) v Utahu (detaily viz sekce ??). Na hladině významnosti $\alpha = 0.05$ testujte nulovou hypotézu, že délka lýtkové kosti z pravé strany u skeletů mužského pohlaví je menší než délka lýtkové kosti z levé strany.

Řešení příkladu 9.8

Načteme datový soubor a pomocí operátoru [] vybereme z datové tabulky údaje o délce lýtkové kosti z pravé strany (tibia.LR), resp. z levé strany (tibia.LL) u mužů (sex == 'm') z populace obyvatel útesových obydlí (pop == 'Cliff Dweller'). Z vektoru naměřených hodnot následně odstraníme chybějící hodnoty a zjistíme rozsah náhodného výběru.

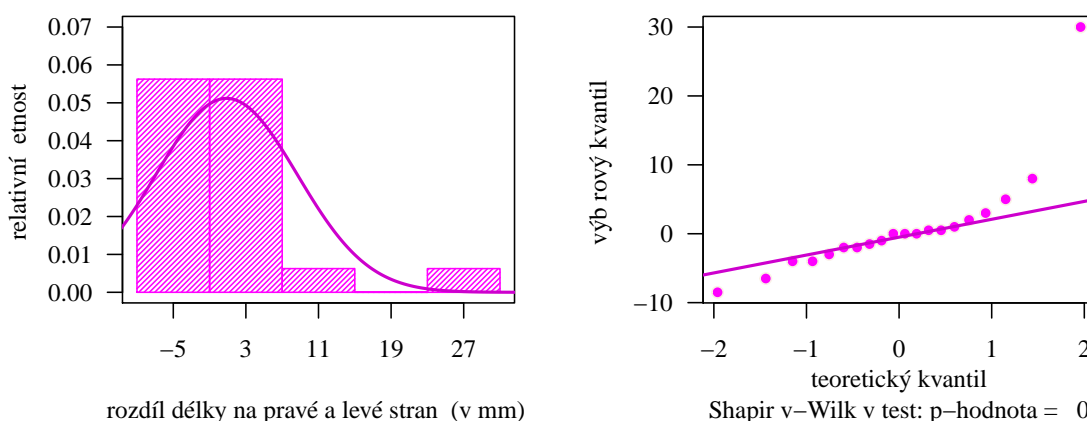
```
233 data <- read.delim('00-Data//21-goldman-measures.csv', sep = ';')
234 # head(data)
235 data.TRL <- data[data$pop == 'Cliff Dweller' & data$sex == 'm',
236                 c('tibia.LR', 'tibia.LL')]
237 data.TRL <- na.omit(data.TRL)
238 dim(data.TRL) # 20
```

Datový soubor obsahuje kompletní údaje o délce lýtkových kostí u 26 mužů z populace obyvatel útesových obydlí na území Utahu.

Naším úkolem ze zadání je porovnat naměřené hodnoty na pravé a levé straně. Jde tedy o měření stejného znaku (délka lýtkové kosti) sledovaného na stejných subjektech (muži). Proto použijeme pro tuto situaci párový test. Prvním krokem k použití párového testu je vytvoření rozdílů hodnot naměřených na pravé a levé straně.

```
239 tibia.LR <- data.TRL$tibia.LR
240 tibia.LL <- data.TRL$tibia.LL
241 tibia.LRL <- tibia.LR - tibia.LL
```

V druhém kroku je potřeba ověřit normální charakter těchto rozdílů. Tímto předpokladem jsme se zabývali již v rámci příkladu ?? v sekci ??, kde jsme zjistili, že náhodný výběr rozdílů mezi délkami lýtkových kostí na pravé a levé straně nepochází z normálního rozdělení (p -hodnota < 0.0001). Pro připomenutí viz obrázek 15.



Obrázek 15: Histogram a diagram rozdílů mezi délkou lýtkové kosti na levé a pravé straně u mužů z populace obyvatel skalních obydlí v Utahu

Z histogramu na obrázku 15 vidíme, že rozdíly mezi pravou a levou stranou jsou výrazně vyšší doleva s

odlehlym pozorováním na pravé straně. Z QQ-diagramu potom vidíme, že toho odlehle pozorování je extrémně vzdálené od referenční přímky. Z příkladu ?? již víme, že toho pozorování narušuje fatálním způsobem normalitu náhodného výběru. V příkladu ?? jsme tuto situaci vyřešili odstraněním odlehleho pozorování a následným provedením parametrického párového testu. Zde budeme pracovat s předpokladem, že naměřené údaje jsou správné a pozorování v souboru ponecháme. Protože však náhodný výběr rozdílů nepochází z normálního rozdělení, nemůžeme hypotézu ze zadání otestovat pomocí parametrického testu, analogicky jako v příkladu ?. K otestování použijeme neparametrický exaktní znaménkový párový test. V zadání příkladu máme explicitně uvedeno, že máme testovat nulovou hypotézu o menší délce lýtkové kosti z pravé strany vzhledem k délce lýtkové kosti z levé strany. To odpovídá tvrzení, že rozdíly získané odečtením naměřených hodnot na levé straně od naměřených hodnot na pravé straně budou menší než nula. Zbývá dodefinovat alternativní hypotézu.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy
 H_0 : Medián rozdílů délky pažní kosti na pravé a levé straně u mužů je menší nebo roven nule.
 H_1 : Medián rozdílů délky pažní kosti na pravé a levé straně u mužů je větší než nula.
- **matematická formulace** nulové a alternativní hypotézy
 H_0 : $\tilde{z} \leq \tilde{z}_0$, kde $\tilde{z}_0 = 0$
 H_1 : $\tilde{z} > \tilde{z}_0$, kde $\tilde{z}_0 = 0$ (pravostranná alternativa)

2. Volba hladiny významnosti

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

3. Testování kritickým oborem

- **Testovací statistika**
V úvodu příkladu jsme vytvořili vektor `tibia.LRL` ($Z_i, i = 1, \dots, 20$) jako vektor rozdílů mezi délkami lýtkových kostí na pravé a levé straně. Nyní je potřeba stanovit rozdíly $Z_i - \tilde{z}_0$. Protože ale $\tilde{z}_0 = 0$, jsou rozdíly $Z_i - \tilde{z}_0$ rovny původnímu vektoru rozdílů Z_i . Hodnoty délky lýtkové kosti na pravé straně (X_i), resp. na levé straně (Y_i), rozdíly Z_i , rozdíly $Z_i - \tilde{z}_0$ a znaménko posledních uvedených rozdílů jsou pro názornost uvedeny v tabulce 9.

Tabulka 9: Naměřené hodnoty X_i, Y_i , rozdíly $Z_i = X_i - Y_i$, rozdíly $Z_i - \tilde{z}_0$ a znaménka těchto rozdílů

měření	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
X_i	376	370	346	388	380	335	372.5	388	364.0	361	370	396.5	373.0	378	368	370	374	356.0	355	337
Y_i	379	370	350	386	380	334	372.0	388	365.5	362	374	396.0	379.5	370	365	372	376	364.5	350	307
Z_i	-3	0	-4	2	0	1	0.5	0	-1.5	-1	-4	0.5	-6.5	8	3	-2	-2	-8.5	5	30
$Z_i - \tilde{z}_0$	-3	0	-4	2	0	1	0.5	0	-1.5	-1	-4	0.5	-6.5	8	3	-2	-2	-8.5	5	30
+/-	-	-	-	+	-	+	+	-	-	-	-	+	-	+	+	-	-	-	+	+

Z tabulky 9 vidíme, že celkem osm rozdílů $Z_i - \tilde{z}_0, i = 1, \dots, 20$, je kladných. Hodnota testovací statistiky S_E bude tedy rovná 8.

$$S_E = \sum_{i=1}^n I(Z_i - \tilde{z}_0 > 0) = \sum_{i=1}^{20} I(Z_i - 0 > 0) = 8.$$

```

242 z0 <- 0
243 I <- (tibia.LRL > z0)
244 tab <- data.frame(rbind("Xi" = tibia.LR, "Yi" = tibia.LL,
245                        "Zi" = tibia.LRL, "Zi-z0" = tibia.LRL - z0, "+/-" = I))
246 names(tab) <- 1 : 20
247 SE <- sum(I) # 8

```

- **Kritický obor**

Z tabulky 9 dále vidíme, že z celkového počtu $n = 20$ rozdílů jsou tři rozdílů nulové, a tedy počet nenulových rozdílů $m = 17$. Kritický obor bude potom tvořen hodnotou kvantilu $b_{m,1/2}(1 - \alpha)$ a nekonečnem.

$$\begin{aligned} W &= \langle b_{m,1/2}(1 - \alpha); \infty \rangle \\ &= \langle b_{17,1/2}(1 - 0.05); \infty \rangle \\ &= \langle b_{17,1/2}(0.95); \infty \rangle \\ &= \langle 12; \infty \rangle \end{aligned}$$

```
248 n <- length(tibia.LRL)
249 m <- sum(tibia.LRL - z0 != 0) # 17
250 alpha <- 0.05
251 q <- qbinom(1 - alpha, m, 1 / 2) # 12
```

- **Závěr testování**

Protože realizace testovací statistiky $s_E = 8$ nenáleží do kritického oboru, tj. $s_E \notin W$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

4. Testování intervalem spolehlivosti

- **Interval spolehlivosti**

Abychom mohli stanovit hranice 95% intervalu spolehlivosti, musíme nejprve rozdílů mezi naměřenými délkami pravé a levé strany vzestupně seřadit. Seřazené rozdílů jsou k dispozici v tabulce 10.

Tabulka 10: Seřazené rozdílů mezi délkami lýtkových kostí na pravé a levé straně

pořadí	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
seřazené Z_i	-8.5	-6.5	-4.0	-4.0	-3.0	-2.0	-2.0	-1.5	-1.0	0.0	0.0	0.0	0.5	0.5	1.0	2.0	3.0	5.0	8.0	30.0

Hranice intervalu spolehlivosti potom tvoří hodnota, která se v seřazeném vektoru hodnot nachází na $(b_{n,1/2}(\alpha))$ -té pozici, a nekonečno.

$$\begin{aligned} (d, h) &= \left(Z^{(b_{n,1/2}(\alpha))}; \infty \right) \\ &= \left(Z^{(b_{17,1/2}(0.05))}; \infty \right) \\ &= \left(Z^{(6)}; \infty \right) \\ &= (-2; \infty) \end{aligned}$$

```
252 qbinom(alpha, n, 1 / 2) # 6
253 tibia.LRLs[6] # -2
```

- **Závěr testování**

Protože $\tilde{z}_0 = 0$ náleží do 95% empirického jednostranného intervalu spolehlivosti, tj. $\tilde{z}_0 = 0 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

5. Testování p -hodnotou

- **p -hodnota**

$$p\text{-hodnota} = \Pr(S_E \geq s_E) = \Pr(S_E \geq 8) = 1 - \Pr(S_E < 8) = 1 - \Pr(S_E \leq 7) = 0.6854706 \doteq 0.6855$$

- **Závěr testování**

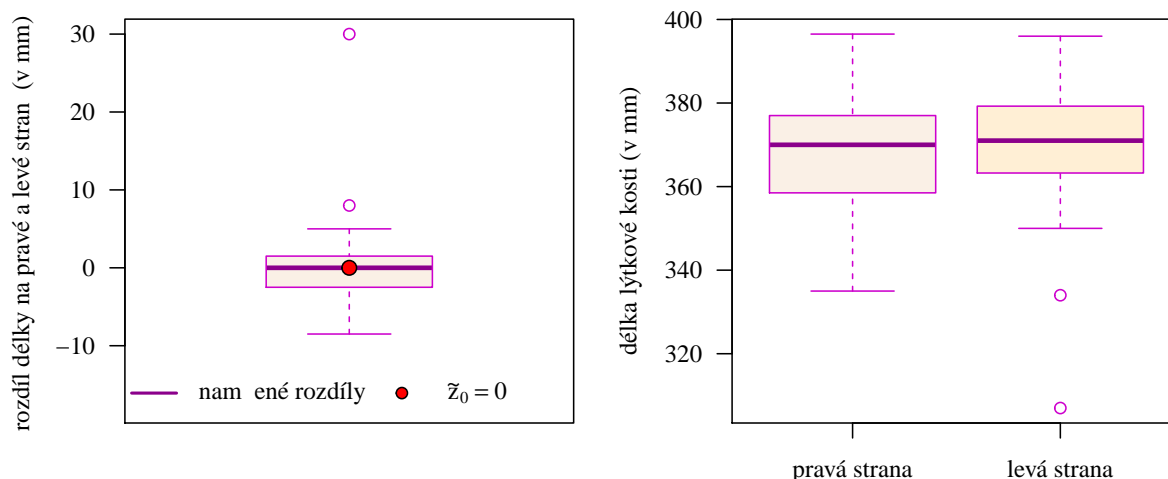
Protože p -hodnota = 0.6855 je větší než $\alpha = 0.05$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

6. Interpretace výsledků

Na základě všech tří typů testování na hladině významnosti $\alpha = 0.05$ nezamítáme nulovou hypotézu, že délka lýtkové kosti z pravé strany u skeletů mužského pohlaví je menší než délka lýtkové kosti z levé strany. Uvědomme si, že zamítnutí nulové hypotézy by vedlo k závěru, že délka lýtkové kosti z pravé strany je statisticky významně větší než délka lýtkové kosti z levé strany. Nezamítnutí nulové hypotézy tedy vede k závěru, že délka lýtkové kosti z pravé strany není statisticky významně větší než délka lýtkové kosti z levé strany.

7. Grafická vizualizace výsledku testování

Porovnání měření na pravé a levé straně vizualizujeme pomocí krabicového diagramu, přičemž opět se můžeme rozhodnout, zda dáme přednost diagramu porovnávajícímu rozdíly mezi pravou a levou stranou s konstantou $\tilde{z}_0 = 0$ (obrázek 16 vlevo), nebo diagramu porovnávajícímu naměřené hodnoty na pravé straně s hodnotami naměřenými na levé straně (obrázek 16 vpravo).



Obrázek 16: Krabicový diagram rozdílů mezi délkou lýtkové kosti na levé a pravé straně u mužů z populace obyvatel skalních obydlí v Utahu

Poznámka: Znaménkový exaktní párový test provedeme pomocí funkce `SIGN.test()` z knihovny `BSDA`. Vstupními parametry budou vektor naměřených hodnot délek lýtkové kosti na pravé straně (`tibia.LR`), vektor naměřených hodnot délek lýtkové kosti na levé straně (`tibia.LL`), specifikace párového testu (`paired = T`), hodnota hladiny významnosti α zadaná prostřednictvím koeficientu spolehlivosti $1 - \alpha$ (`conf.level = 0.95`) a pravostranný typ alternativní hypotézy (`alternative = 'greater'`). Součástí výstupu je hodnota mediánu rozdílů naměřených výšek na pravé a levé straně

```
255 BSDA::SIGN.test(tibia.LR, tibia.LL, paired = T, conf.level = 0.95,
256 alt = 'greater') # IS interpolovany
```

median of $x-y = 0$, hodnota testovací statistiky $s = 8$, interpolované hranice 95% empirického jednostranného intervalu spolehlivosti -2 a Inf pro rozdíl mezi délkami na pravé a levé straně a p -hodnota $p\text{-value} = 0.6855$. Jediné, co musíme stanovit zvlášť, je dolní hranice kritického oboru. ★

Příklad 9.9. Znaménkový párový exaktní test

Mějme datový soubor 21-goldman-measures.csv obsahující údaje o anteroposteriorním průměru hlavice stehenní kosti v mm na pravé straně (femur.HDR) a na levé straně (femur.HDL) u skeletů aljašské populace kmene Ipituaq (detaily viz sekce ??). Na hladině významnosti $\alpha = 0.10$ zjistěte, zda je u skeletů ženského pohlaví anteroposteriorní průměr hlavice stehenní kosti z levé strany menší než z pravé strany.

Řešení příkladu 9.9

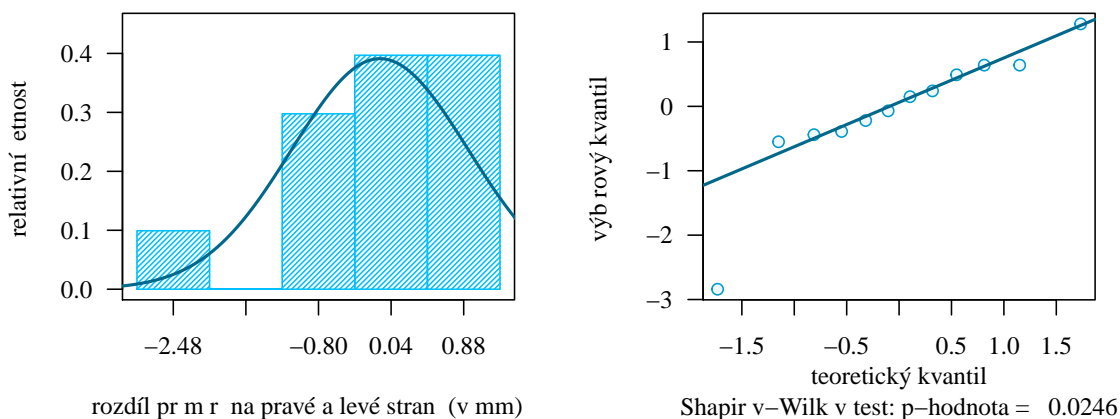
Načteme datový soubor a pomocí operátoru `[]` vybereme z datové tabulky údaje o délce pažní kosti z levé strany (humer.LL), resp. z pravé strany (humer.LR) u žen (`sex == 'f'`) z římského pohřebiště v Poundbury (`pop == 'Poundbury'`). Z vektoru naměřených hodnot odstraníme chybějící hodnoty a zjistíme rozsah náhodného výběru.

```
257 data <- read.delim('00-Data//21-goldman-measures.csv', sep = ';')
258 head(data)
259 data.HDLR <- data[data$pop == 'Ipituaq' & data$sex == 'f', c('femur.HDR', 'femur.HDL')]
260 data.HDLR <- na.omit(data.HDLR)
261 dim(data.HDLR) # 12x2
```

Datový soubor obsahuje kompletní údaje o anteroposteriorním průměru hlavice stehenní kosti z levé i pravé strany u 12 skeletů žen z aljašské populace kmene Ipituaq.

Naším úkolem ze zadání je porovnat naměřené hodnoty na levé a pravé straně. Jde tedy o měření stejného znaku (anteroposteriorní průměr) sledovaného na stejných subjektech (ženy), proto použijeme pro tuto situaci párový test. Nejprve tedy vytvoříme rozdíly hodnot anteroposteriorních průměrů hlavice stehenní kosti naměřených na levé a pravé straně a následně ověříme normální rozdělení těchto rozdílů. K tomu využijeme Shapiro-Wilkovým testem ($\alpha = 0.05$), QQ-diagram a histogram, přičemž naměřené hodnoty rozdělíme do pěti ekvidistatních intervalů s šířkou 0.84 mm prostřednictvím stanovených hranic $-2.9, -2.06, \dots, 1.30$ (viz obrázek 17).

```
262 femur.HDL <- data.HDLR$femur.HDL
263 femur.HDR <- data.HDLR$femur.HDR
264 femur.HDLR <- femur.HDL - femur.HDR
```



Obrázek 17: Histogram a diagram rozdílů mezi anteroposteriorním průměrem hlavice stehenní kosti na levé a pravé straně u skeletů žen z aljašské populace kmene Ipituaq

Protože p -hodnota = 0.0246 je menší než 0.05, nulovou hypotézu o normalitě rozdílů zamítáme na hladině významnosti $\alpha = 0.05$. Z histogramu na obrázku 17 vidíme, že normalitu pravděpodobně fatálně způsobem ovlivňuje odlehlý

rozdíl pravé a levé strany u desátého měření (-2.48). První, co bychom v takovémto případě měli udělat, je zkontrolovat, zda byla měření na obou stranách zaznamenána správně. Pokud ne, údaj opravíme a je možné, že normalita rozdílů se vyspraví. V našem případě vycházíme z předpokladu, že oba údaje byly zaznamenány správně a tedy odlehlý rozdíl je důsledkem měření na reálných datech. Zemřelý jedinec mohl být například nějakým způsobem hendikepovaný, což mohlo vést k nerovnoměrnému vývoji stehenních kostí, apod. Protože máme pouze 12 kompletních údajů, je v našem zájmu zachovat odlehlé pozorování v datovém souboru a porovnat pravou i levou stranu s pomocí všech 12 měření. Náhodný výběr rozdílů na levé a pravé straně nicméně nepochází z normálního rozdělení, což musíme zohlednit při výběru vhodného párového testu.

Protože náhodný výběr rozdílů nepochází z normálního rozdělení, nemůžeme k ověření otázky ze zadání použít parametrický párový test. Použijeme tedy jeho neparametrickou alternativu, a sice exaktní znaménkový párový test. Naším úkolem je zjistit, zda je u skeletů ženského pohlaví anteroposteriorní průměr hlavice stehenní kosti z levé strany menší než anteroposteriorní průměr hlavice stehenní kosti z pravé strany. Tato věta je zněním alternativní hypotézy, přičemž nulovou hypotézu dodefinujeme jako doplněk k tomuto tvrzení.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : Medián rozdílů anteroposteriorního průměru hlavice stehenní kosti na levé a pravé straně u žen z kmene Ipituaq je větší nebo rovný nule.

H_1 : Medián rozdílů anteroposteriorního průměru hlavice stehenní kosti na levé a pravé straně u žen z kmene Ipituaq je menší než nula.

- **matematická formulace** nulové a alternativní hypotézy

H_0 : $\tilde{z} \geq \tilde{z}_0$, kde $\tilde{z}_0 = 0$

H_1 : $\tilde{z} < \tilde{z}_0$, kde $\tilde{z}_0 = 0$ (levostranná alternativa)

2. Volba hladiny významnosti

- Hladinu významnosti volíme v souladu se zadáním jako $\alpha = 0.10$.

3. Testování kritickým oborem

- **Testovací statistika**

V úvodu příkladu jsme vytvořili vektor femur.HDLR (Z_i) jako vektor rozdílů mezi anteroposteriorním průměrem na levé a pravé straně. Nyní je potřeba stanovit rozdíl $Z_i - \tilde{z}_0$ a určit počet kladných rozdílů $Z_i - \tilde{z}_0$ a počet nenulových rozdílů m . Hodnoty anteroposteriorních průměrů hlavice stehenní kosti na levé straně (X_i), resp. na pravé straně (Y_i), rozdílů Z_i , rozdílů $Z_i - \tilde{z}_0$ a znaménko posledních uvedených rozdílů jsou pro názornost uvedeny v tabulce 11.

Tabulka 11: Naměřené hodnoty X_i , Y_i , rozdílů $Z_i = X_i - Y_i$, rozdílů $Z_i - \tilde{z}_0$ a znaménka těchto rozdílů

měření	1	2	3	4	5	6	7	8	9	10	11	12
X_i	40.58	43.22	38.68	41.47	40.21	41.39	41.27	38.00	44.02	39.91	38.91	40.72
Y_i	41.13	41.94	38.04	41.86	40.43	40.75	41.03	37.85	44.09	42.75	39.35	40.23
Z_i	-0.55	1.28	0.64	-0.39	-0.22	0.64	0.24	0.15	-0.07	-2.84	-0.44	0.49
$Z_i - \tilde{z}_0$	-0.55	1.28	0.64	-0.39	-0.22	0.64	0.24	0.15	-0.07	-2.84	-0.44	0.49
+/-	-	+	+	-	-	+	+	+	-	-	-	+

Z tabulky 11 vidíme, že z celkového počtu 12 rozdílů není žádný rozdíl nulový, tj. $m = 12$ a celkem šest rozdílů je kladných. Hodnota testovací statistiky S_E definované jako počet kladných rozdílů bude tedy rovná šesti.

$$S_E = \sum_{i=1}^n I(Z_i - \tilde{z}_0 > 0) = \sum_{i=1}^{12} I(Z_i - 0 > 0) = 6.$$

```

265 z0 <- 0
266 I <- (femur.HDLR > z0)
267 tab <- data.frame(rbind("Xi" = femur.HDL, "Yi" = femur.HDR,
268 "Zi" = femur.HDLR, "Zi-z0" = femur.HDLR - z0, "+/-" = I))
269 SE <- sum(I) # 6
270 n <- length(femur.HDLR) # 12
271 m <- sum(femur.HDLR - z0 != 0) # 12

```

- **Kritický obor**

$$\begin{aligned}
W &= (-\infty; b_{m,1/2}(\alpha) - 1) \\
&= (-\infty; b_{12,1/2}(0.10) - 1) \\
&= (-\infty; 4 - 1) \\
&= (-\infty; 3)
\end{aligned}$$

```

272 alpha <- 0.10
273 q <- qbinom(alpha, m, 1 / 2) - 1 # 3

```

- **Závěr testování**

Protože realizace testovací statistiky $s_E = 6$ nenáleží do kritického oboru, tj. $s_E \notin W$, H_0 nezamítáme na hladině významnosti $\alpha = 0.10$.

4. Testování intervalem spolehlivosti

- **Interval spolehlivosti**

Abychom mohli stanovit hranice 90% intervalu spolehlivosti, musíme nejprve rozdíly mezi naměřenými anteroposteriorními průměry na levé a pravé straně vzestupně seřadit.

Hranice intervalu spolehlivosti potom tvoří mínus nekonečno a hodnota, která se v seřazeném vektoru hodnot nachází na $(b_{n,1/2}(1 - \alpha) + 1)$ -té pozici.

$$\begin{aligned}
(d, h) &= (-\infty; Z^{(b_{n,1/2}(1-\alpha)+1)}) \\
&= (-\infty; Z^{(b_{12,1/2}(1-0.10)+1)}) \\
&= (-\infty; Z^{(b_{12,1/2}(0.90)+1)}) \\
&= (-\infty; Z^{(9)}) \\
&= (-\infty; 0.49)
\end{aligned}$$

```

274 qbinom(1 - alpha, n, 1 / 2) + 1 # 9
275 femur.HDLRs[9] # 0.49

```

- **Závěr testování**

Protože $\tilde{z}_0 = 0$ náleží do 90% empirického pravostranného intervalu spolehlivosti, tj. $\tilde{x}_0 = 0 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0.10$.

5. Testování p -hodnotou

- p -hodnota

$$p\text{-hodnota} = \Pr(S_E \leq s_E) = \Pr(S_E \leq 6) = 0.612793 \times 0.6128$$

276 `pbinom(SE, m, 1 / 2) # 0.612793`

- **Závěr testování**

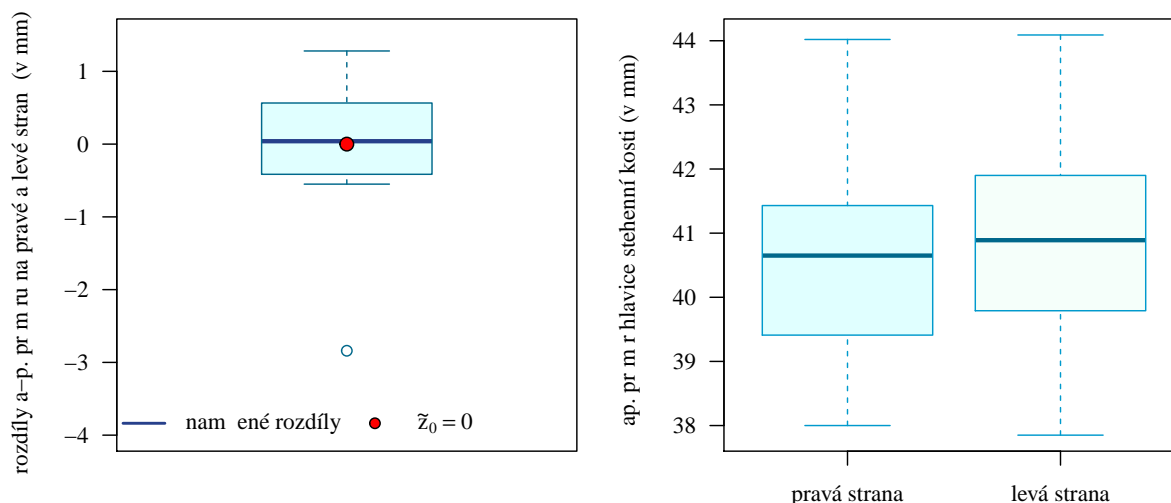
Protože p -hodnota = 0.6128×10^{-7} je větší než $\alpha = 0.10$, H_0 nezamítáme na hladině významnosti $\alpha = 0.10$.

6. Interpretace výsledků

Na základě všech tří typů testování nezamítáme nulovou hypotézu na hladině významnosti $\alpha = 0.10$. Anteroposteriorní průměr hlavice stehenní kosti na levé straně u žen z aljašské populace jmene Ipituaq není statisticky významně menší než na pravé straně.

7. Grafická vizualizace výsledku testování

Rozdíl v anteroposteriorním průměru hlavice stehenní kosti na pravé a levé straně vizualizujeme pomocí krabicového diagramu (viz obrázek 18).



Obrázek 18: Krabicový diagram rozdílů mezi anteroposteriorním průměrem hlavice stehenní kosti na pravé a levé straně u skeletů žen z aljašské populace kmene Ipituaq

Poznámka: Znaménkový párový test můžeme provést pomocí funkce `SIGN.test()`. Vstupními parametry budou vektor naměřených hodnot anteroposteriorních průměrů hlavice stehenní kosti na levé straně (`femur.HDL`), vektor naměřených hodnot anteroposteriorních průměrů hlavice stehenní kosti na pravé straně (`femur.HDR`), volba párového testu (`paired = T`), hodnota hladiny významnosti α zadaná prostřednictvím koeficientu spolehlivosti $1 - \alpha$ (`conf.level = 0.90`) a levostranný typ alternativní hypotézy (`alternative = 'less'`). Součástí výstupu je hodnota

277 `BSDA::SIGN.test(femur.HDL, femur.HDR, paired = T,`
 278 `conf.level = 0.90, alternative = 'less')`

mediánu rozdílů naměřených anteroposteriorních průměrů na levé a pravé straně `median of x-y = 0.04`, hodnota testovací statistiky $S = 6$, interpolované hranice 90% empirického pravo intervalu spolehlivosti `-lnf a 0.4341414` pro rozdíl naměřených anteroposteriorních průměrů na levé a pravé straně a p -hodnota `p-value = 0.6128`. Jediné, co musíme stanovit zvlášť, je horní hranice kritického oboru. ★

Dependent-samples Sign-Test	279
data: femur.HDL and femur.HDR	280
S = 6, p-value = 0.6128	281
alternative hypothesis: true median difference is less than 0	282
90 percent confidence interval:	283
-Inf 0.4341414	284
sample estimates:	285
median of x-y	286
0.04	287
Achieved and Interpolated Confidence Intervals:	288
	289
	290
	291
	292
	293
	294
	295
	296

Příklad 9.10. Znaménkový párový asymptotický test

Máme datový soubor 21-goldman-measures.csv obsahující údaje o epikondylární šířce pažní kosti na pravé straně (humer.EBR) a na levé straně (humer.EBL) u mužů (sex = 'm' a sex == 'm?') a žen (sex = 'f' a sex == 'f?') pohřbených v oblasti Indian Knoll v Kentucky (viz sekce ??). Na hladině významnosti $\alpha = 0.01$ zjistíte, zda se střední hodnota epikondylární šířky pažní kosti u mužů na pravé straně a levé straně liší. K testování využijete všechna kompletní měření týkající se mužů (skupiny muž (sex == 'm') i pravděpodobně muž (sex == 'm?')).

Řešení příkladu 9.10

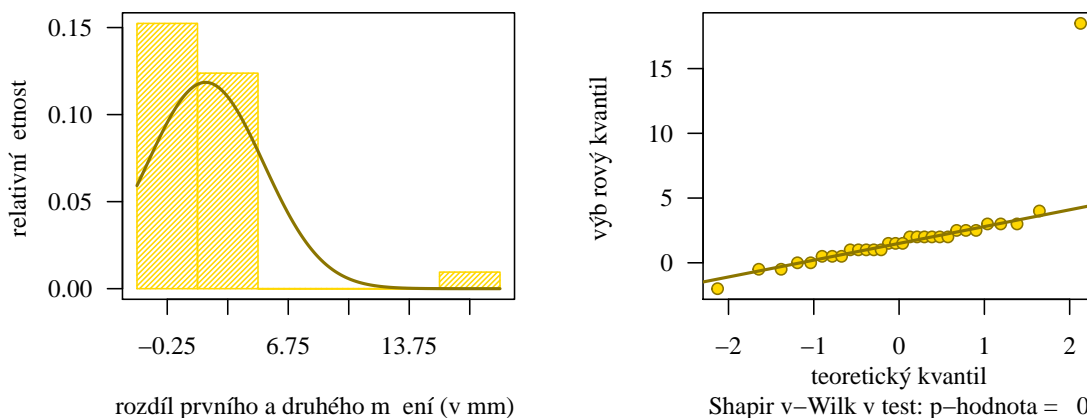
Načteme datový soubor a pomocí operátoru [] z něj vybereme pouze údaje o epikondylární šířce pažní kosti na pravé straně (humer.EBR) a na levé straně (humer.EBL) týkající se pouze mužů, přičemž v souladu se zadáním uvažujeme obě skupiny mužů (sex %in% c('m', 'm?')), z pohřebiště Indian Knoll v Kentucky (pop == 'Indian Knoll'). Údaje vložíme do proměnné data.m. Z datové tabulky následně odstraníme chybějící údaje a zjistíme rozsah náhodného výběru.

```
297 data <- read.delim('00-Data//21-goldman-measures.csv', sep = ';')
298 #head(data)
299 data.m <- data[data$sex %in% c('m', 'm?') & data$pop == 'Indian Knoll',
300             c('humer.EBL', 'humer.EBR')]
301 data.m <- na.omit(data.m)
302 dim(data.m) # 30 x 2
```

Datový soubor obsahuje kompletní údaje o epikondylární šířce pažní kosti na pravé i levé straně u 30 mužů z populace z oblasti Indian Knoll. Jelikož máme za úkol porovnat párový znak, tj. měření provedená na jednom subjektu jak z pravé strany tak z levé strany, použijeme k tomuto porovnání párový test. Prvním krokem k provedení tohoto testu je vytvoření rozdílů hodnot naměřených na pravé a levé straně. V druhém kroku je potřeba ověřit předpoklad normality těchto rozdílů.

```
303 humer.EBR <- data.m$humer.EBR
304 humer.EBL <- data.m$humer.EBL
305 humer.EBRL <- humer.EBR - humer.EBL
```

Rozsah náhodného výběru ($n = 30$) je přesně na horní hranici použití Shapirova-Wilkova testu ($\alpha = 0.05$) k ověření předpokladu normality rozdílů. Graficky zhodnotíme normalitu QQ-diagramem a histogramem. Naměřené hodnoty rozdělíme do šesti ekvidistantních intervalů se šířkou 3.5 mm prostřednictvím stanovených hranic $-2, 1.5, \dots, 19$ (viz obrázek 19).



Obrázek 19: Histogram a QQ-diagram rozdílů epikondylární šířky pažní kosti na pravé a levé straně

Protože p -hodnota vyšla < 0.0001 , což je menší než 0.05, nulovou hypotézu o normalitě rozdílů zamítáme na hladině významnosti $\alpha = 0.05$. Náhodný výběr rozdílů tedy nepochází z normálního rozdělení. Z histogramu a QQ-diagramu je patrné, že normalitu rozdílů fatálním způsobem ovlivňuje extrémní rozdíl s hodnotou 18.49. Ten je způsoben mimořádně nízkou hodnotou epikondylární šířky (36.51 mm) naměřené na levé straně u v pořadí 29. muže. Můžete si vyzkoušet, že po odstranění této hodnoty by náhodný výběr rozdílů předpoklad normality splňoval. V praxi, kde bychom epikondylární šířky měřili sami, bychom nejprve zkontrolovali, zda byl rozměr změřený správně, či zda nedošlo k chybě při záznamu hodnoty. Pokud bychom chybu v měření nezaznamenali, máme dvě možnosti. První možností je pozorování z výběru odstranit, a získat tak soubor 29 kompletních měření z pravé a levé strany. Nicméně tím náhodný výběr idealizujeme, protože vyhadujeme byť extrémní, ale správně naměřené pozorování. Datový soubor tím navíc ztrácí charakter náhodného výběru. Navíc v praxi často nemáme tolik pozorování, abychom si mohli dovolit libovolné z nich ze souboru odstranit. Proto zde využijeme druhou možnost, která znamená ponechání extrémního pozorování v datovém souboru a zvolení vhodné neparametrické metody párového testu, která zohledňuje porušení předpokladu normality nohodného výběru rozdílů mezi pravou a levou stranou.

Otázku ze zadání tedy ověříme pomocí neparametrického znaménkového párového testu. Vzhledem k rozsahu náhodného výběru použijeme jeho asymptotickou variantu. Naším úkolem je zjistit, zda se střední hodnota epikondylární šířky pažní kosti u mužů na pravé straně a levé straně liší. Toto tvrzení bude zněním alternativní hypotézy, neboť odlišnost implikuje nerovnost a nerovnost je vždy součástí alternativní hypotézy. Nulová hypotéza tedy bude naopak tvrdit, že střední hodnota epikondylární šířky pažní kosti u mužů na pravé straně a levé straně se neliší, tedy že epikondylární šířky pažní kosti u mužů na pravé straně a levé straně jsou shodné. Dále je potřeba si uvědomit, že možnost testovat hypotézu o středních hodnotách jsme ztratili při volbě neparametrického testu. Namísto toho budeme testovat nulovou hypotézu o shodě mediánu epikondylární šířky na pravé straně s mediánem epikondylární šířky na levé straně, což je analogie nulové hypotézy, že medián rozdílů měření na pravé a levé straně je rovný nule.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : Medián rozdílů epikondylární šířky u mužů o oblasti Indian Knoll na pravé a levé straně je rovný nule.

H_1 : Medián rozdílů epikondylární šířky u mužů o oblasti Indian Knoll na pravé a levé straně není rovný nule.

- **matematická formulace** nulové a alternativní hypotézy

H_0 : $\tilde{z} = \tilde{z}_0$, kde $\tilde{z}_0 = 0$

H_1 : $\tilde{z} \neq \tilde{z}_0$, kde $\tilde{z}_0 = 0$ (oboustranná alternativa)

2. Volba hladiny významnosti

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

3. Testování kritickým oborem

- **Testovací statistika**

V úvodu příkladu jsme vytvořili vektor `humer.EBRL` (Z_i) jako vektor rozdílů mezi pravou a levou stranou. Nyní je potřeba stanovit rozdíl $Z_i - \tilde{z}_0$ a určit počet kladných rozdílů $Z_i - \tilde{z}_0$ a počet nenulových rozdílů m .

```
306 z0 <- 0
307 I <- (humer.EBRL > z0)
308 SE <- sum(I) # 25
309 n <- length(humer.EBRL) # 30
310 m <- sum(humer.EBRL - z0 != 0) # 28
```

Z celkového počtu 30 rozdílů jsou dva rozdílů nulové, tj. počet nenulových rozdílů $m = 28$, a 25 rozdílů je

kladných ($S_E = 25$).

$$S_E = \sum_{i=1}^n I(Z_i - \tilde{z}_0 > 0) = \sum_{i=1}^{30} I(Z_i - 0 > 0) = 25.$$

Nyní vypočítáme testovací statistiku asymptotické varianty znaménkového testu podle vzorce 9.2.

$$S_A = \frac{S_E - \frac{m}{2}}{\sqrt{\frac{m}{4}}} = \frac{25 - \frac{28}{2}}{\sqrt{\frac{28}{4}}} = \frac{0}{25 - 14} \sqrt{7} = \frac{11}{2.645751} = 4.15761 \doteq 4.1576$$

```
311 SA <- (SE - m / 2) / sqrt(m / 4) # 4.157609
```

• Kritický obor

$$\begin{aligned} W &= (-\infty; u_{\alpha/2}) \cup (u_{1-\alpha/2}; \infty) \\ &= (-\infty; u_{0.01/2}) \cup (u_{1-0.01/2}; \infty) \\ &= (-\infty; u_{0.005}) \cup (u_{0.995}; \infty) \\ &= (-\infty; -2.5758) \cup (2.5758; \infty) \end{aligned}$$

```
312 alpha <- 0.01
313 qnorm(alpha / 2) # -2.575829
314 qnorm(1 - alpha / 2) # 2.575829
```

• Závěr testování

Protože realizace testovací statistiky $s_A = 4.1576$ náleží do kritického oboru, tj. $s_A \in W$, H_0 zamítáme na hladině významnosti $\alpha = 0.01$.

4. Testování intervalem spolehlivosti

• Interval spolehlivosti

Pomocí příkazu `sort()` nejprve vzestupně seřadíme naměřené hodnoty Z_i , $i = 1, \dots, 30$. Hranice 99% intervalu spolehlivosti budou potom $(C_{1-\alpha/2})$ -tá a $(n + 1 - C_{1-\alpha/2})$ -tá hodnota v seřazeném vektoru naměřených hodnot, kde

$$\begin{aligned} C_{1-\alpha/2} &= \frac{n}{2} - u_{1-\alpha/2} \sqrt{\frac{n}{4}} \\ &= \frac{30}{2} - u_{1-0.01/2} \sqrt{\frac{30}{4}} \\ &= 15 - u_{0.995} \sqrt{7.5} \\ &= 15 - 2.575829 \times 2.738613 \\ &= 7.945801 \doteq 8 \end{aligned}$$

Interval spolehlivosti má potom následující tvar

$$\begin{aligned}(d, h) &= \left(X^{(C_{1-\alpha/2})}; X^{(n+1-C_{1-\alpha/2})} \right) \\ &= \left(X^{(8)}; X^{(30+1-8)} \right) \\ &= \left(X^{(8)}; X^{(23)} \right) \\ &= (0.5; 2.5)\end{aligned}$$

```
315 humer.EBRLs <- sort(humer.EBRL)
316 C1 <- round(n / 2 - qnorm(1 - alpha / 2) * sqrt(n / 4)) # 8
317 C2 <- round(n + 1 - C1) # 23
318 humer.EBRLs[C1] # 0.5
319 humer.EBRLs[C2] # 2.5
```

- **Závěr testování**

Protože $\tilde{z}_0 = 0$ nenáleží do 99% empirického oboustranného intervalu spolehlivosti, tj. $\tilde{z}_0 = 0 \in IS$, H_0 zamítáme na hladině významnosti $\alpha = 0.01$.

5. Testování p -hodnotou

- p -hodnota

$$\begin{aligned}p\text{-hodnota} &= 2 \min\{\Pr(S_A \leq s_A), \Pr(S_A \geq s_A)\} \\ &= 2 \min\{\Pr(S_A \leq 4.157609), \Pr(S_A \geq 4.157609)\} \\ &= 2 \min\{\Pr(S_A \leq 4.157609), 1 - \Pr(S_A \leq 4.157609)\} \\ &= 2 \min\{0.9999839, 1.607978 \times 10^{-5}\} \\ &= 2 \times 1.607978 \times 10^{-5} = 3.215956 \times 10^{-5} \doteq 3.2160 \times 10^{-5}\end{aligned}$$

```
320 p.hodnota <- 2 * min(pnorm(SA), 1 - pnorm(SA)) # 3.215956e-05
```

- **Závěr testování**

Protože p -hodnota = 3.2160×10^{-5} je menší než $\alpha = 0.01$, H_0 zamítáme na hladině významnosti $\alpha = 0.01$.

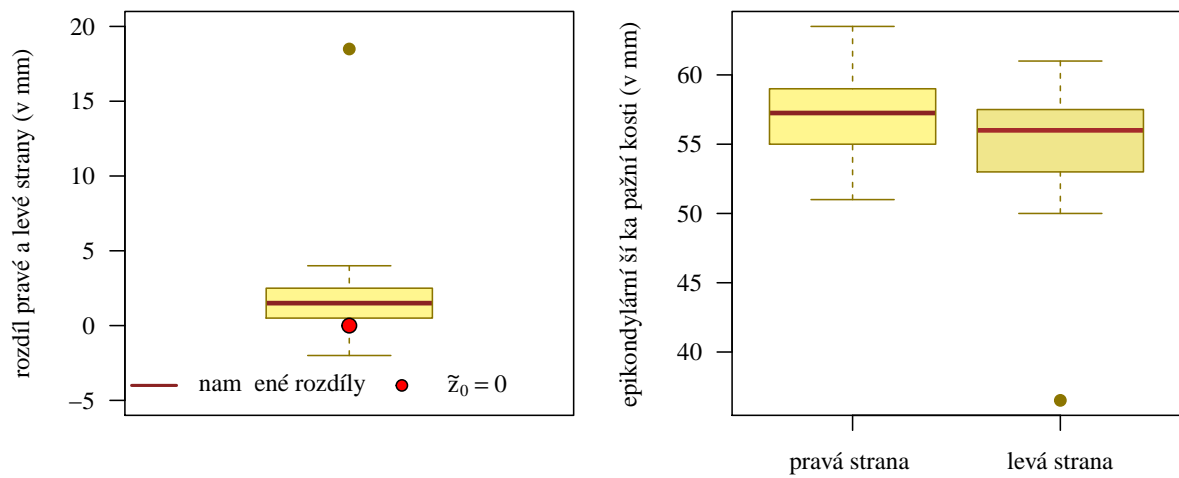
6. Interpretace výsledků

Za základě všech tří typů testování zamítáme nulovou hypotézu na hladině významnosti $\alpha = 0.01$. Mezi epikondylární šířkou na pravé straně a na levé straně u mužů z oblasti Indian Knoll existuje statisticky významný rozdíl.

7. Grafická vizualizace výsledku testování

Rozdíl v epikondylárních šířkách z pravé a z levé straně můžeme vizualizovat pomocí krabicového diagramu (viz obrázek 20).





Obrázek 20: Krabicový diagram rozdílů epikondylární šířky u mužů z oblasti Indian Knoll na pravé a levé straně

Příklad 9.11. Wilcoxonův párový asymptotický test

Máme datový soubor 02-paired-means-clavicle.txt obsahující údaje o hodnotách vertikálního průměru středu délky těla klíční kosti z pravé a levé strany (*clavicula*) z pohřebiště u Sv. Jakuba v Brně, převážně z období středověku, naměřené jedním výzkumníkem ve dvou opakovaných měřeních (viz sekce ??). Hodnoty naměřené při prvním opakování jsou uloženy v proměnné `simd.1`, hodnoty naměřené při druhém opakování jsou uloženy v proměnné `simd.2`. Můžeme zjistit, že aritmetický průměr hodnot získaných v rámci prvního měření je větší než aritmetický průměr hodnot získaných v rámci druhého měření. Na hladině významnosti $\alpha = 0.05$ zjistíte, zda je střední hodnota prvního měření větší než střední hodnota druhého měření vertikálního průměru délky těla klíční kosti na levé straně provedené tímto výzkumníkem.

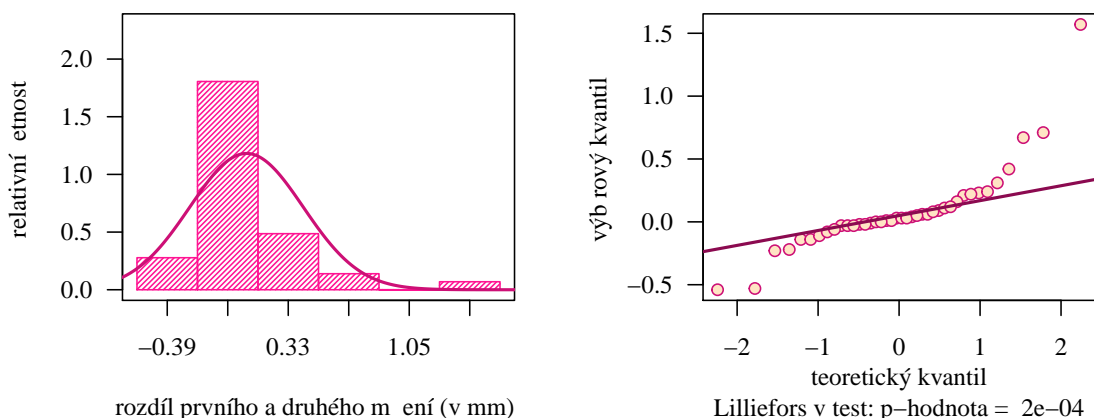
Řešení příkladu 9.11

Načteme datový soubor. pomocí operátoru `[]` z něj vybereme pouze údaje naměřené sledovaným výzkumníkem (sloupce `simd.1` a `simd.2`) na levé straně `side == 'L'`. Údaje vložíme do proměnné `data.12L`. Z datové tabulky následně odstraníme chybějící údaje a zjistíme rozsah náhodného výběru.

```
321 data <- read.delim('00-Data//02-paired-means-clavicle.txt')
322 #head(data)
323 data.12L <- data[data$side == 'L', c('simd.1', 'simd.2')]
324 data.12L <- na.omit(data.12L)
325 dim(data.12L) # 40 x 2
```

Datový soubor obsahuje 40 opakovaných měření vertikálního průměru středu délky těla klíční kosti z levé strany. Jelikož chceme porovnat opakovaná měření provedená na jednom subjektu, použijeme k tomuto porovnání párový test. Prvním krokem k provedení tohoto testu je vytvoření rozdílů hodnot získaných v prvním a druhém měření. V druhém kroku je potřeba ověřit předpoklad normálního rozdělení těchto rozdílů. Testováním normality náhodného výběru rozdílů jsme se zabývali v rámci příkladu ?? v sekci ??, přičemž jsme zjistili, že tento náhodný výběr nepochází z normálního rozdělení (p -hodnota = 0.0002), a to ani po odstranění nejvíce odlehlého rozdílu (p -hodnota = 0.0302). Pro připomenutí viz obrázek 21 zobrazující histogram a QQ-diagram pro původní sadu 40 rozdílů.

```
326 simd.1L <- data.12L$simd.1
327 simd.2L <- data.12L$simd.2
328 simd.12L <- simd.1L - simd.2L
```



Obrázek 21: Histogram a diagram rozdílů prvního a druhého měření vertikálního průměru ve středu délky těla klíční kosti na levé straně získaných jedním výzkumníkem

Protože předpoklad normality rozdílů prvního a druhého měření není splněn, a k jeho splnění nedošlo ani po

odstranění nejdlehlšího pozorování, vrátíme se zpátky k vektoru obsahujícímu všech 40 rozdílů. Otázku ze zadání ověříme pomocí neparametrické alternativy párového testu, a sice pomocí asymptotického znaménkového párového testu. Naším úkolem je zjistit, zda je střední hodnota prvního měření větší než střední hodnota druhého měření. Toto tvrzení bude zněním alternativní hypotézy, neboť v zadání není o znění nulové hypotézy žádná zmínka. Analogické znění alternativní hypotézy je, že střední hodnota rozdílů, získaných odečtením hodnot naměřených v druhém měření od hodnot naměřených v prvním měření, bude větší než nula. Nulovou hypotézu naformulujeme jako doplněk k této hypotéze. Namísto středních hodnot se opět zaměříme na jejich neparametrické alternativy a sice na mediány, které budou figurovat v konečném znění obou hypotéz.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : Medián rozdílů prvního a druhého měření vertikálního průměru těla klíční kosti na levé straně měřených jedním výzkumníkem je menší než nula.

H_1 : Medián rozdílů prvního a druhého měření vertikálního průměru těla klíční kosti na levé straně měřených jedním výzkumníkem je větší než nula.

- **matematická formulace** nulové a alternativní hypotézy

$H_0 : \tilde{z} \leq \tilde{z}_0$, kde $\tilde{z}_0 = 0$

$H_1 : \tilde{z} > \tilde{z}_0$, kde $\tilde{z}_0 = 0$ (pravostranná alternativa)

2. Volba hladiny významnosti

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

3. Testování kritickým oborem

- **Testovací statistika**

V úvodu příkladu jsme vytvořili vektor `simd.12L` (Z_i) jako vektor rozdílů mezi naměřenými hodnotami v rámci prvního a druhého měření na levé straně. Nyní je potřeba stanovit rozdíl $Z_i - \tilde{z}_0$ a určit počet kladných rozdílů $Z_i - \tilde{z}_0$ a počet nenulových rozdílů m .

```
329 z0 <- 0
330 I <- (simd.12L > z0)
331 SE <- sum(I) # 23
332 n <- length(simd.12L) # 40
333 m <- sum(simd.12L - z0 != 0) # 38
```

Z celkového počtu 40 rozdílů jsou dva rozdílů nenulové, tj. $m = 38$ a 23 rozdílů je kladných, tj. $S_E = 23$.

$$S_E = \sum_{i=1}^n I(Z_i - \tilde{z}_0 > 0) = \sum_{i=1}^{40} I(Z_i - 0 > 0) = 23.$$

Dále vypočítáme testovací statistiku asymptotické varianty znaménkového testu S_A podle vzorce 9.2.

$$S_A = \frac{S_E - \frac{m}{2}}{\sqrt{\frac{m}{4}}} = \frac{23 - \frac{38}{2}}{\sqrt{\frac{38}{4}}} = \frac{0}{23 - 19} \sqrt{9.5} = \frac{4}{3.082207} = 1.297771 \doteq 1.2978$$

```
334 SA <- (SE - m / 2) / sqrt(m / 4) # 1.297771
```

- Kritický obor

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

```

335 alpha <- 0.05
336 qnorm(1 - alpha) # 1.644854

```

- Závěr testování

Protože realizace testovací statistiky $s_A = 1.297771$ nenáleží do kritického oboru, tj. $s_A \notin W$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

4. Testování intervalem spolehlivosti

- Interval spolehlivosti

Pomocí příkazu `sort()` nejprve vzestupně seřadíme naměřené hodnoty Z_i , $i = 1, \dots, 40$. Hranice 95% intervalu spolehlivosti budou potom $(C_{1-\alpha})$ -tá hodnota v seřazeném vektoru naměřených hodnot, kde

$$\begin{aligned}
 C_{1-\alpha} &= \frac{n}{2} - u_{1-\alpha} \sqrt{\frac{n}{4}} \\
 &= \frac{40}{2} - u_{1-0.05} \sqrt{\frac{40}{4}} \\
 &= 20 - u_{0.95} \sqrt{10} \\
 &= 20 - 1.644854 \times 3.162278 \\
 &= 14.79851 \doteq 15
 \end{aligned}$$

a plus nekonečno.

Interval spolehlivosti má potom následující tvar

$$\begin{aligned}
 (d, h) &= \left(X^{(C_{1-\alpha})}; \infty \right) \\
 &= \left(X^{(15)}; \infty \right) \\
 &= \left(X^{(15)}; \infty \right) \\
 &= (-0.01; \infty)
 \end{aligned}$$

```

337 simd.12Ls <- sort(simd.12L)
338 C1 <- round(n / 2 - qnorm(1 - alpha) * sqrt(n / 4)) # 14
339 simd.12Ls[C1] # -0.01

```

- Závěr testování

Protože $\tilde{z}_0 = 0$ náleží do 95% empirického oboustranného intervalu spolehlivosti, tj. $\tilde{z}_0 = 0 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

5. Testování p -hodnotou

- p -hodnota

$$p\text{-hodnota} = \Pr(S_A \geq s_A) = \Pr(S_A \geq 1.297771) = 1 - \Pr(S_A \leq 1.297771) = 0.09718296 \doteq 0.09718$$

```
340 p.hodnota <- 1 - pnorm(SA) # 0.09718296
```

- **Závěr testování**

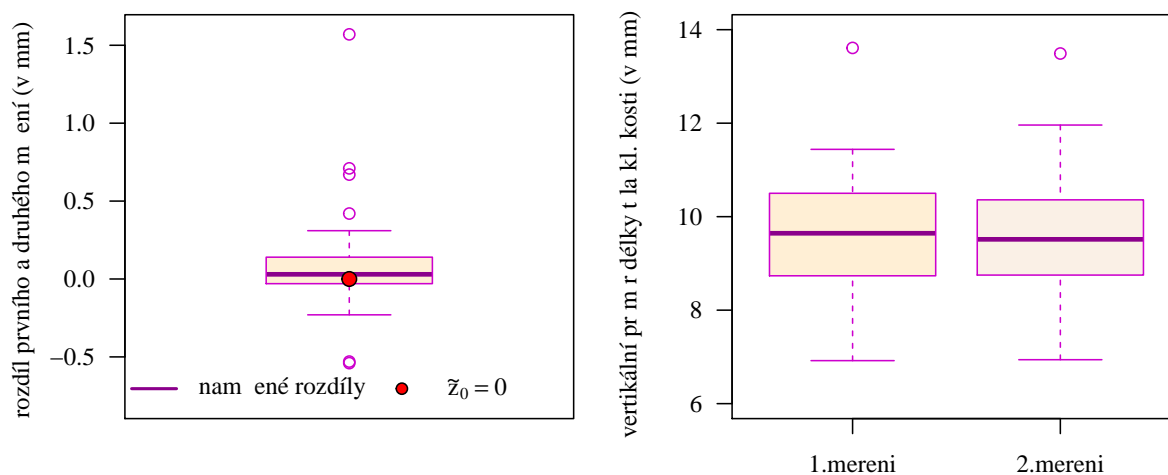
Protože p -hodnota = 0.09718 je větší než $\alpha = 0.05$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

6. Interpretace výsledků

Za základě všech tří typů testování nezamítáme nulovou hypotézu na hladině významnosti $\alpha = 0.05$. Hodnoty získané v prvním měření nejsou statisticky významně větší než hodnoty získané v druhém měření.

7. Grafická vizualizace výsledku testování

Porovnání naměřených hodnot získaných v rámci prvního a druhého měření zobrazíme pomocí krabicového diagramu (viz obrázek 22).



Obrázek 22: Krabicový diagram rozdílů prvního a druhého měření vertikálního průměru ve středu délky těla klíční kosti na levé straně získaných jedním výzkumníkem



Příklad 9.12. Znaménkový párový asymptotický test

Máme datový soubor 03-paired-means-clavicle2.txt obsahující údaje o délkách klíční kosti (*clavicula*) z pravé strany (`length.R`) a levé strany (`length.L`) z anglického souboru dokumentovaných skeletů (Parsons, 1916, viz soubor D-03-paired-means-clavicle2). Na hladině významnosti $\alpha = 0.05$ testujte hypotézu, že střední hodnota délky klíční kosti u žen z pravé strany je větší než střední hodnota délky klíční kosti u žen z levé strany.

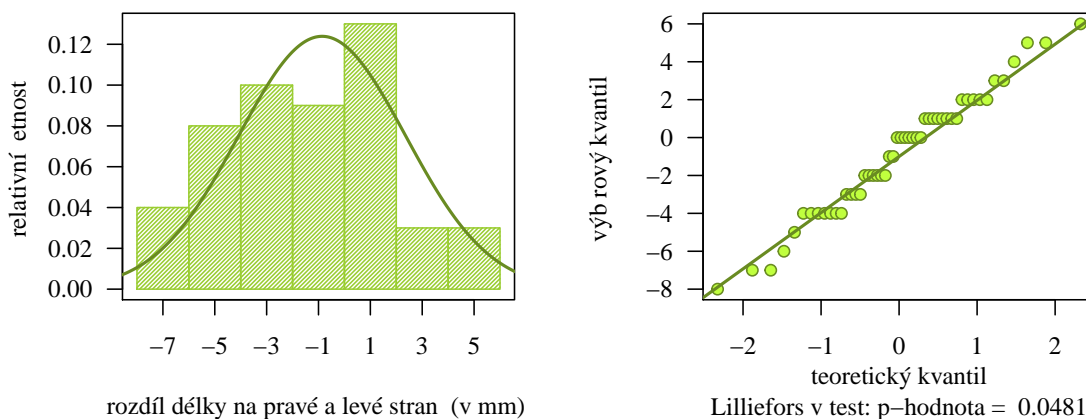
Řešení příkladu 9.12

Nejprve načteme datový soubor a pomocí operátoru `[]` z něj vybereme pouze údaje o délce klíční kosti z pravé strany (sloupce `length.R` resp. z levé strany (`length.L`) u žen `sex == 'f'`. Údaje vložíme do proměnné `length.RLF`. Z datové tabulky následně odstraníme chybějící hodnoty a zjistíme rozsah náhodného výběru.

```
341 data <- read.delim('00-Data//03-paired-means-clavicle2.txt')
342 #head(data)
343 data.RLF <- data[data$sex == 'f', c('length.R', 'length.L')]
344 data.RLF <- na.omit(data.RLF)
345 dim(data.RLF) # 50
```

Datový soubor obsahuje údaje o délce klíční kosti z pravé a levé strany u 50 mužů. Úkolem ze zadání je porovnat naměřené hodnoty na pravé a levé straně. Jde tedy o měření stejného znaku (délka klíční kosti) sledovaného na stejných subjektech (ženy), proto použijeme na tuto situaci párový test. Prvním krokem k tohoto testu je vytvoření rozdílů hodnot naměřených na pravé a levé straně. V druhém kroku je potřeba ověřit předpoklad normality těchto rozdílů.

Předpoklad normality ověříme Lillieforsovým testem ($\alpha = 0.05$) v kombinaci s QQ-diagramem a histogramem (viz obrázek 23). Datový soubor rozdělíme do sedmi ekvidistantních intervalů s šířkou 2 mm prostřednictvím stanovených hranic $-8, -6, \dots, 6$.



Obrázek 23: Histogram a QQ-diagram rozdílů délky těla klíčních kostí na levé a pravé straně žen

Protože p -hodnota = 0.0481 je menší než 0.05, nulovou hypotézu o normalitě dat, zamítáme na hladině významnosti $\alpha = 0.05$. Z histogramu vidíme, že rozdíly pravé a levé strany vykazují plošší trend než očekávaná křivka hustoty normálního rozdělení. Taktéž zde pozorujeme mírné vyšikmení hodnot směrem doleva s odlehklými pozorováními na pravé straně. Lillieforsův test tyto nedostatky vyhodnotil vzhledem k poměrně vysokému rozsahu náhodného výběru jako fatální pro normální charakter náhodného výběru. Náhodný výběr rozdílů délek klíčních kostí z levé a pravé strany tedy nepochází z normálního rozdělení.

Předpoklad normality pro použití parametrického testu není splněn, proto není možné tvrzení ze zadání ověřit pomocí parametrického párového testu. Otázku ze zadání tedy ověříme pomocí neparametrického asymptotického

znaménkového párového testu. Naším úkolem je tetsovat hypotézu, že střední hodnota délky klíční kosti u žen z pravé strany je větší než střední hodnota délky klíční kosti u žen z levé strany. Toto tvrzení je zněním nulové hypotézy, protože pojmem *hypotéza* zmíněným v zadání je vždy myšlena nulová hypotéza. Dále je potřeba si uvědomit, že možnost testovat hypotézu o středních hodnotách jsme ztratili při rozhodnutí použít k testování neparametrický test. Pojem *střední hodnoty* nahradíme ve slovní formulaci hypotéz pojmem *medián*.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : Medián rozdílů délky klíční kosti z pravé a z levé strany u žen je větší nebo roven nule.

H_1 : Medián rozdílů délky klíční kosti z pravé a z levé strany u žen je menší než nula.

- **matematická formulace** nulové a alternativní hypotézy

H_0 : $\tilde{z} \geq \tilde{z}_0$, kde $\tilde{z}_0 = 0$

H_1 : $\tilde{z} < \tilde{z}_0$, kde $\tilde{z}_0 = 0$ (levostranná alternativa)

2. Volba hladiny významnosti

- Hladinu významnosti podle zadání zvolíme jako $\alpha = 0.05$.

3. Testování kritickým oborem

- **Testovací statistika**

V úvodu příkladu jsme vytvořili vektor `length.RLF` (Z_i) jako vektor rozdílů mezi naměřenými hodnotami na pravé a levé straně. Nyní je potřeba stanovit rozdíl $Z_i - \tilde{z}_0$ a určit počet kladných rozdílů $Z_i - \tilde{z}_0$ a počet nenulových rozdílů m .

```
346 z0 <- 0
347 I <- (length.RLF > z0)
348 SE <- sum(I) # 19
349 n <- length(length.RLF) # 50
350 m <- sum(length.RLF - z0 != 0) # 43
```

Z celkového počtu 50 rozdílů je sedm rozdílů rovných nule, tj. počet nenulových rozdílů $m = 43$ a 19 rozdílů je kladných, tj. $S_E = 19$.

$$S_E = \sum_{i=1}^n I(Z_i - \tilde{z}_0 > 0) = \sum_{i=1}^{50} I(Z_i - 0 > 0) = 19.$$

Nyní vypočítáme testovací statistiku asymptotické varianty znaménkového testu podle vzorce 9.2.

$$S_A = \frac{S_E - \frac{m}{2}}{\sqrt{\frac{m}{4}}} = \frac{19 - \frac{43}{2}}{\sqrt{\frac{43}{4}}} = \frac{0}{19 - 21.5} \sqrt{10.75} = \frac{-2.5}{3.278719} = -0.7624929 \doteq -0.7625$$

```
351 SA <- (SE - m / 2) / sqrt(m / 4) # -0.7624929
```

- **Kritický obor**

$$\begin{aligned} W &= (-\infty; u_\alpha) \\ &= (-\infty; u_{0.05}) \\ &= (-\infty; -1.644854) \end{aligned}$$

```
352 alpha <- 0.05
353 qnorm(alpha) # -1.644854
```

- **Závěr testování**

Protože realizace testovací statistiky $s_A = -0.16013$ nenáleží do kritického oboru, tj. $s_A \notin W$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

4. Testování intervalem spolehlivosti

- **Interval spolehlivosti**

Pomocí příkazu `sort()` nejprve vzestupně seřadíme naměřené hodnoty Z_i , $i = 1, \dots, 50$. Hranice 95% pravostranného intervalu spolehlivosti budou potom mínus nekonečno a $(n + 1 - C_{1-\alpha})$ -tá hodnota v seřazeném vektoru naměřených hodnot, kde

$$\begin{aligned} C_{1-\alpha} &= \frac{n}{2} - u_{1-\alpha} \sqrt{\frac{n}{4}} \\ &= \frac{50}{2} - u_{1-0.05} \sqrt{\frac{50}{4}} \\ &= 25 - u_{0.95} \sqrt{12.5} \\ &= 25 - 1.644854 \times 3.535534 \\ &= 19.18456 \doteq 19 \end{aligned}$$

Interval spolehlivosti má potom následující tvar

$$\begin{aligned} (d, h) &= \left(-\infty; X^{(n+1-C_{1-\alpha/2})} \right) \\ &= \left(-\infty; X^{(50+1-19)} \right) \\ &= \left(-\infty; X^{(32)} \right) \\ &= (-\infty; 1) \end{aligned}$$

```
354 length.RLFs <- sort(length.RLF)
355 C1 <- round(n / 2 - qnorm(1 - alpha) * sqrt(n / 4)) # 19
356 C2 <- round(n + 1 - C1) # 32
357 length.RLFs[C2] # 1
```

- **Závěr testování**

Protože $\tilde{z}_0 = 0$ náleží do 95% empirického pravostranného intervalu spolehlivosti, tj. $\tilde{z}_0 = 0 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

5. Testování p -hodnotou

- **p -hodnota**

$$p\text{-hodnota} = \Pr(S_A \leq s_A) = \Pr(S_A \leq -0.7624929) = 0.222883 \doteq 0.2229$$

```
358 p.hodnota <- pnorm(SA) # 0.2229
```


- **Závěr testování**

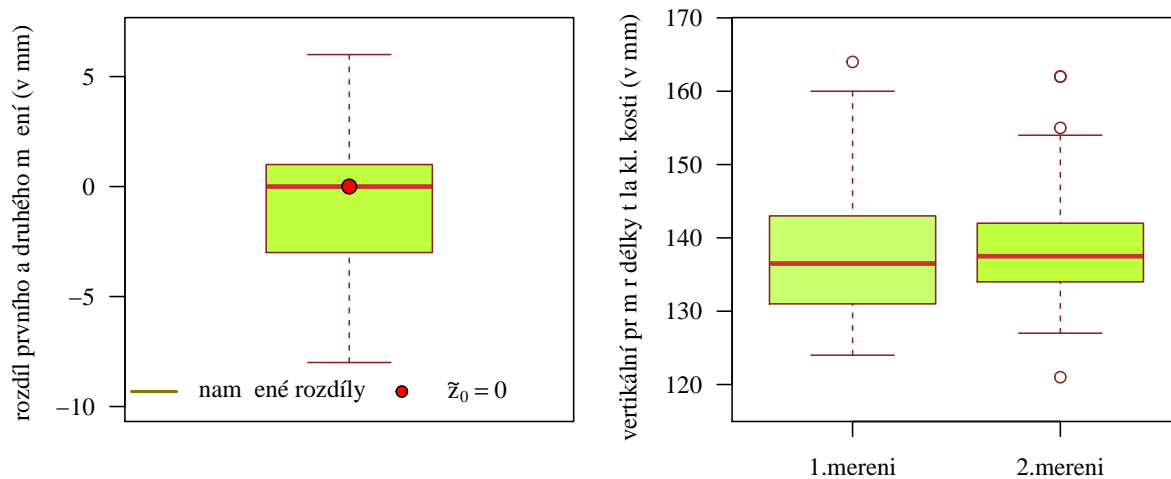
Protože p -hodnota = 0.2229 je větší než $\alpha = 0.05$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

6. **Interpretace výsledků**

Za základě všech tří typů testování nezamítáme nulovou hypotézu na hladině významnosti $\alpha = 0.05$. Délka klíční kosti z pravé strany není u žen statisticky významně menší než délka klíční kosti z levé strany.

7. **Grafická vizualizace výsledku testování**

Porovnání naměřených hodnot na pravé a levé straně vizualizujeme pomocí krabicového diagramu (viz obrázek 24).



Obrázek 24: Krabicový diagram rozdílů délky těla klíčních kostí na levé a pravé straně žen



9.4 Wilcoxonův jednovýběrový exaktní test

Nechť X_1, \dots, X_{1n} , $n \geq 2$ je náhodný výběr ze spojitého rozdělení s hustotou $f(x)$, která je symetrická okolo mediánu \tilde{x} a nechť \tilde{x}_0 je konstanta. Na hladině významnosti α testujeme jednu z následujících tří hypotéz oproti příslušné alternativní hypotéze.


$$\begin{array}{lll} H_{01} : \tilde{x} = \tilde{x}_0 & \text{oproti} & H_{11} : \tilde{x} \neq \tilde{x}_0 \quad (\text{oboustranná alt.}) \\ H_{02} : \tilde{x} \leq \tilde{x}_0 & \text{oproti} & H_{12} : \tilde{x} > \tilde{x}_0 \quad (\text{pravostranná alt.}) \\ H_{03} : \tilde{x} \geq \tilde{x}_0 & \text{oproti} & H_{13} : \tilde{x} < \tilde{x}_0 \quad (\text{levostranná alt.}) \end{array}$$

Test nazýváme Wilcoxonův jednovýběrový exaktní test o mediánu \tilde{x} . Testovací statistika má tvar

$$S_E = \sum_{i=1}^m R_i I(X_i - \tilde{x}_0 > 0) \quad (9.4)$$

kde m je počet nenulových rozdílů $X_i - \tilde{x}_0$, R_i je pořadí absolutní hodnoty rozdílu i -té náhodné veličiny X_i a konstanty \tilde{x}_0 , tj. R_i je pořadí seřazených rozdílů $|X_i - \tilde{x}_0|$, $i = 1, \dots, m$, $I(X_i - \tilde{x}_0 > 0)$ je indikační funkce, která nabývá hodnoty 1, pokud $X_i - \tilde{x}_0 > 0$, a hodnoty 0 jinak. Statistika S_E je potom součet pořadí absolutních hodnot rozdílů $|X_i - \tilde{x}_0|$, pro něž je rozdíl $X_i - \tilde{x}_0$ kladný. Kritický obor podle zvolené alternativní hypotézy má tvar

$$\begin{array}{ll} H_{11} : \tilde{x} \neq \tilde{x}_0 & W = (-\infty; s_m(\alpha/2) - 1) \cup \langle s_m(1 - \alpha/2); \infty \rangle \\ H_{12} : \tilde{x} > \tilde{x}_0 & W = \langle s_m(1 - \alpha); \infty \rangle \\ H_{13} : \tilde{x} < \tilde{x}_0 & W = (-\infty; s_m(\alpha) - 1) \end{array}$$

kde $s_m(\alpha/2)$, $s_m(1 - \alpha/2)$, $s_m(\alpha)$ a $s_m(1 - \alpha)$ jsou tabelované kvantily pro jednovýběrový Wilcoxonův test, jejichž hodnoty získáme pomocí softwaru  a implementované funkce `qsignrank()`.


Interval spolehlivosti má podle zvolené alternativní hypotézy jeden z následujících tvarů

$$\begin{array}{ll} H_{11} : \tilde{x} \neq \tilde{x}_0 & (d, h) = (V^{(s_m(\alpha/2))}; V^{(s_m(1-\alpha/2)+1)}) \\ H_{12} : \tilde{x} > \tilde{x}_0 & (d, \infty) = (V^{(s_m(\alpha))}; \infty) \\ H_{13} : \tilde{x} < \tilde{x}_0 & (-\infty, h) = (-\infty; V^{(s_m(1-\alpha)+1)}) \end{array}$$

kde $V^{(1)} < V^{(2)} < \dots < V^{(\frac{m(m+1)}{2})}$ značí posloupnost vzestupně seřazených $\frac{m(m+1)}{2}$ Walshových průměrů $\frac{(X_i + X_j)}{2}$, $i = 1, \dots, m$, $j = 1, \dots, m$, $j \leq i$ a $V^{(k)}$ značí k -tý seřazený Walshův průměr. Posloupnost Walshových průměrů získáme příkazem `owa` z knihovny `NSM3`.

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

$$\begin{array}{ll} H_{11} : \tilde{x} \neq \tilde{x}_0 & p\text{-hodnota} = 2 \min\{\Pr(S_E \leq s_E), \Pr(S_E \geq s_E)\} \\ H_{12} : \tilde{x} > \tilde{x}_0 & p\text{-hodnota} = \Pr(S_E \geq s_E) \\ H_{13} : \tilde{x} < \tilde{x}_0 & p\text{-hodnota} = \Pr(S_E \leq s_E) \end{array}$$

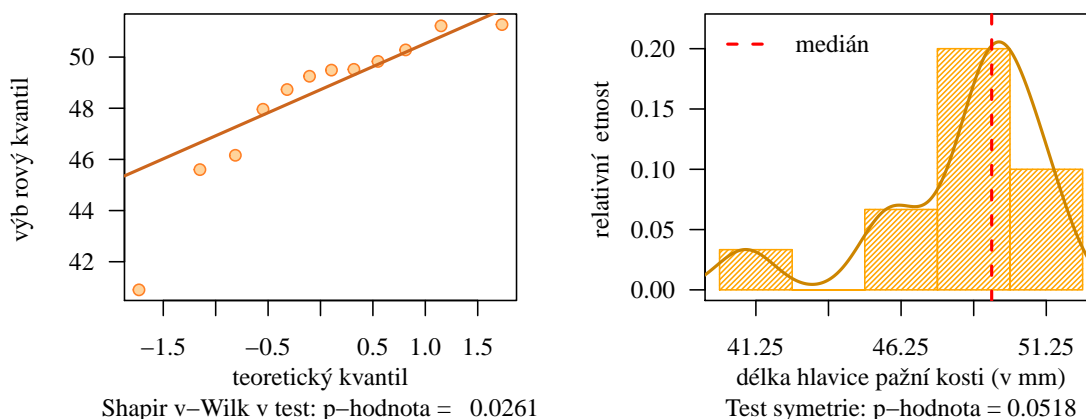
kde S_E je náhodná veličina, s_E je realizace testovací statistiky S_E (viz vzorec 9.4), tedy konkrétní číslo, a $\Pr(S_E \leq s_E)$ je distribuční funkce tabelovaného rozdělení pro jednovýběrový Wilcoxonův test, jejíž hodnotu získáme pomocí  a implementované funkce `psignrank()`.

Příklad 9.13. Wilcoxonův jednovýběrový exaktní test

Mějme datový soubor 21-goldman-measures.csv a proměnnou humer.DR popisující největší délku hlavice pažní kosti z pravé strany v mm u skeletů z období raného středověku z oblasti *Staubing* (viz sekce ??). Dále máme k dispozici údaje ze studie (Mall et al.) z roku 1999, v rámci které byla měřena největší délka hlavice pažní kosti u mužů současné německé populace ($m_m = 50.0$ mm, $n_m = 64$). Na hladině významnosti $\alpha = 0.05$ zjistěte, zda existuje rozdíl mezi největší délkou hlavice pažní kosti z pravé strany u mužů z raně středověké německé populace a u mužů současné německé populace.

Řešení příkladu 9.13

Zadání příkladu je shodné se zadáním příkladu 9.1. Datový soubor obsahující údaje o největší délce hlavice pažní kosti z pravé strany u 12 mužů nepochází z normálního rozdělení (viz QQ-diagram, obrázek 25 vlevo), proto na prozkoumání otázky ze zadání použijeme neparametrický jednovýběrový test. Tentokrát bychom však chtěli nulovou hypotézu otestovat pomocí Wilcoxonova jednovýběrového testu. K použití tohoto testu však musí být splněn předpoklad symetrického rozdělení naměřených hodnot okolo mediánu. Ten ověříme pomocí testu symetrie (příkaz `symmetry.test()` z knihovny `lawstat`) uvedeného v sekci ?? (hladinu významnosti zvolíme $\alpha = 0.05$). Graficky vizualizujeme míru symetrie naměřených hodnot okolo mediánu pomocí histogramu superponovaného jádrovým odhadem křivky hustoty (viz kapitola ??) se zvýrazněnou hodnotou mediánu pomocí vertikální přímky (viz obrázek 25 vpravo).



Obrázek 25: QQ-diagram (vlevo) a histogram naměřených hodnot největší délky hlavice pažní kosti na pravé straně u skeletů mužů ze starověké populace z oblasti Staubing superponovaný křivkou jadrového odhadu hustoty (vpravo)

Protože p -hodnota = 0.0518 je větší než 0.05, nezamítáme hypotézu o symetrii okolo mediánu na hladině významnosti $\alpha = 0.05$. Z histogramu na obrázku 25 vidíme, že data nejsou ukázkově symetrická okolo mediánu, jsou vyšikmená doprava s prodlouženým levým koncem. S ohledem na nízký rozsah náhodného výběru jsou však tyto prohřešky vzhledem k symetrii ještě přijatelné. Naměřené hodnoty největší délky hlavice pažní kosti z pravé strany tedy považujeme za symetricky rozdělené okolo mediánu.

Protože hypotéza o symetrii data okolo mediánu nebyla zamítnuta, můžeme k otestování otázky ze zadání použít Wilcoxonův jednovýběrový parametrický test. Analogicky jako v příkladu 9.1 provedeme testování v posloupnosti sedmi kroků.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : Medián největší délky hlavice pažní kosti raně středověké německé populace je shodný s mediánem největší délky hlavice pažní kosti na pravé straně mužů současné německé populace.

H_1 : Medián největší délky hlavice pažní kosti raně středověké německé populace není shodný s mediánem největší délky hlavice pažní kosti na pravé straně mužů současné německé populace.

- **matematická formulace** nulové a alternativní hypotézy

$$H_0 : \tilde{x} = \tilde{x}_0, \text{ kde } \tilde{x}_0 = 50.00$$

$$H_1 : \tilde{x} \neq \tilde{x}_0, \text{ kde } \tilde{x}_0 = 50.00 \text{ (oboustranná alternativa)}$$

2. Volba hladiny významnosti

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

3. Testování kritickým oborem

- **Testovací statistika**

Nejprve vypočítáme vektor rozdílů naměřených hodnot X_i a konstanty \tilde{x}_0 , tj. $X_i - \tilde{x}_0 = X_i - 50$. Následně se zaměříme na stanovení počtu nenulových rozdílů, kde zjistíme, že žádný z rozdílů $X_i - 50$ není nulový, tedy počet nenulových rozdílů $m = n = 12$. Pro dodržení obecného postupu vytvoříme nyní vektor *humer.DRM0* obsahující ty naměřené délky pažních kostí z pravé strany, pro než rozdíl $X_i - \tilde{x}_0$ nebyl rovný nule. V tomto případě bude vektor *humer.DRM0* shodný s vektorem *humer.DRM*, protože žádný z rozdílů $X_i - \tilde{x}_0$ nule rovný nebyl. V dalších krocích testování budeme nicméně používat vektor *humer.DRM0*. Následně vypočítáme absolutní hodnoty nenulových rozdílů, tj. $|X_i - \tilde{x}_0|$, $i = 1, \dots, m$ a stanovíme pořadí těchto rozdílů v absolutních hodnotách (viz tabulka 12).

Tabulka 12: Naměřené hodnoty X_i , rozdíly $X_i - \tilde{x}_0$, znaménka těchto rozdílů, absolutní hodnoty nenulových rozdílů $|X_i - \tilde{x}_0|$ a pořadí nenulových rozdílů v absolutních hodnotách R_i

měření	1	2	3	4	5	6	7	8	9	10	11	12
X_i	49.83	47.96	49.49	48.73	49.25	40.9	51.22	51.27	46.16	49.52	50.28	45.6
$X_i - \tilde{x}_0$	-0.17	-2.04	-0.51	-1.27	-0.75	-9.1	1.22	1.27	-3.84	-0.48	0.28	-4.4
I	-	-	-	-	-	-	+	+	-	-	+	-
$ X_i - \tilde{x}_0 $	0.17	2.04	0.51	1.27	0.75	9.1	1.22	1.27	3.84	0.48	0.28	4.4
R_i	1.00	9.00	4.00	7.50	5.00	12.0	6.00	7.50	10.00	3.00	2.00	11.0

Z tabulky 12 vidíme, že celkem tři rozdíly $X_i - \tilde{x}_0$, $i = 1, \dots, m$ jsou kladné, přičemž pořadí těchto rozdílů v absolutních hodnotách jsou 2, 6 a 7.5. Hodnota testovací statistiky S_E , která je definovaná jako součet pořadí absolutních hodnot kladných rozdílů, bude tedy rovná 15.5.

$$S_E = \sum_{i=1}^{12} R_i I(X_i - \tilde{x}_0 > 0) = 2 + 6 + 7.5 = 15.5$$

- **Kritický obor**

$$\begin{aligned} W &= (-\infty; s_m(\alpha/2) - 1) \cup \langle s_m(1 - \alpha/2); \infty \rangle \\ &= (-\infty; s_{12}(0.05/2) - 1) \cup \langle s_{12}(1 - 0.05/2); \infty \rangle \\ &= (-\infty; s_{12}(0.025) - 1) \cup \langle s_{12}(0.975); \infty \rangle \\ &= (-\infty; 14 - 1) \cup \langle 64; \infty \rangle \\ &= (-\infty; 13) \cup \langle 64; \infty \rangle \end{aligned}$$

- **Závěr testování**

Protože realizace testovací statistiky $s_E = 15.5$ nenáleží do kritického oboru, tj. $s_E \notin W$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

4. Testování intervalem spolehlivosti

- **Interval spolehlivosti**

Abychom mohli stanovit hranice 95% intervalu spolehlivosti, musíme nejprve vypočítat $\frac{m(m+1)}{2}$ Walshových průměrů $\frac{X_i + X_j}{2}$, $i = 1, \dots, m$ a ty následně vzestupně seřadit. Walshovy průměry získáme

```

359 x0 <- 50
360 n <- length(humer.DRM)
361 m <- sum(humer.DRM - x0 != 0)
362 humer.DRMO <- humer.DRM
363 humer.DRMO[humer.DRMO - x0 == 0] <- NA
364
365 I <- (humer.DRMO - x0 > 0)
366 Ri <- rank(abs(humer.DRMO - x0), na.last = "keep")
367 tab <- rbind("Xi" = humer.DRM, "Xi-x0" = humer.DRM - x0,
368             "I" = I, "|Xi-x0|" = abs(humer.DRMO - x0),
369             "Ri" = Ri)
370 tab <- data.frame(tab)
371 names(tab) <- 1 : 12
372 #           1      2      3      4      5      6      7      8      9     10     11     12
373 # Xi      49.83 47.96 49.49 48.73 49.25 40.9 51.22 51.27 46.16 49.52 50.28 45.6
374 # Xi-x0   -0.17 -2.04 -0.51 -1.27 -0.75 -9.1  1.22  1.27 -3.84 -0.48  0.28 -4.4
375 # I        0.00  0.00  0.00  0.00  0.00  0.0  1.00  1.00  0.00  0.00  1.00  0.0
376 # |Xi-x0|  0.17  2.04  0.51  1.27  0.75  9.1  1.22  1.27  3.84  0.48  0.28  4.4
377 # Ri       1.00  9.00  4.00  7.50  5.00 12.0  6.00  7.50 10.00  3.00  2.00 11.0
378 SE <- sum(Ri * I) # 15.5

```

```

379 alpha <- 0.05
380 qsignrank(alpha / 2, m) - 1 # 13
381 qsignrank(1 - alpha / 2, m) # 64

```

pomocí funkce `owa` z knihovny `NSM3`. Prvním argumentem této funkce bude vektor nul o délce $m = 12$, druhým argumentem bude vektor `humer.DRMO` o délce $m = 12$. Získané Walschovy průměry seřadíme příkazem `sort()`.

```

382 walsh <- NSM3::owa(rep(0, 12), humer.DRMO)$owa
383 V <- sort(walsh)
384 # [ 1] 40.900 43.250 43.530 44.430 44.815 45.075 45.195 45.210 45.365 45.590 45.600
385 # [12] 45.880 46.060 46.085 46.160 46.780 47.060 47.165 47.425 47.445 47.545 47.560
386 # [23] 47.705 47.715 47.825 47.840 47.940 47.960 47.995 48.220 48.345 48.410 48.435
387 # [34] 48.605 48.690 48.715 48.725 48.730 48.740 48.895 48.990 49.110 49.120 49.125
388 # [45] 49.250 49.280 49.370 49.385 49.490 49.505 49.505 49.520 49.540 49.590 49.615
389 # [56] 49.660 49.675 49.765 49.830 49.885 49.900 49.975 50.000 50.055 50.235 50.260
390 # [67] 50.280 50.355 50.370 50.380 50.395 50.525 50.550 50.750 50.775 51.220 51.245
391 # [78] 51.270

```

Hranice intervalu spolehlivosti potom tvoří ty hodnoty, které se v seřazeném vektoru Walshových průměrů V_s nachází na $(s_m(\alpha/2))$ -té pozici a na $(s_m(1 - \alpha/2) + 1)$ -té pozici. Hodnoty $s_m(\alpha/2)$ a $s_m(1 - \alpha/2) + 1$ nalezneme opět pomocí funkce `qsignrank()`.

$$\begin{aligned}
(d, h) &= \left(V^{(s_m(\alpha/2))}; V^{(s_m(1-\alpha/2)+1)} \right) \\
&= \left(V^{(s_{12}(0.05/2))}; V^{(s_{12}(1-0.05/2)+1)} \right) \\
&= \left(V^{(14)}; V^{(64)} \right) \\
&= (46.085; 50.235)
\end{aligned}$$

- Závěr testování

```

392 qsignrank(alpha / 2, m) # 14
393 qsignrank(1 - alpha / 2, m) + 1 # 64
394 V[14] # 46.085
395 V[65] # 50.235

```

Protože $\tilde{x}_0 = 50.00$ náleží do 95% empirického oboustranného intervalu spolehlivosti, tj. $\tilde{x}_0 = 50.00 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

5. Testování p -hodnotou

Průslušnou p -hodnotu vypočítáme pomocí vzorce $2 \min\{\Pr(S_E \leq s_E), \Pr(S_E \geq s_E)\}$. Zde si uvědomme, že realizace testovací statistiky $s_E = 15.5$. Zároveň S_E je diskrétní náhodná veličina. Z vlastností pravděpodobnostní funkce diskrétních náhodných veličin víme, že $\Pr(S_E \leq 15.5) = \Pr(S_E \leq 15)$ a $\Pr(S_E \geq 15.5) = \Pr(S_E \geq 16) = 1 - \Pr(S_E < 16) = \Pr(S_E \leq 15)$. Viz kapitola ??.

- p -hodnota

$$\begin{aligned}
p\text{-hodnota} &= 2 \min\{\Pr(S_E \leq s_E), \Pr(S_E \geq s_E)\} \\
&= 2 \min\{\Pr(S_E \leq 15), 1 - \Pr(S_E \leq 15)\} \\
&= 2 \min\{0.03198242, 0.9680176\} \\
&= 2 \times 0.03198242 = 0.06396484 \doteq 0.06396
\end{aligned}$$

```

396 2 * min (psignrank(15, m), 1 - psignrank(15, m)) # 0.06396484

```

- Závěr testování

Protože p -hodnota = 0.06396 je větší než $\alpha = 0.05$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

6. Interpretace výsledků

Za základě všech tří typů testování nezamítáme nulovou hypotézu na hladině významnosti $\alpha = 0.05$. Mezi největší délkou hlavice pažní kosti na pravé straně u mužů raně středověké a současné populace neexistuje statisticky významný rozdíl. Ke stejnému závěru jsme dospěli také v příkladu 9.1.

7. Grafická vizualizace výsledku testování

Grafické porovnání náhodného výběru s konstantou $\tilde{x}_0 = 50.00$ bychom provedli analogicky jako v příkladu 9.1 pomocí krabicového diagramu (viz obrázek 2).

Poznámka: Wilcoxonův jednovýběrový exaktní test můžeme provést pomocí funkce `wilcox.test()`. Vstupními parametry budou vektor reprezentující náhodný výběr (`humer.DRM`), hodnota parametru \tilde{x}_0 z nulové hypotézy zadaná argumentem `mu = 50.00`, požadavek na výpočet hranic intervalu spolehlivosti zadaný nastavením argumentu `conf.int = T`, hodnota hladiny významnosti α zadaná prostřednictvím koeficientu spolehlivosti $1 - \alpha$ nastavením hodnoty argumentu `conf.level = 0.95`, typ zvolené alternativní hypotézy (oboustranná) zadaný pomocí argumentu `alternative = 'two.sided'` a argument `correct = F`, který zakazuje aplikování spojitě korekce na výsledné intervaly spolehlivosti a p -hodnotu.

```

397 wilcox.test(humer.DRM, mu = 50.0, conf.int = T, conf.level = 0.95,
398             alternative = 'two.sided', correct = F) # interpolovany IS i p-hodnota

```

Součástí výstupu je hodnota mediánu náhodného výběru `median of x = 48.79584`, hodnota testovací statistiky `V = 15.5`, interpolované hranice 95% Waldova empirického oboustranného intervalu spolehlivosti 46.085 a 50.235 a interpolovaná p -hodnota `p-value = 0.07108` adaptovaná na situaci, kdy jsou některé rozdíly $|X_i - \tilde{x}_0|$ shodné (viz tabulka

Wilcoxon signed rank test	399
	400
data: humer.DRM	401
V = 15.5, p-value = 0.06515	402
alternative hypothesis: true location is not equal to 50	403
95 percent confidence interval:	404
46.15993 50.05504	405
sample estimates:	406
(pseudo)median	407
48.79584	408
	409

12, kde $|X_4 - \tilde{x}_0| = |X_7 - \tilde{x}_0| = 1.27$). Jediné, co musíme stanovit zvlášť, jsou dolní a horní hranice kritického oboru.

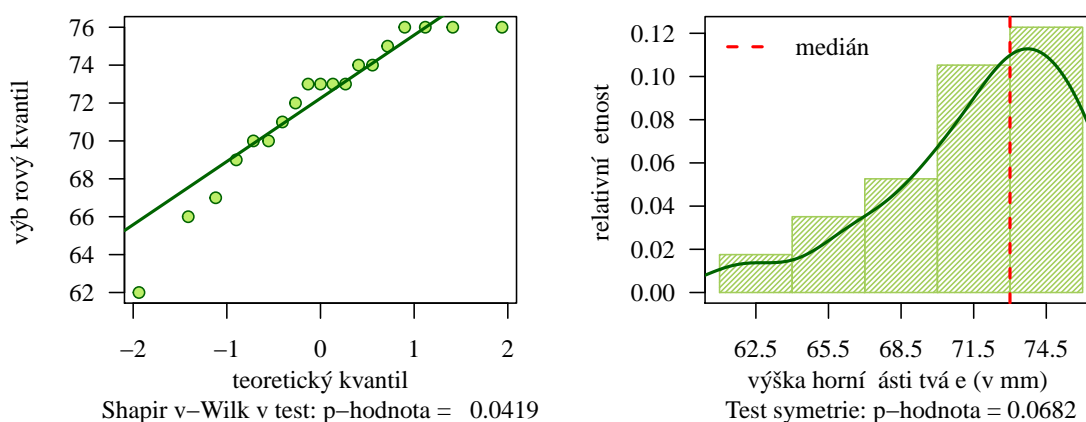
Poznámka: Výstup funkce `wilcox.test()` v tomto případě provází varovná hláška `Warning messages: In wilcox.test.default (humer.DRM, mu = 50, conf.int = T, conf.level = 0.95, : cannot compute exact p-value with ties`. Tato hláška nás upozorňuje právě na výskyt duplicitních rozdílů $|X_4 - \tilde{x}_0| = |X_7 - \tilde{x}_0| = 1.27$ a na použití modifikovaného postupu na výpočet p -hodnoty. ★

Příklad 9.14. Wilcoxonův jednovýběrový exaktní test

Mějme datový soubor 15-anova-means-skull.txt a proměnnou upface.H popisující výšku horní části tváře mužů německé populace (viz sekce ??). Dále máme k dispozici údaje o výšce horní části tváře mužů Černjachovské kultury na území dnešní Ukrajiny ($m_{ck} = 70.00$ mm, $n_{ck} = 99$). Na hladině významnosti $\alpha = 0.10$ testujte hypotézu, že výška horní části tváře německé mužské populace je menší nebo rovna výšce horní části tváře mužské populace Černjachovské kultury.

Řešení příkladu 9.14

Zadání příkladu je shodné se zadáním příkladu 9.2. Datový soubor obsahující údaje o výšce horní části tváře 19 mužů německé populace nepochází z normálního rozdělení (p -hodnota = 0.0419, viz obrázek 9.14), proto na prozkoumání otázky ze zadání použijeme neparametrický jednovýběrový test. Opět bychom chtěli nulovou hypotézu otestovat pomocí Wilcoxonova jednovýběrového testu, k čemuž je potřeba ověřit předpoklad symetrického rozdělení naměřených hodnot okolo mediánu. Ten ověříme pomocí testu symetrie ($\alpha = 0.05$). Graficky vizualizujeme míru symetrie naměřených hodnot okolo mediánu pomocí histogramu superponovaného jádrovým odhadem křivky hustoty (viz kapitola ??) se zvýrazněnou hodnotou mediánu pomocí vertikální přímky (viz obrázek 26).



Obrázek 26: QQ-diagram (vlevo) a histogram naměřených hodnot výšky horní části tváře u mužů německé populace superponovaný křivkou jádrového odhadu hustoty (vpravo)

Protože p -hodnota = 0.0682 je větší než 0.05, hypotézu o symetrickém rozdělení náhodného výběru výšek horní části tváře mužů okolo mediánu nezamítáme na hladině významnosti $\alpha = 0.05$. Z grafu 26 není symetrie okolo mediánu příliš zjevná, hodnoty jsou vyšikmené směrem doprava. U náhodného výběru většího rozsahu by to byl problém, vzhledem k nízkému rozsahu náhodného výběru je však test symetrie schovávavější a hypotéza o symetrii nebyla ještě zamítnuta. Náhodný výběr výšek horní části tváře u mužů považujeme tedy za symetricky rozdělený okolo mediánu.

Protože hypotéza o symetrii data okolo mediánu nebyla zamítnuta, můžeme k otestování otázky ze zadání použít Wilcoxonův jednovýběrový parametrický test. Analogicky jako v příkladu 9.2 testujeme (nulovou) hypotézu v polysloupnosti sedmi kroků.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : Medián výšky horní části tváře německé mužské populace je menší nebo roven mediánu výšky horní části tváře mužské populace z Černjachovské kultury.

H_1 : Medián výšky horní části tváře německé mužské populace je větší než medián výšky horní části tváře mužské populace z Černjachovské kultury.

- **matematická formulace** nulové a alternativní hypotézy

H_0 : $\tilde{x} \leq \tilde{x}_0$, kde $\tilde{x}_0 = 70.00$

H_1 : $\tilde{x} > \tilde{x}_0$, kde $\tilde{x}_0 = 70.00$ (pravostranná alternativa)

2. Volba hladiny významnosti

- Hladinu významnosti volíme v souladu se zadáním jako $\alpha = 0.10$.

3. Testování kritickým oborem

• Testovací statistika

Nejprve vypočítáme vektor rozdílů naměřených hodnot X_i a konstanty \tilde{x}_0 , tj. $X_i - \tilde{x}_0 = X_i - 70$. Následně se zaměříme na stanovení počtu nenulových rozdílů, kde zjistíme, že z celkového počtu $n = 19$ rozdílů jsou dva rozdílů nulové, a tedy počet nenulových rozdílů $m = 17$. Dále vytvoříme vektor `upface.HN0` obsahující ty naměřené výšky horní části tváře, pro něž rozdíl $X_i - \tilde{x}_0$ nebyl rovný nule. Délka vektoru `upface.HN0` bude rovná $m = 17$. V dalších krocích testování budeme používat právě vektor `upface.HN0`. Následně vypočítáme absolutní hodnoty nenulových rozdílů, tj. $|X_i - \tilde{x}_0|$, $i = 1, \dots, m$ a stanovíme pořadí těchto rozdílů v absolutních hodnotách (viz tabulka 13).

Tabulka 13: Naměřené hodnoty X_i , rozdíly $X_i - \tilde{x}_0$, znaménka těchto rozdílů, absolutní hodnoty rozdílů $|X_i - \tilde{x}_0|$ a pořadí rozdílů v absolutních hodnotách R_i

měření	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
X_i	73	73	67	75	70.0	62	76.0	73	71.0	66	76.0	73	74	74	70.0	76.0	72	69.0	76.0
$X_i - \tilde{x}_0$	3	3	-3	5	0.0	-8	6.0	3	1.0	-4	6.0	3	4	4	0.0	6.0	2	-1.0	6.0
I	+	+	-	+	-	-	+	+	+	-	+	+	+	+	-	+	+	-	+
$ X_i - \tilde{x}_0 $	3	3	3	5	-	8	6.0	3	1.0	4	6.0	3	4	4	-	6.0	2	1.0	6.0
R_i	6	6	6	12	-	17	14.5	6	1.5	10	14.5	6	10	10	14.5	-	3	1.5	14.5

Z tabulky 13 vidíme, že celkem 13 rozdílů $X_i - \tilde{x}_0$, $i = 1, \dots, 17$ je kladných, přičemž pořadí těchto rozdílů v absolutních hodnotách jsou 6, 6, 12, 14.5, 6, 1.5, 14.5, 6, 10, 10, 14.5, 3, 14.5. Hodnota testovací statistiky S_E bude tedy rovná součtu těchto pořadí.

$$S_E = \sum_{i=1}^{17} R_i I(X_i - \tilde{x}_0 > 0) = 6 + 6 + 12 + 14.5 + 6 + 1.5 + 14.5 + 6 + 10 + 10 + 14.5 + 3 + 14.5 = 118.5$$

Tabulku analogickou tabulce 13 nyní vytvoříme také pomocí softwaru `R`. Z výše uvedeného textu víme, že vektor X_i (`upface.HN`), a $X_i - 70$ mají délku $n = 19$. Nicméně vektory I , $|X_i - 70|$ a R_i mají mít délku $m = 17$. Zároveň ale víme, že `R` umí vytvořit tabulku pouze ze řádků, které mají stejnou délku. Situaci tedy vyřešíme tak, že vytvoříme vektor `upface.HN0`, který bude mít délku 19, nicméně pozice, kde $X_i - 70 = 0$, nahradíme hodnotami NA. Indikační proměnná I , pro kterou platí $I = 1$ pokud $Z_i - 0 > 0$ a $I = 0$ jinak bude mít opět délku 19, ale dvě pozice budou opět nabývat hodnoty NA. Stejná situace nastane u vektoru pořadí rozdílů v absolutních hodnotách R_i . Vektor pořadí stanovíme pomocí funkce `rank()` s volbou argumentu `na.last = "keep"`, která zajistí, aby NA hodnoty získali jako pořadí rovněž NA hodnotu.

- **Kritický obor** Horní a dolní hranici kritického oboru stanovíme pomocí funkce `qsignrank()`.

$$\begin{aligned} W &= \langle s_m(1 - \alpha); \infty \rangle \\ &= \langle s_{17}(1 - 0.10); \infty \rangle \\ &= \langle s_{17}(0.9); \infty \rangle \\ &= \langle 111; \infty \rangle \\ &= \langle 111; \infty \rangle \end{aligned}$$

• Závěr testování

Protože realizace testovací statistiky $s_E = 118.5$ náleží do kritického oboru, tj. $s_E \in W$, H_0 zamítáme na hladině významnosti $\alpha = 0.10$.

4. Testování intervalem spolehlivosti

```

410 # upface.HN1 <- upface.HN
411 x0 <- 70.00
412 n <- length(upface.HN)
413 m <- sum(upface.HN - x0 != 0)
414 upface.HN0 <- upface.HN
415 upface.HN0[upface.HN0 - x0 == 0] <- NA
416
417 I <- (upface.HN0 - x0 > 0)
418 Ri <- rank(abs(upface.HN0 - x0), na.last = "keep")
419 tab <- rbind("Xi" = upface.HN, "Xi-x0" = upface.HN - x0,
420             "I" = I, "|Xi-x0|" = abs(upface.HN0 - x0),
421             "Ri" = Ri)
422 tab <- data.frame(tab)
423 names(tab) <- 1 : n
424 #
425 # Xi      73 73 67 75 70 62 76.0 73 71.0 66 76.0 73 74 74 70 76.0 72 69.0 76.0
426 # Xi-x0   3 3 -3 5 0 -8 6.0 3 1.0 -4 6.0 3 4 4 0 6.0 2 -1.0 6.0
427 # I       1 1 0 1 NA 0 1.0 1 1.0 0 1.0 1 1 1 NA 1.0 1 0.0 1.0
428 # |Xi-x0| 3 3 3 5 NA 8 6.0 3 1.0 4 6.0 3 4 4 NA 6.0 2 1.0 6.0
429 # Ri      6 6 6 12 NA 17 14.5 6 1.5 10 14.5 6 10 10 NA 14.5 3 1.5 14.5
430 SE <- sum(Ri * I, na.rm = T) # 118.5

431 alpha <- 0.10
432 qsignrank(1 - alpha, m) # 104

```

- **Interval spolehlivosti**

Abychom mohli stanovit hranice 90% levostranného intervalu spolehlivosti, musíme nejprve vypočítat $\frac{m(m+1)}{2}$ Walshových průměrů $\frac{X_i+X_j}{2}$, $i = 1, \dots, m$. Walshovy průměry získáme pomocí funkce `owa` z knihovny `NSM3`. Prvním argumentem této funkce bude vektor nul o délce $m = 17$, druhým argumentem bude vektor `upface.HN0` o délce $m = 17$. Získané Walshovy průměry seřadíme příkazem `sort()`.

```

433 walsh <- NSM3::owa(rep(0, m), na.omit(upface.HN0))$owa
434 V <- sort(walsh)

```

Hranice intervalu spolehlivosti potom tvoří hodnota, která se v seřazeném vektoru Walshových průměrů V nachází na $(s_m(\alpha))$ -té pozici a nekonečno. Hodnotu $s_m(\alpha)$ nalezneme pomocí funkce `qsignrank()`.

$$\begin{aligned}
 (d, h) &= \left(V^{(s_m(\alpha))}; \infty \right) \\
 &= \left(V^{(s_{17}(0.10))}; \infty \right) \\
 &= \left(V^{(49)}; \infty \right) \\
 &= (71; \infty)
 \end{aligned}$$

- **Závěr testování**

Protože $\tilde{x}_0 = 70.00$ nenáleží do 90% empirického levostranného intervalu spolehlivosti, tj. $\tilde{x}_0 = 70.00 \notin IS, H_0$ zamítáme na hladině významnosti $\alpha = 0.10$.

5. Testování p -hodnotou

Příslušnou p -hodnotu vypočítáme pomocí vzorce $\Pr(S_E \geq s_E)$. Zde si uvědomme, že realizace testovací statistiky $s_E = 118.5$. Zároveň S_E je diskrétní náhodná veličina. Z vlastností pravděpodobnostní funkce diskrétních náhodných veličin víme, že $\Pr(S_E \geq 118.5) = \Pr(S_E \geq 119) = 1 - \Pr(S_E < 119) = 1 - \Pr(S_E \leq 118)$. Viz kapitola ??.

```

435 alpha <- 0.10
436 qsignrank(alpha, m) # 49
437 V[49]

```

- **p -hodnota**

$$p\text{-hodnota} = \Pr(S_E \geq s_E) = \Pr(S_E \geq 118.5) = 1 - \Pr(S_E \leq 118) = 1 - 0.9776154 = 0.02238464 \doteq 0.02238$$

```

438 1 - psignrank(118, m) # 0.02238464

```

- **Závěr testování**

Protože p -hodnota = 0.02238 je menší než $\alpha = 0.10$, H_0 zamítáme na hladině významnosti $\alpha = 0.10$.

6. Interpretace výsledků

Za základě všech tří typů testování zamítáme nulovou hypotézu na hladině významnosti $\alpha = 0.10$. Výška horní části tváře mužů německé populace je statisticky významně větší než výška horní části tváře mužů Černjachovské populace z oblasti dnešní Ukrajiny. Ke stejnému závěru jsme dospěli také v příkladu 9.2.

7. Grafická vizualizace výsledku testování

Významný rozdíl horní části tváře mužů mezi oběma populacemi vizualizujeme analogicky jako v příkladu 9.2 krabicovým diagramem (viz obrázek 4).

Poznámka: K provedení Wilcoxonova jednovýběrového exaktního testu použijeme funkci `wilcox.test()`. Vstupními parametry budou původní vektor délky $n = 19$ reprezentující náhodný výběr (`upface.HN`), hodnota parametru \tilde{x}_0 z nulové hypotézy (`mu = 70.00`), argument `conf.int = T` zadávající požadavek na výpočet intervalu spolehlivosti, hodnota hladiny významnosti α zadaná prostřednictvím koeficientu spolehlivosti $1 - \alpha$ (`conf.level = 0.90`), typ zvolené alternativní hypotézy (pravostranná, `alternative = 'greater'`) a argument `correct = F`, který zakazuje aplikování spojité korekce na výsledné intervaly spolehlivosti a p -hodnotu.

```

439 wilcox.test(upface.HN, mu = 70.00, conf.int = T, conf.level = 0.90,
440             alt = 'greater', correct = F) # IS i p-hodonota interpolovane

```

Součástí výstupu je hodnota mediánu náhodného výběru `median of x = 72.99994`, hodnota testovací statistiky `s = 118.5`, interpolované hranice 90% empirického levostranného intervalu spolehlivosti `71.00003` a `Inf` a interpolovaná p -hodnota `p-value = 0.02286`. Jediné, co musíme stanovit zvlášť, je dolní hranice kritického oboru.

★

```
Wilcoxon signed rank test
data: upface.HN
V = 118.5, p-value = 0.02286
alternative hypothesis: true location is greater than 70
90 percent confidence interval:
 71.00003      Inf
sample estimates:
(pseudo)median
 72.99994
```

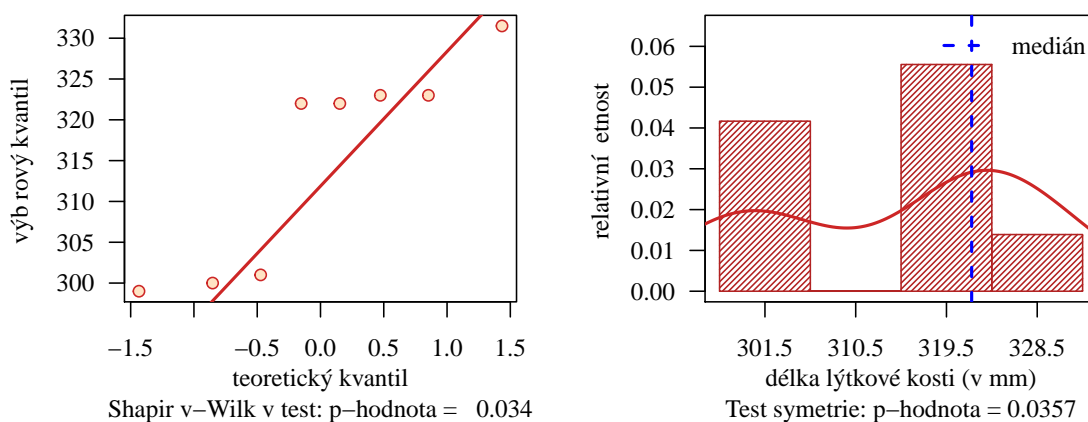
441
442
443
444
445
446
447
448
449
450
451

Příklad 9.15. Wilxonův jednovýběrový exaktní test

Mějme datový soubor 21-goldman-measures.csv a proměnnou tibia.LR popisující největší délku lýtkové kosti z pravé strany v mm u skeletů z období neolitu z oblasti *Yoshigo* (viz sekce ??). Dále máme k dispozici údaje ze studie (Hasegawa et al.) z roku 2009, v rámci které byla měřena největší délka lýtkové kosti z pravé strany u žen současné japonské populace ($m_f = 329.40$ mm, $s_f = 17.3$ mm, $n_f = 342$). Na hladině významnosti $\alpha = 0.01$ zjistěte, zda je největší délka lýtkové kosti z pravé strany u žen z neolitické japonské populace významně menší než u žen současné japonské populace.

Řešení příkladu 9.15

Zadání příkladu je shodné se zadáním příkladu 9.3. Datový soubor obsahující údaje o největší délce lýtkové kosti z pravé strany nepochází z normálního rozdělení (p -hodnota = 0.034, viz obrázek 9.3), proto na prozkoumání otázky ze zadání použijeme neparametrický jednovýběrový test. Opět bychom chtěli nulovou hypotézu otestovat pomocí Wilxonova jednovýběrového testu, k čemuž je potřeba ověřit předpoklad symetrického rozdělení naměřených hodnot okolo mediánu. Ten ověříme pomocí testu symetrie ($\alpha = 0.05$) a histogramem superponovaným jádrovým odhadem křivky hustoty se zvýrazněnou hodnotou mediánu pomocí vertikální přímky (viz obrázek 27).



Obrázek 27: QQ-diagram (vlevo) a histogram naměřených hodnot délky lýtkové kosti na pravé straně u skeletů žen populace z oblasti Youshigo Shell Mound superponovaný křivkou jadrového odhadu hustoty (vpravo)

Protože p -hodnota = 0.0357 je menší než 0.05, zamítáme hypotézu o symetrii naměřených hodnot okolo mediánu na hladině významnosti $\alpha = 0.05$. Z grafu 27 vidíme, že data jsou nesymetrická okolo mediánu, a to natolik že i přes nízký rozsah náhodného výběru dochází k zamítnutí hypotézy o symetrii. Naměřené hodnoty největší délky lýtkové kosti z pravé strany u žen nejsou symetricky rozdělené okolo mediánu. Předpoklad pro použití Wilxonova testu není splněn a proto tento test nemůžeme v tomto případě použít. Musíme se tedy spokojit s výsledkem znaménkového testu použitého v rámci příkladu 9.3. ★

9.5 Wilcoxonův jednovýběrový asymptotický test

Pro náhodné výběry o rozsazích $n \geq 30$ máme možnost použít k otestování nulové hypotézy asymptotickou variantu testu. Tuto variantu nazýváme Wilcoxonův jednovýběrový asymptotický test o mediánu \tilde{x} . Testovací statistika asymptotického testu má tvar

$$S_A = \frac{S_E - \frac{m(m+1)}{4}}{\sqrt{\frac{m(m+1)(2m+1)}{24}}} \quad (9.5)$$

kde S_E je statistika definovaná vztahem 9.4 a m je počet nenulových rozdílů $X_i - \tilde{x}_0$. Za platnosti nulové hypotézy pochází statistika S_A ze standardizovaného normálního rozdělení, tj.

$$S_A = \frac{S_E - \frac{m(m+1)}{4}}{\sqrt{\frac{m(m+1)(2m+1)}{24}}} \stackrel{H_0}{\sim} N(0, 1).$$

Kritický obor podle zvolené alternativní hypotézy má tvar

$$\begin{aligned} H_{11} : \tilde{x} &\neq \tilde{x}_0 & W &= (-\infty; u_{\alpha/2}) \cup (u_{1-\alpha/2}; \infty) \\ H_{12} : \tilde{x} &> \tilde{x}_0 & W &= (u_{1-\alpha}; \infty) \\ H_{13} : \tilde{x} &< \tilde{x}_0 & W &= (-\infty; u_{\alpha}) \end{aligned}$$

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

Interval spolehlivosti má podle zvolené alternativní hypotézy jeden z následujících tvarů

$$\begin{aligned} H_{11} : \tilde{x} &\neq \tilde{x}_0 & (d, h) &= (V^{(C_{1-\alpha/2})}; V^{(C_{\alpha/2})}) \\ H_{12} : \tilde{x} &> \tilde{x}_0 & (d, \infty) &= (V^{(C_{1-\alpha})}; \infty) \\ H_{13} : \tilde{x} &< \tilde{x}_0 & (-\infty, h) &= (-\infty; V^{(C_{\alpha})}) \end{aligned}$$

kde $C_{1-\alpha/2} = \frac{m(m+1)}{4} - u_{1-\alpha/2} \sqrt{\frac{m(m+1)(2m+1)}{24}}$, $C_{\alpha/2} = \frac{m(m+1)}{4} - u_{\alpha/2} \sqrt{\frac{m(m+1)(2m+1)}{24}}$, $C_{1-\alpha} = \frac{m(m+1)}{4} - u_{1-\alpha} \sqrt{\frac{m(m+1)(2m+1)}{24}}$, $C_{\alpha} = \frac{m(m+1)}{4} - u_{\alpha} \sqrt{\frac{m(m+1)(2m+1)}{24}}$, $V(1) \leq \dots \leq V^{(\frac{m(m+1)}{2})}$ značí posloupnost vzestupně seřazených $\frac{m(m+1)}{2}$ Walshových průměrů $\frac{(X_i + X_j)}{2}$, $i = 1, \dots, m$, $j = 1, \dots, m$, $j \leq i$ a $V^{(k)}$ značí k -tý seřazený Walshův průměr. Posloupnost Walshových průměrů získáme příkazem `owa` z knihovny `NSM3`.

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

$$\begin{aligned} H_{11} : \tilde{x} &\neq \tilde{x}_0 & p\text{-hodnota} &= 2 \min\{\Pr(S_A \leq s_A), \Pr(S_A \geq s_A)\} \\ H_{12} : \tilde{x} &> \tilde{x}_0 & p\text{-hodnota} &= \Pr(S_A \geq s_A) \\ H_{13} : \tilde{x} &< \tilde{x}_0 & p\text{-hodnota} &= \Pr(S_A \leq s_A) \end{aligned}$$

kde S_A je náhodná veličina, s_A je realizace testovací statistiky S_A (viz vzorec 9.5), tedy konkrétní číslo, a $\Pr(S_A \leq s_A)$ je distribuční funkce standardizovaného normálního rozdělení, jejíž hodnotu získáme pomocí \mathcal{R} a implementované funkce `pnorm()`.

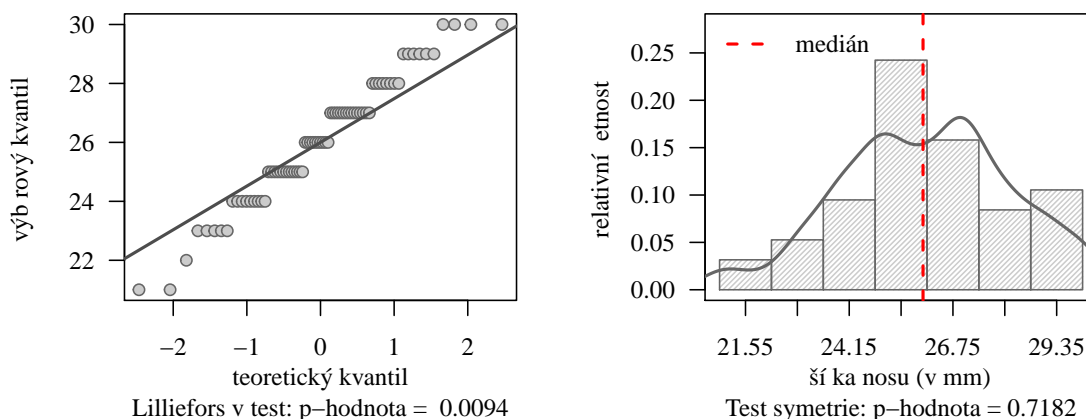
Příklad 9.16. Wilcoxonův jednovýběrový asymptotický test

Mějme datový soubor 19-more-samples-correlations-skull.txt a proměnnou nose.B popisující šířku nosu v mm (viz sekce ??). Dále máme k dispozici údaje o šířce nosu mužů současné malajské populace ($m_m = 26.90$ mm, $n_m = 45$) uveřejněné ve studii (Ibrahim, 2017). Na hladině významnosti $\alpha = 0.01$ testujte hypotézu o shodě šířky nosu starověké malajské populace a současné malajské populace.

Řešení příkladu 9.16

Zadání příkladu je shodné se zadáním příkladu 9.4. Datový soubor obsahující údaje o největší šířce nosu 73 mužů malajské populace nepochází z normálního rozdělení (p -hodnota = 0.0094, viz obrázek 9.16), proto na otestování hypotézy ze zadání použijeme neparametrický jednovýběrový test. Nulovou hypotézu bychom tentokrát chtěli otestovat pomocí Wilcoxonova jednovýběrového testu. Abychom tento test mohli použít, je potřeba nejprve ověřit předpoklad symetrického rozdělení naměřených hodnot okolo mediánu. Ten ověříme pomocí testu symetrie ($\alpha = 0.05$). Graficky vizualizujeme míru symetrie naměřených hodnot okolo mediánu pomocí histogramu superponovaného jádrovým odhadem křivky hustoty (viz kapitola ??) se zvýrazněnou hodnotou mediánu pomocí vertikální příčky (viz obrázek 28).

```
452 data <- read.delim('00-Data//19-more-samples-correlations-skull.txt')
453 nose.BM <- data[data$pop == 'mal', 'nose.B']
454 nose.BM <- na.omit(nose.BM)
455 n <- length(nose.BM)
```



Obrázek 28: Míra symetrie naměřených hodnot šířky nosu mužů malajské populace okolo mediánu

Na základě testu symetrie hypotézu o symetrickém rozdělení naměřených hodnot největších šířek nosu u mužů malajské populace nezamítáme na hladině významnosti $\alpha = 0.05$ (p -hodnota = 0.7182). Histogramu na obrázku 28 taktéž naznačuje, že data jsou okolo mediánu rozmístěná symetricky. Předpoklad k použití Wilcoxonova neparametrického jednovýběrového testu je tedy splněn, a my můžeme přistoupit k testování hypotézy ze zadání. Ta je, stejně jako alternativní hypotéza a hladina významnosti α analogická jako v řešení příkladu 9.4. Testování provedeme v posloupnosti sedmi kroků.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : Medián šířky nosu mužů malajské populace je shodný s mediánem šířky nosu mužů současné malajské populace.

H_1 : Medián šířky nosu mužů malajské populace není shodný s mediánem šířky nosu mužů současné malajské populace.

- **matematická formulace** nulové a alternativní hypotézy

$H_0 : \tilde{x} = \tilde{x}_0$, kde $\tilde{x}_0 = 26.90$

$H_1 : \tilde{x} \neq \tilde{x}_0$, kde $\tilde{x}_0 = 26.90$ (levostranná alternativa)

2. Volba hladiny významnosti

- Hladina významnosti $\alpha = 0.01$.

3. Testování kritickým oborem

- **Testovací statistika**

Nejprve vypočítáme vektor rozdílů naměřených hodnot X_i a konstanty \tilde{x}_0 , tj. $X_i - \tilde{x}_0 = X_i - 26.90$. Následně se zaměříme na stanovení počtu nenulových rozdílů, kde zjistíme, že žádný z rozdílů $X_i - 26.90$ není nulový, tedy počet nenulových rozdílů $m = n = 73$. Pro dodržení obecného postupu vytvoříme nyní vektor nose.BM0 obsahující ty naměřené šířky nosu, pro něž rozdíl $X_i - \tilde{x}_0$ nebyl rovný nule. V tomto případě bude vektor nose.BM0 shodný s vektorem nose.BM, protože žádný z rozdílů $X_i - \tilde{x}_0$ nulový nebyl. V dalších krocích testování budeme nicméně používat vektor nose.BM0. Následně vypočítáme absolutní hodnoty nenulových rozdílů, tj. $|X_i - \tilde{x}_0|$, $i = 1, \dots, m$ a stanovíme pořadí těchto rozdílů v absolutních hodnotách.

```
456 x0 <- 26.90
457 n <- length(nose.BM)
458 m <- sum(nose.BM - x0 != 0)
459 nose.BM0 <- nose.BM
460 nose.BM0[nose.BM0 - x0 ==0] <- NA
461
462 I <- (nose.BM0 - x0 > 0)
463 Ri <- rank(abs(nose.BM0 - x0), na.last = "keep")
464 tab <- rbind("Xi" = nose.BM, "Xi-x0" = nose.BM - x0,
465             "I" = I, "|Xi-x0|" = abs(nose.BM0 - x0),
466             "Ri" = Ri)
467 tab <- data.frame(tab)
468 names(tab) <- 1 : n
```

K výpočtu testovací statistiky S_A musíme nejprve vypočítat testovací statistiku S_E . Protože kladných rozdílů máme celkem 33, přičemž pořadí absolutních hodnot těchto rozdílů jsou 72.5, 20.5, 71, 40, 8, 57, 40, 57, 29.5, 20.5, 63.5, 49.5, 57, 57, 20.5, 20.5, 20.5, 49.5, 68, 8, 49.5, 40, 49.5, 40, 8, 8, 8, 20.5, 40, 49.5, 20.5, 8, 20.5, 40, 40, 8, 8, 68, 57, 8, 20.5, 63.5, 29.5, 29.5, 49.5, 8, 40, 57, 40, 63.5, 63.5, 29.5, 40, 20.5, 8, 68, 8, 72.5, 40, 40, 8, 57, 29.5, 40, 29.5, 57, 29.5, 68, 29.5, 8, 57, 8 a 68, bude hodnota testovací statistiky S_E , definované jako součet pořadí absolutních hodnot $|X_i - 26.90|$ kladných rozdílů, rovna 907, tj.

$$S_E = \sum_{i=1}^{12} R_i I(X_i - \tilde{x}_0 > 0) = 72.5 + 20.5 + 71 + \dots + 57 + 8 + 68 = 907.$$

Nyní dopočítáme hodnotu testovací statistiky S_A .

$$\begin{aligned} S_A &= \frac{S_E - \frac{m(m+1)}{4}}{\sqrt{\frac{m(m+1)(2m+1)}{24}}} = \frac{907 - \frac{73(73+1)}{4}}{\sqrt{\frac{73(73+1)(2 \times 73+1)}{24}}} = \frac{907 - \frac{73 \times 74}{4}}{\sqrt{\frac{73 \times 74 \times 147}{24}}} = \frac{907 - 1350.5}{\sqrt{33087.25}} = \frac{-443.5}{181.899} \\ &= -2.438166 \doteq -2.4382 \end{aligned}$$


```

469 SE <- sum(Ri * I) # 907
470 SA <- (SE - m * (m + 1) / 4) / sqrt(m * (m + 1) * (2 * m + 1) / 24) # -2.438166

```

- **Kritický obor**

Protože $S_A \sim N(0, 1)$, horní a dolní hranici kritického oboru nalezneme pomocí funkce `qnorm()` jako $\alpha/2$ kvantil a $1 - \alpha/2$ kvantil standardizovaného normálního rozdělení $N(0, 1)$.

$$\begin{aligned}
W &= (-\infty; u_{\alpha/2}) \cup (u_{1-\alpha/2}; \infty) \\
&= (-\infty; u_{0.01/2}) \cup (u_{1-0.01/2}; \infty) \\
&= (-\infty; u_{0.005}) \cup (u_{0.995}; \infty) \\
&= (-\infty; -2.575829) \cup (2.575829; \infty) \\
&= (-\infty; -2.5758) \cup (2.5758; \infty)
\end{aligned}$$

```

471 alpha <- 0.01
472 qnorm(alpha / 2) # -2.575829
473 qnorm(1 - alpha / 2) # 2.575829

```

- **Závěr testování**

Protože realizace testovací statistiky $s_A = -2.4382$ nenáleží do kritického oboru, tj. $s_A \notin W$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

4. Testování intervalem spolehlivosti

- **Interval spolehlivosti**

Abychom mohli stanovit hranice 99% intervalu spolehlivosti, musíme nejprve vypočítat $\frac{m(m+1)}{2}$ Walshových průměrů $\frac{X_i+X_j}{2}$, $i = 1, \dots, m$, $j = 1, \dots, m$, $j \leq i$ a ty následně vzestupně seřadit. Walshovy průměry získáme pomocí funkce `owa` z knihovny `NSM3`. Prvním argumentem této funkce bude vektor nul o délce $m = 73$, druhým argumentem bude vektor `nose.BM0` o délce $m = 73$. Získané Walshovy průměry seřadíme příkazem `sort()`. Následně vypočítáme pozicové souřadnice $C_{1-\alpha/2}$ a $C_{\alpha/2}$.

$$\begin{aligned}
C_{1-\alpha/2} &= \frac{m(m+1)}{4} - u_{1-\alpha/2} \sqrt{\frac{m(m+1)(2m+1)}{24}} \\
&= \frac{73(73+1)}{4} - u_{1-0.01/2} \sqrt{\frac{73(73+1)(2 \times 73+1)}{24}} \\
&= \frac{73 \times 74}{4} - u_{0.995} \sqrt{\frac{73 \times 74 \times 147}{24}} \\
&= 1350.5 - 2.575829 \sqrt{33087.25} \\
&= 1350.5 - 2.575829 \times 181.899 = 881.9593 \doteq 882
\end{aligned}$$

$$\begin{aligned}
C_{\alpha/2} &= \frac{m(m+1)}{4} - u_{\alpha/2} \sqrt{\frac{m(m+1)(2m+1)}{24}} \\
&= \frac{73(73+1)}{4} - u_{0.01/2} \sqrt{\frac{73(73+1)(2 \times 73+1)}{24}} \\
&= \frac{73 \times 74}{4} - u_{0.005} \sqrt{\frac{73 \times 74 \times 147}{24}} \\
&= 1350.5 + 2.575829 \sqrt{33087.25} \\
&= 1350.5 + 2.575829 \times 181.899 = 1819.041 \doteq 1819
\end{aligned}$$

Hranice intervalu spolehlivosti potom tvoří ty hodnoty, které se v seřazeném vektoru Walshových průměrů V nachází na 882. pozici a na 1819. pozici.

$$\begin{aligned}
(d, h) &= \left(V^{(C_{1-\alpha/2})}; V^{(C_{\alpha/2})} \right) \\
&= \left(V^{(882)}; V^{(1819)} \right) \\
&= (25.5; 27.0)
\end{aligned}$$

```

474 walsh <- NSM3::owa(rep(0, m), nose.BM0)$owa
475 V <- sort(walsh)
476 C1 <- m * (m + 1) / 4 - qnorm(1 - alpha / 2) * sqrt(m * (m + 1) * (2 * m + 1) / 24) #
      881.9592
477 C2 <- m * (m + 1) / 4 - qnorm(alpha / 2) * sqrt(m * (m + 1) * (2 * m + 1) / 24) #
      1819.041
478 V[882] # 25.5
479 V[1819] # 27

```

- **Závěr testování**

Protože $\tilde{x}_0 = 26.90$ náleží do 99% empirického oboustranného intervalu spolehlivosti, tj. $\tilde{x}_0 = 50.00 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0.01$.

5. Testování p -hodnotou

Příslušnou p -hodnotu vypočítáme pomocí vzorce $2 \min\{\Pr(S_A \leq s_A), \Pr(S_A \geq s_A)\}$. Zde si uvědomme, že realizace testovací statistiky $s_A = -2.4382$. Zároveň S_A je spojitá náhodná veličina, neboť pochází ze spojitého (standardizovaného normálního) rozdělení $N(0, 1)$. Z vlastností pravděpodobnostní funkce diskrétních náhodných veličin víme, že $\Pr(S_A \geq -2.4382) = 1 - \Pr(S_A < -2.4382) = \Pr(S_A \leq -2.4382)$. Viz kapitola ??.

- **p -hodnota**

$$\begin{aligned}
p\text{-hodnota} &= 2 \min\{\Pr(S_A \leq s_A), \Pr(S_A \geq s_A)\} \\
&= 2 \min\{\Pr(S_A \leq -2.4382), 1 - \Pr(S_A \leq -2.4382)\} \\
&= 2 \min\{0.03198242, 0.9680176\} \\
&= 2 \times 0.03198242 = 0.01476199 \doteq 0.01476
\end{aligned}$$

- **Závěr testování**

Protože p -hodnota = 0.01476 je větší než $\alpha = 0.01$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

```
480 2 * min (pnorm(SA), 1 - pnorm(SA)) # 0.01476199
```

6. Interpretace výsledků

Za základě všech tří typů testování nezamítáme nulovou hypotézu na hladině významnosti $\alpha = 0.01$. Mezi šířkou nosu mužů starověké a současné malajské populace není statisticky významný rozdíl. Ke stejnému závěru jsme dospěli také v příkladu 9.4.

7. Grafická vizualizace výsledku testování

Grafické porovnání šířky nosu obou malajských populací provedeme analogicky jako v příkladu 9.4 krabicovým diagramem (viz obrázek 8).

Poznámka: Wilcoxonův jednovýběrový asymptotický test můžeme provést pomocí funkce `wilcox.test()`. Vstupními parametry budou vektor reprezentující náhodný výběr (`nose.BM`), hodnota parametru \tilde{x}_0 z nulové hypotézy zadaná argumentem `mu = 26.90`, požadavek na výpočet hranic intervalu spolehlivosti zadaný nastavením argumentu `conf.int = T`, hodnota hladiny významnosti α zadaná prostřednictvím koeficientu spolehlivosti $1 - \alpha$ nastavením hodnoty argumentu `conf.level = 0.99`, typ zvolené alternativní hypotézy (oboustranná) zadaný pomocí argumentu `alternative = 'two.sided'` a argument `correct = T`, který zvolí k výpočtu právě asymptotickou variantu Wilcoxonova neparametrického testu.

```
481 wilcox.test(nose.BM, mu = 26.90, alt = 'two.sided', conf.int = T,  
482             conf.level = 0.99, correct = T) # interpolovany IS i p-hodnota
```

```
Wilcoxon signed rank test with continuity correction  
  
data: nose.BM  
V = 907, p-value = 0.01462  
alternative hypothesis: true location is not equal to 26.9  
99 percent confidence interval:  
 25.49996 26.99993  
sample estimates:  
(pseudo)median  
 26.00002
```

483
484
485
486
487
488
489
490
491
492
493

Součástí výstupu je hodnota testovací statistiky S_E (nikoli S_A) $V = 907$, interpolované hranice 99% Waldova empirického oboustranného intervalu spolehlivosti 25.49996 a 26.99993 a interpolovaná p -hodnota $p\text{-value} = 0.01462$. Jediné, co musíme stanovit zvlášť, jsou dolní a horní hranice kritického oboru.



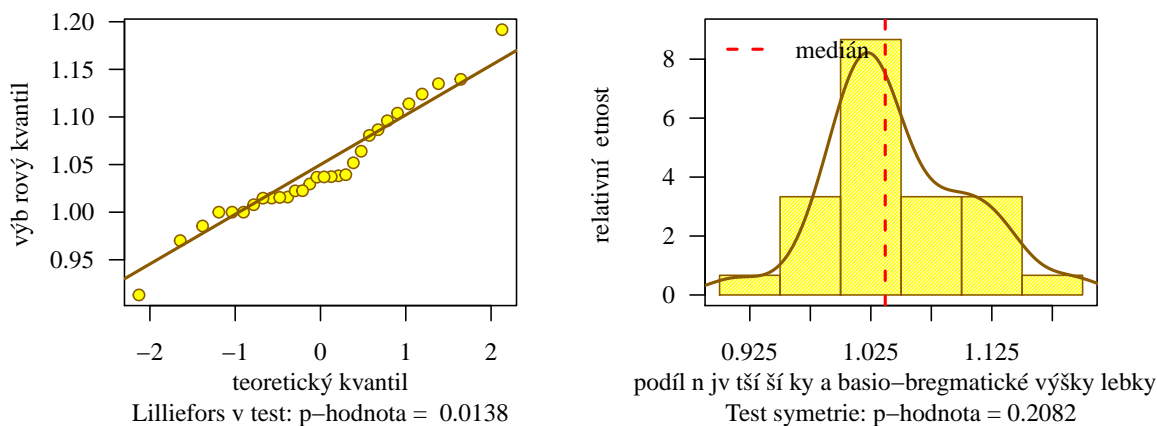
Příklad 9.17. Wilcoxonův jednovýběrový asymptotický test

Mějme datový soubor 28-one-sample-mean-skull-m5.txt a proměnné skull.H popisující basio-bregmatickou výšku lebky v mm a skull.B popisující největší šířku lebky v mm u skeletů mužského pohlaví z egyptské populace (viz sekce ??). Dále máme k dispozici údaje o basio-bregmatické výšce lebky ($m_H = 134.0$), o největší šířce lebky ($m_B = 131.4$) a o podílu největší šířky lebky a basio-bregmatické výšky lebky ($m_{BH} = 0.981$) u skeletů mužské egyptské populace z předdynastické doby. Na hladině významnosti $\alpha = 0.05$ zjistíte, zda je podíl největší šířky lebky a basio-bregmatické výšky lebky u mužů žijících v době okolo 150 let po našem letopočtu větší než u mužů žijících v předdynastické době.

Řešení příkladu 9.17

Zadání příkladu je shodné se zadáním příkladu 9.5. Datový soubor obsahující údaje o podílu největší šířky a basio-bregmatické výšky lebky 30 mužů egyptské populace z období okolo 150 let n.l. nepochází z normálního rozdělení (p -hodnota = 0.0138, viz obrázek ??), proto na otestování hypotézy ze zadání použijeme neparametrický jednovýběrový test. Otázku ze zadání bychom tentokrát chtěli ověřit pomocí Wilcoxonova jednovýběrového testu. Nejprve tedy ověříme předpoklad symetrického rozdělení naměřených podílů okolo jejich mediánu pomocí testu symetrie ($\alpha = 0.05$). Graficky vizualizujeme míru symetrie naměřených podílů okolo mediánu histogramem superponovaným jádrovým odhadem křivky hustoty se zvýrazněnou hodnotou mediánu pomocí vertikální přímkou (viz obrázek 29).

```
494 data <- read.delim('00-Data//28-one-sample-mean-skull-m5.txt')
495 # head(data)
496 data.BH <- data[data$year == 150, c('skull.B', 'skull.H')]
497 data.BH <- na.omit(data.BH)
498 skull.BH <- data.BH$skull.B / data.BH$skull.H
499 n <- length(skull.BH)
```



Obrázek 29: Míra symetrie podílů největší šířky a basio-bregmatické výšky lebky u mužů z egyptské populace okolo mediánu

Na základě testu symetrie hypotézu o symetrickém rozdělení naměřených podílů okolo mediánu nezamítáme na hladině významnosti $\alpha = 0.05$ (p -hodnota = 0.2082). Z histogramu na obrázku 29 můžeme rovněž vidět, že data jsou okolo mediánu rozmístěná symetricky. Předpoklad k použití Wilcoxonova neparametrického jednovýběrového testu je tedy splněn, a my můžeme přistoupit k ověření otázky ze zadání. Ta je, stejně jako v řešení příkladu 9.5 zněním alternativní hypotézy. Stejně tak nulová hypotéza a hladina významnosti jsou stejné jako v příkladu 9.5.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : Medián podílu největší šířky a basio-bregmatické výšky lebky u mužů žijících v době okolo 150 let

n.l. je menší nebo roven mediánu podílu největší šířky a basio-bregmatické výšky lebky u mužů žijících v předdynastické době.

H_1 : *Medián podílu největší šířky a basio-bregmatické výšky lebky u mužů žijících v době okolo 150 let n.l. je větší než medián podílu největší šířky a basio-bregmatické výšky lebky u mužů žijících v předdynastické době.*

- **matematická formulace** nulové a alternativní hypotézy

$$H_0 : \tilde{x} \leq \tilde{x}_0, \text{ kde } \tilde{x}_0 = 0.981$$

$$H_1 : \tilde{x} > \tilde{x}_0, \text{ kde } \tilde{x}_0 = 0.981 \text{ (pravostranná alternativa)}$$

2. Volba hladiny významnosti

- Hladina významnosti $\alpha = 0.05$.

3. Testování kritickým oborem

- **Testovací statistika**

Nejprve vypočítáme vektor rozdílů naměřených podílů X_i a konstanty \tilde{x}_0 , tj. $X_i - \tilde{x}_0 = X_i - 0.981$. Následně se zaměříme na stanovení počtu nenulových rozdílů, kde zjistíme, že žádný z rozdílů $X_i - 0.981$ není nulový, tedy počet nenulových rozdílů $m = n = 30$. Pro dodržení obecného postupu vytvoříme nyní vektor skull.BH0 obsahující ty podíly, pro něž rozdíl $X_i - \tilde{x}_0$ nebyl roven nule. V tomto případě bude vektor skull.BH0 shodný s vektorem skull.BH, protože žádný z rozdílů $X_i - \tilde{x}_0$ nulový nebyl. V dalších krocích testování budeme nicméně používat vektor skull.BH0. Následně vypočítáme absolutní hodnoty nenulových rozdílů, tj. $|X_i - \tilde{x}_0|$, $i = 1, \dots, m$ a stanovíme pořadí těchto rozdílů v absolutních hodnotách.

```
500 x0 <- 0.981
501 n <- length(skull.BH)
502 m <- sum(skull.BH - x0 != 0)
503 skull.BH0 <- skull.BH
504 skull.BH0[skull.BH0 - x0 ==0] <- NA
505
506 I <- (skull.BH0 - x0 > 0)
507 Ri <- rank(abs(skull.BH0 - x0), na.last = "keep")
508 tab <- rbind("Xi" = skull.BH, "Xi-x0" = skull.BH - x0,
509           "I" = I, "|Xi-x0|" = abs(skull.BH0 - x0),
510           "Ri" = Ri)
511 tab <- data.frame(tab)
512 names(tab) <- 1 : n
```

K výpočtu testovací statistiky S_A musíme nejprve vypočítat testovací statistiku S_E . Protože kladných rozdílů máme celkem 28, přičemž pořadí absolutních hodnot těchto rozdílů jsou 26, 17, 10, 23, 1, 4, 4, 13, 30, 14, 4, 11, 20, 16, 25, 8, 21, 27, 7, 9, 28, 22, 18, 24, 6, 15, 29 a 12, bude testovací statistika S_E , definovaná jako součet pořadí absolutních hodnot $|X_i - 0.981|$ kladných rozdílů, rovna 444, tj.

$$S_E = \sum_{i=1}^{12} R_i I(X_i - \tilde{x}_0 > 0) = 26 + 17 + 10 + \dots + 15 + 29 + 12 = 444.$$

Nyní dopočítáme hodnotu testovací statistiky S_A .

$$\begin{aligned} S_A &= \frac{S_E - \frac{m(m+1)}{4}}{\sqrt{\frac{m(m+1)(2m+1)}{24}}} = \frac{444 - \frac{30(30+1)}{4}}{\sqrt{\frac{30(30+1)(2 \times 30+1)}{24}}} = \frac{444 - \frac{30 \times 31}{4}}{\sqrt{\frac{30 \times 31 \times 61}{24}}} = \frac{444 - 232.5}{\sqrt{2363.75}} = \frac{211.5}{48.61841} \\ &= 4.350204 \doteq 4.3502 \end{aligned}$$

```
513 SE <- sum(Ri * I) # 444
514 SA <- (SE - m * (m + 1) / 4) / sqrt(m * (m + 1) * (2 * m + 1) / 24) # 4.350204
```

- **Kritický obor**

Protože $S_A \sim N(0, 1)$, kritický obor tvoří $1 - \alpha$ kvantil standardizovaného normálního rozdělení $N(0, 1)$ a nekonečno.

$$\begin{aligned} W &= \langle u_{1-\alpha}; \infty \rangle \\ &= \langle u_{1-0.05}; \infty \rangle \\ &= \langle u_{0.95}; \infty \rangle \\ &= \langle 1.644854; \infty \rangle \\ &= \langle 1.6449; \infty \rangle \end{aligned}$$

```
515 alpha <- 0.05
516 qnorm(1 - alpha) # 1.644854
```

- **Závěr testování**

Protože realizace testovací statistiky $s_A = 4.3502$ náleží do kritického oboru, tj. $s_A \in W$, H_0 zamítáme na hladině významnosti $\alpha = 0.05$.

4. Testování intervalem spolehlivosti

- **Interval spolehlivosti**

Proti pravostranné alternativní hypotéze postavíme levostranný interval spolehlivosti. Nejprve vypočítáme $\frac{m(m+1)}{2}$ Walshových průměrů $\frac{X_i + X_j}{2}$, $i = 1, \dots, m$, $j = 1, \dots, m$, $j \leq i$ a ty následně vzestupně seřadíme. Dále vypočítáme pozicovou souřadnici $C_{1-\alpha}$.

$$\begin{aligned} C_{1-\alpha} &= \frac{m(m+1)}{4} - u_{1-\alpha} \sqrt{\frac{m(m+1)(2m+1)}{24}} \\ &= \frac{30(30+1)}{4} - u_{1-0.05} \sqrt{\frac{30(30+1)(2 \times 30+1)}{24}} \\ &= \frac{30 \times 31}{4} - u_{0.95} \sqrt{\frac{30 \times 31 \times 61}{24}} \\ &= 232.5 - 1.644854 \sqrt{2363.75} \\ &= 232.5 - 1.644854 \times 48.61841 = 152.5298 \doteq 153 \end{aligned}$$

Hranice intervalu spolehlivosti potom tvoří hodnota, která se v seřazeném vektoru Walshových průměrů V nachází na 153. pozici a nekonečno.

$$\begin{aligned} (d, h) &= \left(V^{(C_{1-\alpha})}; \infty \right) \\ &= \left(V^{(137)}; \infty \right) \\ &= (1.02579; \infty) \qquad \doteq (1.0258; \infty) \end{aligned}$$

```

517 walsh <- NSM3::owa(rep(0, m), skull.BH0)$owa
518 V <- sort(walsh)
519 C1 <- m * (m + 1) / 4 - qnorm(1 - alpha) * sqrt(m * (m + 1) * (2 * m + 1) / 24) #
      152.5298
520 V[153] # 1.02579

```

- **Závěr testování**

Protože $\tilde{x}_0 = 0.981$ nenáleží do 95% empirického oboustranného intervalu spolehlivosti, tj. $\tilde{x}_0 = 0.981 \notin IS$, H_0 zamítáme na hladině významnosti $\alpha = 0.05$.

5. Testování p -hodnotou

Příslušnou p -hodnotu vypočítáme pomocí vzorce $\Pr(S_A \geq s_A)$, kde realizace testovací statistiky $s_A = 4.3502$. Zároveň S_A je spojitá náhodná veličina, neboť pochází ze spojitého (standardizovaného normálního) rozdělení $N(0, 1)$. Z vlastností pravděpodobnostní funkce spojitých náhodných veličin víme, že $\Pr(S_A \geq 4.3502) == 1 - \Pr(S_A < 4.3502) = \Pr(S_A \leq 4.3502)$. Viz kapitola ??.

- **p -hodnota**

$$p\text{-hodnota} = \Pr(S_A \geq s_A) = 1 - \Pr(S_A \leq 4.3502) = 6.800554 \times 10^{-6} \doteq 6.8006 \times 10^{-6}$$

```

521 1 - pnorm(SA) # 6.800554e-06

```

- **Závěr testování**

Protože p -hodnota = 6.8006×10^{-6} je menší než $\alpha = 0.05$, H_0 zamítáme na hladině významnosti $\alpha = 0.05$.

6. Interpretace výsledků

Za základě všech tří typů testování zamítáme nulovou hypotézu na hladině významnosti $\alpha = 0.05$. Podíl největší šířky a basio-bregmatické výšky lebky u mužů z egyptské populace z období okolo 150 let n.l. je statisticky významně větší než u mužů z předdynastické doby. Ke stejnému závěru jsme dospěli také v příkladu 9.5.

7. Grafická vizualizace výsledku testování

Grafické porovnání podílu největší šířky a basio-bregmatické výšky lebky u mužů z obou egyptských populací provedeme analogicky jako v příkladu 9.5 krabicovým diagramem (viz obrázek 10).

Poznámka: Wilcoxonův jednovýběrový asymptotický test můžeme provést pomocí funkce `wilcox.test()`. Vstupní parametry budou vektor reprezentující náhodný výběr podílů (`skull.BH`), hodnota parametru \tilde{x}_0 z nulové hypotézy zadaná argumentem `mu = 0.981`, požadavek na výpočet hranic intervalu spolehlivosti zadaný nastavením argumentu `conf.int = T`, hodnota hladiny významnosti α zadaná prostřednictvím koeficientu spolehlivosti $1 - \alpha$ nastavením hodnoty argumentu `conf.level = 0.95`, typ zvolené alternativní hypotézy (pravostranná) zadaný pomocí argumentu `alternative = 'greater'` a argument `correct = T`, který zvolí k výpočtu právě asymptotickou variantu Wilxonova neparametrického testu.

```

522 wilcox.test(skull.BH, mu = 0.981, alt = 'greater', conf.int = T,
523             conf.level = 0.95, correct = T) # interpolovany IS i p-hodnota

```

Součástí výstupu je hodnota testovací statistiky $S_E V = 444$, interpolované hranice 95% Waldova empirického jednostranného intervalu spolehlivosti 1.025749 a `lnf` a interpolovaná p -hodnota `p-value = 7.112e-06`. Jediné, co musíme stanovit zvlášť, jsou dolní a horní hranice kritického oboru.



```
Wilcoxon signed rank test with continuity correction
data: skull.BH
V = 444, p-value = 7.112e-06
alternative hypothesis: true location is greater than 0.981
95 percent confidence interval:
 1.025749      Inf
sample estimates:
(pseudo)median
      1.04075
```

524
525
526
527
528
529
530
531
532
533
534

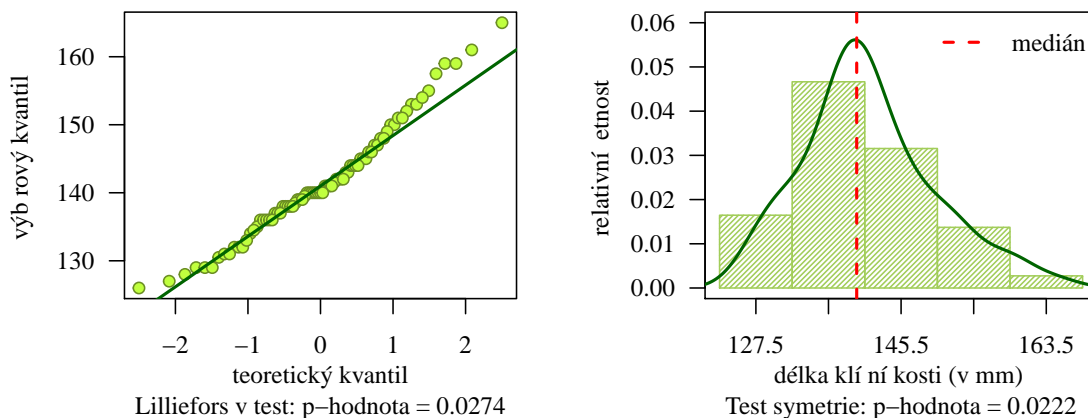
Příklad 9.18. Wilcoxonův jednovýběrový asymptotický test

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 ??). Dále máme k dispozici údaje o největší délce klíční kosti z pravé strany mužů z populace severní Indie ($m_I = 148.0$ mm, $s_I = 8.60$ mm, $n_I = 260$). Na hladině významnosti $\alpha = 0.05$ zjistěte, zda je délka klíční kosti z pravé strany u mužů indické populace z Varanasi menší než u mužů ze severní Indie.

Řešení příkladu 9.18

Zadání příkladu je shodné se zadáním příkladu 9.6. Datový soubor obsahující údaje o největší délce klíční kosti z pravé strany u mužů indické populace z Varanasi nepochází z normálního rozdělení (p -hodnota = 0.0274, viz obrázek 9.18), proto na prozkoumání otázky ze zadání použijeme neparametrický test, tentokrát konkrétně Wilcoxonův jednovýběrový test. K jeho použití je nejprve potřeba pomocí testu symetrie ověřit předpoklad symetrického rozdělení naměřených hodnot okolo jejich mediánu. Míru symetrie naměřených hodnot okolo mediánu si opět znázorníme také graficky (viz obrázek 30).

```
535 data <- read.delim('00-Data//18-more-samples-variances-clavicle.txt')
536 cla.LV <- data[data$pop == 'ind2', 'cla.L']
537 cla.LV <- na.omit(cla.LV)
538 n <- length(cla.LV)
```



Obrázek 30: Míra symetrie naměřených hodnot délky klíční kosti z pravé strany u mužů indické populace z Varanasi okolo mediánu

Na základě histogramu zobrazeného na obrázku 30 bychom usuzovali, že naměřené hodnoty jsou okolo svého mediánu rozloženy symetricky. Test symetrie nicméně hypotézu o symetrickém rozdělení naměřených hodnot zamítá na hladině významnosti $\alpha = 0.05$. Zde je dobré si uvědomit, že test o symetrii zohledňuje míru symetrie vzhledem k rozsahu náhodného výběru. Protože rozsah náhodného výběru je celkem vysoký ($n = 81$), byla míra symetrie, která by se na první pohled zdála uspokojující, vyhodnocena jako nedostatečná. Předpoklad symetrie naměřených délek klíčních kostí z pravé strany nebyl potvrzen, což znamená, že Wilcoxonův jednovýběrový test nemůžeme v tomto případě použít. Musíme se tedy spokojit se závěrem znaménkového testu uvedeným v řešení příkladu 9.6.

★

9.6 Wilcoxonův párový test

Nechť $(X_1, Y_1)^T \dots (X_n, Y_n)^T$ je náhodný výběr z libovolného (ne nutně normálního) dvourozměrného rozdělení. Nechť dále Z_1, \dots, Z_n , $n \geq 2$ je náhodný výběr rozdílů $X - Y$, tj. $Z = (Z_1, \dots, Z_n)^T$, kde $Z_i = X_i - Y_i$, $i = 1, \dots, n$, a nechť tento náhodný výběr pochází ze spojitého rozdělení s hustotou $f(x)$, která je symetrická okolo mediánu \tilde{z} . Konečně, nechť \tilde{z}_0 je konstanta. Na hladině významnosti α testujeme jednu z následujících tří hypotéz oproti příslušné alternativní hypotéze.

$H_{01} : \tilde{z} = \tilde{z}_0$	oproti	$H_{11} : \tilde{z} \neq \tilde{z}_0$	(oboustranná alt.)
$H_{02} : \tilde{z} \leq \tilde{z}_0$	oproti	$H_{12} : \tilde{z} > \tilde{z}_0$	(pravostranná alt.)
$H_{03} : \tilde{z} \geq \tilde{z}_0$	oproti	$H_{13} : \tilde{z} < \tilde{z}_0$	(levostranná alt.)

kde \tilde{z} je medián rozdílů Z_1, \dots, Z_n a \tilde{z}_0 je konstanta, jejíž hodnotu nejčastěji volíme jako $\tilde{z}_0 = 0$. Tato volba odpovídá hypotéze, že rozdíl mezi mediány náhodných veličin X a Y neexistuje (resp. hypotéze, že medián náhodné veličiny X je menší, resp. větší, než medián náhodné veličiny Y). Vzhledem k tomu, že jde finálně o situaci, kdy medián \tilde{z} porovnááme s konstantou \tilde{z}_0 , testujeme hypotézy o rozdílu mediánů $X - Y$ pomocí exaktní nebo asymptotické varianty Wilcoxonova jednovýběrového testu, analogicky jako je uvedeno v sekcích 9.4 a 9.5.

Výše popsaný test, v rámci kterého převádíme problém porovnávání mediánů dvou náhodných veličin X a Y na problém srovnávání mediánu jejich rozdílů Z s konstantou $\tilde{z}_0 = 0$ a následně jej řešíme pomocí exaktní resp. asymptotické varianty Wilcoxonova jednovýběrového testu, nazýváme párový Wilcoxonův test.

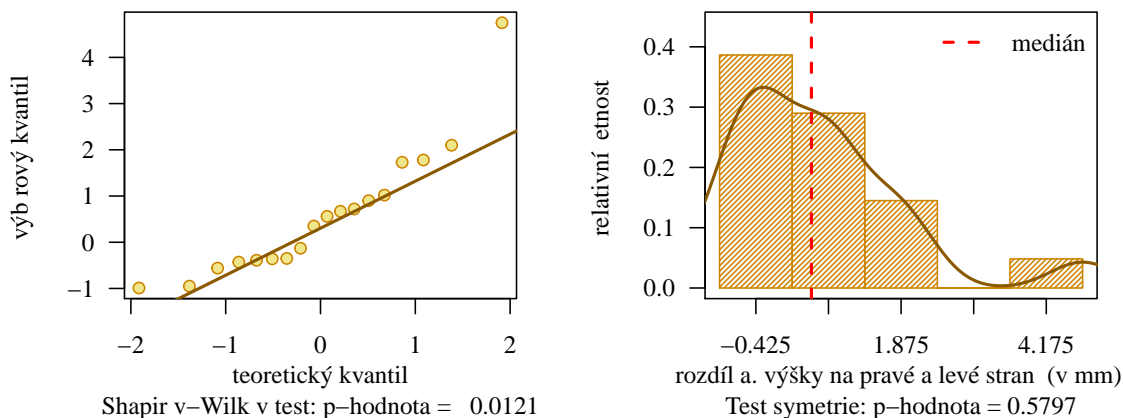
Příklad 9.19. Wilcoxonův párový exaktní test

Mějme datový soubor 21-goldman-measures.csv obsahující údaje o největší výšce kloubní jamky na pravé straně (acetab.HR) a na levé straně (acetab.HL) u skeletů ze starověké egyptské populace (detaily viz sekce ??). Na hladině významnosti $\alpha = 0.05$ zjistěte, zda existuje rozdíl mezi výškou kloubní jamky z pravé a levé strany u skeletů mužského pohlaví.

Řešení příkladu 9.19

Zadání příkladu je shodné se zadáním příkladu 9.7. Rozdíly naměřených největších šířek kloubní jamky z pravé a levé strany u 18 mužů ze starověké egyptské populace nepochází z normálního rozdělení (p -hodnota = 0.0121, viz obrázek 31), proto na prozkoumání otázky ze zadání použijeme neparametrický párový test. Nulovou hypotézu bychom tentokrát chtěli otestovat pomocí Wilcoxonova neparametrického párového testu, k čemuž je potřeba ověřit předpoklad symetrického rozdělení rozdílů okolo mediánu těchto rozdílů. Ten ověříme pomocí testu symetrie ($\alpha = 0.05$). Graficky vizualizujeme míru symetrie naměřených hodnot okolo mediánu pomocí histogramu superponovaného jádrovým odhadem křivky hustoty (viz kapitola ??) se zvýrazněnou hodnotou mediánu pomocí vertikální příčky (viz obrázek 31).

```
539 data <- read.delim('00-Data//21-goldman-measures.csv', sep = ';')
540 data.ARLM <- data[data$pop == 'Dynastic Egyptian, El Hesa' & data$sex == 'm',
541                   c('acetab.HL', 'acetab.HR')]
542 data.ARLM <- na.omit(data.ARLM)
543 acetab.HR <- data.ARLM$acetab.HR
544 acetab.HL <- data.ARLM$acetab.HL
545 acetab.HRL <- acetab.HR - acetab.HL
546 n <- length(acetab.HRL) # 18
```



Obrázek 31: Míra symetrie rozdílů největší acetabulární výšky u mužů na pravé a levé straně

Z histogramu na obrázku 31 bychom usuzovali, že symetrické rozdělení rozdílů okolo jejich mediánu není úplně přesvědčivé, nicméně test o symetrii hypotézu o symetrickém rozdělení rozdílů jednoznačně nezamítá (p -hodnota = 0.5797). Opět připomeňme, že test zohledňuje tentokrát menší rozsah náhodného výběru ($n = 18$), a proto je k zamítnutí hypotézy o symetrii benevolentnější. Předpoklad o symetrii rozdílů největších šířek kloubní jamky z pravé a levé strany je tedy splněn a Wilcoxonův párový test můžeme bez rizika omylu použít. Znění nulové a alternativní hypotézy, stejně jako hladina významnosti jsou analogické, jako v příkladu 9.7.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : Medián rozdílů výšky kloubní jamky na pravé a levé straně u mužů je rovný nule.

H_1 : Medián rozdílů výšky kloubní jamky na pravé a levé straně u mužů není rovný nule.

- **matematická formulace** nulové a alternativní hypotézy

$$H_0 : \tilde{z} = \tilde{z}_0, \text{ kde } \tilde{z}_0 = 0$$

$$H_1 : \tilde{z} \neq \tilde{z}_0, \text{ kde } \tilde{z}_0 = 0 \text{ (oboustranná alternativa)}$$

2. Volba hladiny významnosti

- Hladina významnosti $\alpha = 0.05$.

3. Testování kritickým oborem

- **Testovací statistika**

V úvodu příkladu jsme vytvořili vektor `acetab.HRL` ($Z_i, i = 1, \dots, 18$) jako vektor rozdílů výšky kloubní jamky na pravé straně a na levé straně. Dále jsme stanovili rozdíly $Z_i - \tilde{z}_0$. Protože ale $\tilde{z}_0 = 0$, jsou rozdíly $Z_i - \tilde{z}_0$ rovny původnímu vektoru rozdílů Z_i , tj. $Z_i - \tilde{z}_0 = Z_i - 0 = Z_i$. V dalším kroku se zaměříme na stanovení počtu nenulových rozdílů a zjistíme, že žádný z rozdílů $Z_i - 0$ není nulový (počet nenulových rozdílů $m = n = 18$). Kvůli zachování univerzálnosti postupu nyní vytvoříme vektor `acetab.HRL0` obsahující ty rozdíly mezi délkami látkových kostí na pravé a levé straně, pro než rozdíl $Z_i - \tilde{z}_0$ nebyl rovný nule. Délka tohoto vektoru bude v tomto případě $m = n = 18$ a vektor `acetab.HRL0` bude shodný s vektorem `acetab.HRL`. V dalších krocích testování budeme nicméně kvůli zachování konzistence postupu s postupy uvedenými v dalších příkladech používat právě vektor `acetab.HRL0`. Dále si označíme znaménkem $+$ rozdíly $Z_i - 0$, které jsou kladné a znaménkem $-$ rozdíly $Z_i - 0$, které jsou záporné, nebo 0 . Nakonec vypočítáme absolutní hodnoty rozdílů $|Z_i - 0| = |Z_i|, i = 1, \dots, m$ a stanovíme pořadí těchto rozdílů v absolutních hodnotách. Všechny výše popsané vektory jsou pro názornost uvedeny v tabulce 14.

Tabulka 14: Naměřené hodnoty X_i, Y_i , rozdíly $Z_i = X_i - Y_i$, rozdíly $Z_i - 0$, znaménka těchto rozdílů I , absolutní hodnoty nenulových rozdílů $|Z_i - 0|$ a pořadí R_i nenulových rozdílů v absolutních hodnotách

měření	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
X_i	50.84	52.36	47.80	51.75	46.06	46.64	50.54	47.03	49.02	47.96	48.25	47.78	49.86	49.56	55.41	52.82	52.47	46.02
Y_i	50.97	51.80	48.75	51.40	43.96	44.86	49.64	48.02	49.45	48.52	48.64	48.14	48.13	48.54	50.66	52.10	52.82	45.35
Z_i	-0.13	0.56	-0.95	0.35	2.10	1.78	0.90	-0.99	-0.43	-0.56	-0.39	-0.36	1.73	1.02	4.75	0.72	-0.35	0.67
$Z_i - 0$	-0.13	0.56	-0.95	0.35	2.10	1.78	0.90	-0.99	-0.43	-0.56	-0.39	-0.36	1.73	1.02	4.75	0.72	-0.35	0.67
I	-	+	-	+	+	+	+	-	-	-	-	-	+	+	+	+	-	+
$ Z_i - 0 $	0.13	0.56	0.95	0.35	2.10	1.78	0.90	0.99	0.43	0.56	0.39	0.36	1.73	1.02	4.75	0.72	0.35	0.67
R_i	1	7.5	12	2.5	17	16	11	13	6	7.5	5	4	15	14	18	10	2.5	9

Z tabulky 14 vidíme, že celkem deset rozdílů $Z_i - 0, i = 1, \dots, 18$ je kladných. Pořadí těchto rozdílů v absolutních hodnotách jsou 7.5, 2.5, 17, 16, 11, 15, 14, 18, 10 a 9. Hodnota testovací statistiky S_E , která je definovaná jako součet těchto pořadí, bude tedy rovná 120.

$$S_E = \sum_{i=1}^m I(Z_i - 0 > 0) = 7.5 + 2.5 + 17 + 16 + 11 + 15 + 14 + 18 + 10 + 9 = 120.$$

```

547 z0 <- 0
548 n <- length(acetab.HRL)
549 m <- sum(acetab.HRL - 0 != 0)
550 acetab.HRL0 <- acetab.HRL
551 acetab.HRL0[acetab.HRL0 == 0] <- NA
552 I <- (acetab.HRL0 > z0)
553 Ri <- rank(abs(acetab.HRL0 - z0))
554 tab <- data.frame(rbind("Xi" = acetab.HRL, "Yi" = acetab.HL,
555                          "Zi" = acetab.HRL, "Zi-z0" = acetab.HRL - z0, "I" = I,
556                          "|Zi-z0|" = abs(acetab.HRL0 - z0), "Ri" = Ri))
557 SE <- sum(Ri * I) # 120

```

- Kritický obor

$$\begin{aligned}
 W &= (-\infty; s_m(\alpha/2) - 1) \cup (s_m(1 - \alpha/2); \infty) \\
 &= (-\infty; s_{18}(0.05/2) - 1) \cup (s_{18}(1 - 0.05/2); \infty) \\
 &= (-\infty; s_{18}(0.025) - 1) \cup (s_{18}(0.975); \infty) \\
 &= (-\infty; 41 - 1) \cup (130; \infty) \\
 &= (-\infty; 40) \cup (130; \infty)
 \end{aligned}$$

```

558 alpha <- 0.05
559 q1 <- qsignrank(alpha / 2 , m) - 1 # 40
560 q2 <- qsignrank(1 - alpha / 2 , m) # 130

```

- Závěr testování

Protože realizace testovací statistiky $s_E = 120$ nenáleží do kritického oboru, tj. $s_E \notin W$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

4. Testování intervalem spolehlivosti

- Interval spolehlivosti

Abychom mohli stanovit hranice 95% intervalu spolehlivosti, musíme nejprve vypočítat $\frac{m(m+1)}{2}$ Walshových průměrů $\frac{Z_i+Z_j}{2}$ a ty následně vzestupně seřadit. Walshovy průměry získáme pomocí funkce `owa` z knihovny `NSM3`. Prvním argumentem této funkce bude vektor nul o délce $m = 18$, druhým argumentem bude vektor `acetab.HRL0` o délce $m = 18$ obsahující ty rozdíly mezi výškami kloubní jamky na pravé a levé straně, pro než rozdíl $Z_i - 0$ nebyl rovný nule. Získané Walshovy průměry seřadíme příkazem `sort()`.

Prvním argumentem této funkce bude vektor nul o délce $m = 18$, druhým argumentem bude vektor rozdílů naměřených největších šířek kloubní jamky z pravé a levé strany `acetab.HRL`. Získané Walshovy průměry seřadíme příkazem `sort()`.

```

561 walsh <- NSM3::owa(rep(0, m), acetab.HRL0)$owa
562 (V <- sort(walsh))

```

[1]	-0.990	-0.970	-0.950	-0.775	-0.755	-0.710	-0.690	-0.690	-0.675	-0.670	-0.670	-0.655	563
[13]	-0.650	-0.560	-0.560	-0.540	-0.495	-0.475	-0.460	-0.455	-0.430	-0.410	-0.395	-0.390	564
[25]	-0.390	-0.375	-0.370	-0.360	-0.355	-0.350	-0.345	-0.320	-0.300	-0.280	-0.260	-0.245	565
[37]	-0.240	-0.215	-0.195	-0.160	-0.140	-0.135	-0.130	-0.115	-0.105	-0.045	-0.040	-0.025	566
[49]	-0.020	-0.005	0.000	0.000	0.015	0.035	0.055	0.065	0.080	0.085	0.100	0.105	567
[61]	0.110	0.120	0.140	0.145	0.155	0.160	0.165	0.170	0.180	0.185	0.215	0.230	568
[73]	0.235	0.255	0.270	0.270	0.275	0.295	0.295	0.315	0.330	0.335	0.350	0.370	569
[85]	0.385	0.390	0.395	0.415	0.445	0.455	0.510	0.535	0.555	0.560	0.575	0.585	570
[97]	0.610	0.615	0.625	0.640	0.650	0.670	0.670	0.675	0.685	0.685	0.690	0.695	571
[109]	0.695	0.710	0.715	0.720	0.730	0.770	0.785	0.790	0.800	0.810	0.825	0.835	572
[121]	0.845	0.855	0.870	0.870	0.875	0.900	0.960	0.985	1.020	1.040	1.065	1.145	573
[133]	1.170	1.200	1.225	1.225	1.225	1.250	1.315	1.330	1.340	1.375	1.385	1.400	574
[145]	1.410	1.500	1.560	1.730	1.755	1.780	1.880	1.900	1.915	1.940	2.095	2.100	575
[157]	2.160	2.180	2.195	2.200	2.310	2.550	2.655	2.710	2.735	2.825	2.885	3.240	576
[169]	3.265	3.425	4.750										577

Hranice intervalu spolehlivosti potom tvoří ty hodnoty, které se v seřazeném vektoru Walshových průměrů V nachází na $(s_m(\alpha/2))$ -té pozici a na $(s_m(1 - \alpha/2) + 1)$ -té pozici. Hodnoty $s_m(\alpha/2)$ a $s_m(1 - \alpha/2) + 1$ nalezneme opět pomocí funkce `qsignrank()`.

$$\begin{aligned}
(d, h) &= \left(V^{(s_m(\alpha/2))}; V^{(s_m(1-\alpha/2)+1)} \right) \\
&= \left(V^{(s_{18}(0.05/2))}; V^{(s_{18}(1-0.05/2)+1)} \right) \\
&= \left(V^{(5)}; Z^{(14)} \right) \\
&= (-0.14; 1.065)
\end{aligned}$$

```

578 qsignrank(alpha / 2, m) # 41
579 qsignrank(1 - alpha / 2, m) + 1 # 131
580 V[41] # -0.14
581 V[131] # 1.065

```

- **Závěr testování**

Protože $\tilde{z}_0 = 0$ náleží do 95% empirického oboustranného intervalu spolehlivosti, tj. $\tilde{z}_0 = 0 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

5. Testování p -hodnotou

Příslušnou p -hodnotu vypočítáme pomocí vzorce $2 \min\{\Pr(S_E \leq s_E), \Pr(S_E \geq s_E)\}$, kde $s_E = 120$ je realizace testovací statistiky S_E . Zároveň S_E je diskrétní náhodná veličina, proto $\Pr(S_E \geq 120) = 1 - \Pr(S_E < 120) = 1 - \Pr(S_E \leq 119)$. Viz kapitola ??.

- **p -hodnota**

$$\begin{aligned}
p\text{-hodnota} &= 2 \min\{\Pr(S_E \leq s_E), \Pr(S_E \geq s_E)\} \\
&= 2 \min\{\Pr(S_E \leq 120), \Pr(S_E \geq 120)\} \\
&= 2 \min\{\Pr(S_E \leq 120), 1 - \Pr(S_E \leq 119)\} \\
&= 2 \min\{0.9351311, 0.07075882\} \\
&= 2 \times 0.07075882 = 0.1415176 \doteq 0.1415
\end{aligned}$$

```

582 2 * min (psignrank(120, m), 1 - psignrank(120, m)) # 0.1415176

```

- **Závěr testování**

Protože p -hodnota = 0.1415 je větší než $\alpha = 0.05$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

6. Interpretace výsledků

Za základě všech tří typů testování nezamítáme nulovou hypotézu na hladině významnosti $\alpha = 0.05$. Mezi největší výškou kloubní jamky na pravé a levé straně u mužů starověké egyptské populace neexistuje statisticky významný rozdíl. Ke stejnému závěru jsme dospěli také v příkladu 9.7.

7. Grafická vizualizace výsledku testování

Grafické porovnání měření na pravé a levé straně bychom provedli analogicky jako v příkladu 9.7 pomocí jednoho ze dvou uvedených krabicových diagramů (viz obrázek 14).

Poznámka: Wilcoxonův párový test můžeme provést pomocí funkce `wilcox.test()`. Vstupními parametry budou vektor naměřených hodnot výšek kloubní jamky na pravé straně (`acetab.HR`), vektor naměřených hodnot výšek kloubní jamky na levé straně (`acetab.HL`), argument `paired = T` určující, že oba vektory považujeme za párová pozorování, požadavek na výpočet hranic intervalu spolehlivosti zadaný nastavením argumentu `conf.int = T`, hodnota hladiny významnosti α zadaná prostřednictvím koeficientu spolehlivosti $1 - \alpha$ nastavením hodnoty argumentu `conf.level =`

```
583 wilcox.test(acetab.HR, acetab.HL, paired = T, conf.int = T, conf.level = 0.95,  
584 alternative = 'two.sided', correct = F) # interpolovany IS i p-hodnota
```

```
Wilcoxon signed rank test  
  
data: acetab.HR and acetab.HL  
V = 120, p-value = 0.1329  
alternative hypothesis: true location shift is not equal to 0  
95 percent confidence interval:  
 -0.1400411  1.0650215  
sample estimates:  
(pseudo)median  
 0.390035
```

585
586
587
588
589
590
591
592
593
594
595

0.95, typ zvolené alternativní hypotézy (oboustranná) zadaný pomocí argumentu `alternative = 'two.sided'` a argument `correct = F`, který zakazuje aplikování spojité korekce na výsledné intervaly spolehlivosti a p -hodnotu.

Součástí výstupu je hodnota testovací statistiky $V = 120$, interpolované hranice 95% empirického oboustranného intervalu spolehlivosti -0.1400411 a 1.0650215 pro rozdíl naměřených výšek na pravé a levé straně, které jsou mírně přesnější, než námi stanovené hranice intervalu spolehlivosti (zpřesnění hranic bylo provedeno procesem nazývaným *interpolate*), a p -hodnota `p-value = 0.1329`. Jediné, co musíme stanovit zvlášť, jsou dolní a horní hranice kritického oboru. Všimněme si, že součástí výstupu je také hodnota pseudomediánu `(pseudo)median = 0.390035`. Tento pseudomedián je vypočítaný jako interpolovaná prostřední hodnota z posloupnosti seřazených Walshových průměrů. My však preferujeme hodnotu klasického mediánu (`median = 0.455`), kterou bychom získali příkazem `median()`, nebo příkazem `quantile()`, stejně, jako jsme si ukazovali v kapitole ??.

Druhá možnost provedení Wilcoxonova párového testu je opět použití funkce `wilcox.test()`, kde vstupními parametry budou vektor rozdílů naměřených hodnot na pravé a levé straně (`acetab.HRL`), argument `mu = 0` určující, že rozdílů porovnávané s konstantou $\tilde{z}_0 = 0$, požadavek na výpočet hranic intervalu spolehlivosti zadaný nastavením argumentu `conf.int = T`, hodnota hladiny významnosti α zadaná prostřednictvím koeficientu spolehlivosti $1 - \alpha$ nastavením hodnoty argumentu `conf.level = 0.95`, typ zvolené alternativní hypotézy (oboustranná) zadaný pomocí argumentu `alternative = 'two.sided'` a argument `correct = F`, který zakazuje aplikování spojité korekce na výsledné intervaly spolehlivosti a p -hodnotu. Argument `paired = T` tentokrát vynecháváme, neboť jde o zápis, kde jsme párový test porovnáující pravou a levou stranu převedli na jednovýřevový neparametrický test porovnáující rozdílů Z_i s konstantou $\tilde{z}_0 = 0$.

```
596 wilcox.test(acetab.HRL, mu = 0, conf.int = T, conf.level = 0.95,  
597 alternative = 'two.sided', correct = F) # interpolovany IS i p-hodnota
```

Výstup tohoto příkazu je totožný s výše uvedeným výstupem. Záleží tedy na nás, jakou syntaxi k zadání exaktního Wilcoxonova párového testu použijeme.

Poznámka: Výstup funkce `wilcox.test()` v obou výše uvedených případech provází dvě varovné hlášky. První hláška ve znění: `In wilcox.test.default(acetab.HR, acetab.HL, paired = T, conf.int = T, : cannot compute exact p-value with ties` nás upozorňuje na výskyt duplicitních rozdílů $|Z_2| = |Z_{10}| = 0.56$ a $|Z_4| = |Z_{17}| = 0.35$ a na použití modifikovaného postupu na výpočet p -hodnoty. Druhá hláška ve znění: `In wilcox.test.default(acetab.HR, acetab.HL, paired = T, conf.int = T, : cannot compute exact confidence interval with ties` nás opět upozorňuje na výskyt duplicitních rozdílů a na použití modifikovaného postupu na výpočet hranic intervalů spolehlivosti.

★

```
Wilcoxon signed rank test
data: acetab.HRL
V = 120, p-value = 0.1329
alternative hypothesis: true location is not equal to 0
95 percent confidence interval:
 -0.1400411  1.0650215
sample estimates:
(pseudo)median
 0.390035
```

598
599
600
601
602
603
604
605
606
607
608

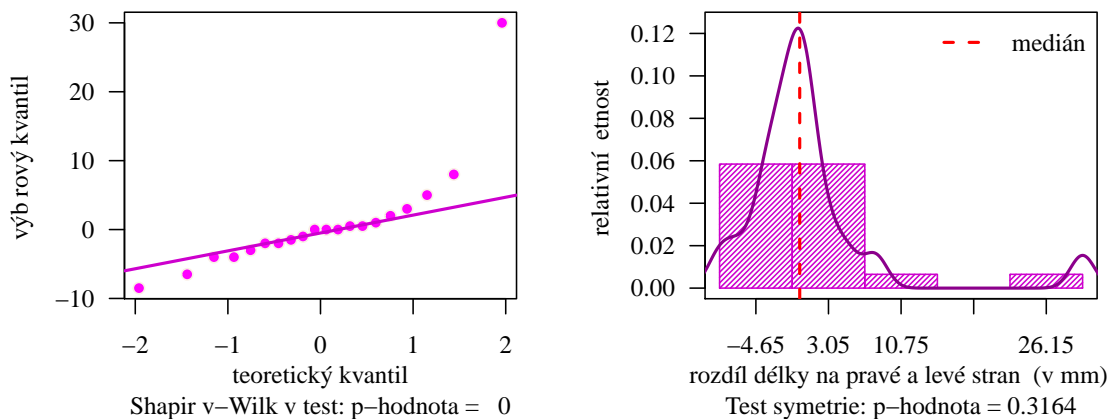
Příklad 9.20. Wilcoxonův párový exaktní test

Mějme datový soubor 21-goldman-measures.csv obsahující údaje o délkách lýtkových kostí v mm na pravé straně (tibia.LR) a na levé straně (tibia.LL) u skeletů obyvatel útesových obydlí (Cliff Dwellings) v Utahu (details viz sekce ??). Na hladině významnosti $\alpha = 0.05$ testujte nulovou hypotézu, že délka lýtkové kosti z pravé strany u skeletů mužského pohlaví je menší než délka lýtkové kosti z levé strany.

Řešení příkladu 9.20

Zadání příkladu je shodné se zadáním příkladu 9.8. Rozdíly naměřených délek lýtkových kostí z pravé a levé strany u 20 mužů z populace obyvatel útesových obydlí nepochází z normálního rozdělení (p -hodnota < 0.0001 , viz obrázek 32), proto na prozkoumání otázky ze zadání použijeme opět neparametrický párový test. Narozdíl od příkladu 9.8 nyní použijeme Wilcoxonův neparametrický párový test. Nejprve tedy testem symetrie ($\alpha = 0.05$) ověříme předpoklad symetrického rozdělení rozdílů okolo jejich mediánu. Graficky vizualizujeme míru symetrie rozdílů okolo mediánu pomocí histogramu superponovaného jádrovým odhadem křivky hustoty (viz obrázek 32).

```
609 data <- read.delim('00-Data//21-goldman-measures.csv', sep = ';')
610 # head(data)
611 data.TRLM <- data[data$pop == 'Cliff Dweller' & data$sex == 'm', c('tibia.LL', 'tibia.LR
  ')]
612 data.TRLM <- na.omit(data.TRLM)
613 tibia.LR <- data.TRLM$tibia.LR
614 tibia.LL <- data.TRLM$tibia.LL
615 tibia.LRL <- tibia.LR - tibia.LL
616 n <- length(tibia.LRL) # 20
```



Obrázek 32: Míra symetrie rozdílů délky lýtkové kosti u mužů na pravé a levé straně

Z QQ-diagramu i histogramu vidíme, že rozdíly z normálního rozdělení skutečně nepochází. Ačkoli na základě grafické vizualizace bychom se k předpokladu, že rozdíly jsou okolo svého mediánu rozmístěny symetricky, příliš nepřikláněli, test symetrie hypotézu o symetrii nezamítá (p -hodnota = 0.3164). Grafická vizualizace může být zkreslená nízkým počtem pozorování, proto se kloníme k závěru testu o symetrii a předpoklad o symetrii rozdílů okolo mediánu považujeme za splněný. Znění nulové a alternativní hypotézy je analogické jako v řešení příkladu 9.8, stejně jako hladina významnosti α .

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : Medián rozdílů délky pažní kosti na pravé a levé straně u mužů je menší nebo roven nule.

H_1 : Medián rozdílů délky pažní kosti na pravé a levé straně u mužů je větší než nula.

- **matematická formulace** nulové a alternativní hypotézy

$$H_0 : \tilde{z} \leq \tilde{z}_0, \text{ kde } \tilde{z}_0 = 0$$

$$H_1 : \tilde{z} > \tilde{z}_0, \text{ kde } \tilde{z}_0 = 0 \text{ (pravostranná alternativa)}$$

2. Volba hladiny významnosti

- Hladina významnosti $\alpha = 0.05$.

3. Testování kritickým oborem

- **Testovací statistika**

V úvodu příkladu jsme vytvořili vektor `tibia.LRL` ($Z_i, i = 1, \dots, 20$) jako vektor rozdílů mezi délkami lýtkových kostí na pravé a levé straně. Nyní je potřeba stanovit rozdíly $Z_i - \tilde{z}_0$. Protože ale opět $\tilde{z}_0 = 0$, jsou rozdíly $Z_i - \tilde{z}_0$ rovny původnímu vektoru Z_i . V dalším kroku se zaměříme na stanovení počtu nenulových rozdílů a zjistíme, že celkem 17 rozdílů $Z_i - 0$ je nenulových ($m = 17$). Dále vytvoříme vektor `tibia.LRLO` obsahující ty rozdíly mezi délkami látkových kostí na pravé a levé straně, pro než rozdíl $Z_i - 0$ nebyl rovný nule. Délka tohoto vektoru bude v tomto případě $m = 17$. V dalších krocích testování budeme používat právě vektor `tibia.LRLO`. Nyní označíme znaménkem $+$ kladné rozdíly $Z_i - 0 > 0$ a znaménkem $-$ záporné rozdíly $Z_i - 0 < 0$ nebo nulu $Z_i - 0 = 0$, vypočítáme absolutní hodnoty všech rozdílů, tj. $|Z_i - 0| = |Z_i|, i = 1, \dots, m$ a stanovíme pořadí R_i všech rozdílů v absolutních hodnotách (viz tabulka 15).

Tabulka 15: Naměřené hodnoty X_i, Y_i , rozdíly $Z_i = X_i - Y_i$, rozdíly $Z_i - 0$, znaménka těchto rozdílů I , absolutní hodnoty rozdílů $|Z_i - 0|$ a pořadí rozdílů v absolutních hodnotách

měření	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
X_i	376	370	346	388	380	335	372.5	388	364.0	361	370	396.5	373.0	378	368	370	374	356.0	355	337
Y_i	379	370	350	386	380	334	372.0	388	365.5	362	374	396.0	379.5	370	365	372	376	364.5	350	307
Z_i	-3	0	-4	2	0	1	0.5	0	-1.5	-1	-4	0.5	-6.5	8	3	-2	-2	-8.5	5	30
$Z_i - 0$	-3	0	-4	2	0	1	0.5	0	-1.5	-1	-4	0.5	-6.5	8	3	-2	-2	-8.5	5	30
I	-	X	-	+	X	+	+	X	-	-	-	+	-	+	+	-	-	-	+	+
$ Z_i - 0 $	3	X	4	2	X	1	0.5	X	1.5	1	4	0.5	6.5	8	3	2	2	8.5	5	30
R_i	9.5	X	11.5	7	X	3.5	1.5	X	5	3.5	11.5	1.5	14	15	9.5	7	7	16	13	17

Tabulku analogickou tabulce 15 nyní vytvoříme také pomocí softwaru \mathbb{R} . Z výše uvedeného textu víme, že vektory X_i (`tibia.LR`), Y_i (`tibia.LL`), Z_i (`tibia.LRL`) a $Z_i - 0$ mají délku $n = 20$. Nicméně vektory I , $|Z_i - 0|$ a R_i mají mít délku $m = 17$. Zároveň ale víme, že \mathbb{R} umí vytvořit tabulku pouze ze řádků, které mají stejnou délku. Situaci tedy vyřešíme tak, že vytvoříme vektor `tibia.LRLO`, který bude mít délku 20, nicméně pozice, kde $Z_i - 0 = 0$, nahradíme hodnotami `NA`. Indikační proměnná `I`, pro kterou platí $I = 1$ pokud $Z_i - 0 > 0$ a $I = 0$ jinak bude mít opět délku 20, ale tři pozice budou opět nabývat hodnoty `NA`. Stejná situace nastane u vektoru pořadí rozdílů v absolutních hodnotách `Ri`. Vektor pořadí stanovíme pomocí funkce `rank()` s volbou argumentu `na.last = "keep"`, která zajistí, aby `NA` hodnoty získali jako pořadí rovněž `NA` hodnotu. Pro větší přehlednost níže vypíšeme pouze prvních 15 sloupců z vytvořené tabulky.

```

617 z0 <- 0
618 n <- length(tibia.LRL) # 20
619 m <- sum(tibia.LRL - z0 != 0) # 17
620 tibia.LRLO <- tibia.LRL
621 tibia.LRLO[tibia.LRLO == 0] <- NA
622 I <- (tibia.LRLO - 0 > z0)
623 Ri <- rank(abs(tibia.LRLO - 0), na.last = "keep")
624 tab <- data.frame(rbind("Xi" = tibia.LR, "Yi" = tibia.LL, "Zi" = tibia.LRL,
625 "Zi-z0" = tibia.LRL - z0, "I" = I,
626 "|Zi-0|" = abs(tibia.LRLO - 0), "Ri" = Ri))
627 names(tab) <- 1:20
628 tab[, 1:15]

```

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
Xi	376.0	370	346.0	388	380	335.0	372.5	388	364.0	361.0	370.0	396.5	373.0	378	368.0	629
Yi	379.0	370	350.0	386	380	334.0	372.0	388	365.5	362.0	374.0	396.0	379.5	370	365.0	630
Zi	-3.0	0	-4.0	2	0	1.0	0.5	0	-1.5	-1.0	-4.0	0.5	-6.5	8	3.0	631
Zi-z0	-3.0	0	-4.0	2	0	1.0	0.5	0	-1.5	-1.0	-4.0	0.5	-6.5	8	3.0	632
I	0.0	NA	0.0	1	NA	1.0	1.0	NA	0.0	0.0	0.0	1.0	0.0	1	1.0	633
Zi-0	3.0	NA	4.0	2	NA	1.0	0.5	NA	1.5	1.0	4.0	0.5	6.5	8	3.0	634
Ri	9.5	NA	11.5	7	NA	3.5	1.5	NA	5.0	3.5	11.5	1.5	14.0	15	9.5	635
																636

Z tabulky 15 vidíme, že z celkového počtu sedmnácti nenulových rozdílů je celkem osm rozdílů kladných, přičemž pořadí těchto rozdílů v posloupnosti absolutních hodnot rozdílů jsou 7, 3.5, 1.5, 1.5, 15, 9.5, 13, 17. Hodnota testovací statistiky S_E , definovaná jako součet těchto pořadí, bude tedy rovná 68.

$$S_E = \sum_{i=1}^m I(Z_i - 0 > 0) = 7 + 3.5 + 1.5 + 1.5 + 15 + 9.5 + 13 + 17 = 68.$$

```
637 SE <- sum(Ri * I, na.rm = T) # 68
```

- **Kritický obor**

Dolní hranici kritického oboru vypočítáme pomocí kvantilové funkce `qsignrank()`.

$$\begin{aligned} W &= \langle s_m(1 - \alpha); \infty \rangle \\ &= \langle s_{17}(1 - 0.05); \infty \rangle \\ &= \langle s_{17}(0.95); \infty \rangle \\ &= \langle 111; \infty \rangle \end{aligned}$$

```
638 alpha <- 0.05
639 q <- qsignrank(1 - alpha, m) # 111
```

- **Závěr testování**

Protože realizace testovací statistiky $s_E = 92$ nenáleží do kritického oboru, tj. $s_E \notin W$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

4. Testování intervalem spolehlivosti

- **Interval spolehlivosti**

Proti pravostranné alternativní hypotéze postavíme levostranný interval spolehlivosti. Abychom mohli stanovit hranice 95% levostranného intervalu spolehlivosti, musíme nejprve vypočítat $\frac{m(m-1)}{2}$ Walshových průměrů, pomocí funkce `owa` z knihovny `NSM3`. Prvním argumentem této funkce bude vektor nul o délce $m = 17$, druhým argumentem bude vektor `tibia.LRL0`, ze kterého nejprve příkazem `na.omit()` odstraníme NA hodnoty. Délka vektoru bude tedy rovněž $m = 17$. Získané Walshovy průměry seřadíme příkazem `sort()`. Hranice intervalu spolehlivosti potom tvoří hodnota, která se v seřazeném vektoru $\frac{m(m+1)}{2}$ Walshových průměrů nachází na $(s_m(\alpha))$ -té pozici, a nekonečno. Hodnotu $s_m(\alpha)$ stanovíme pomocí funkce `qsignrank()`.

$$\begin{aligned}
 (d, h) &= \left(V^{(s_m(\alpha))}; \infty \right) \\
 &= \left(V^{(s_{17}(0.05))}; \infty \right) \\
 &= \left(V^{(42)}; \infty \right) \\
 &= (-2.25; \infty)
 \end{aligned}$$

```

640 walsh <- NSM3::owa(rep(0, m), na.omit(tibia.LRL0))$owa
641 V <- sort(walsh)
642 qsignrank(alpha, m) # 42
643 V[42] # -2.25

```

- **Závěr testování**

Protože $\tilde{z}_0 = 0$ náleží do 95% empirického jednostranného intervalu spolehlivosti, tj. $\tilde{z}_0 = 0 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

5. Testování p -hodnotou

Příslušnou p -hodnotu vypočítáme pomocí vzorce $\Pr(S_E \geq s_E)$, kde S_E je diskrétní náhodná veličina a s_E je realizace této náhodné veličiny, tj. $s_E = 68$.

- **p -hodnota**

$$\begin{aligned}
 p\text{-hodnota} &= \Pr(S_E \geq s_E) = \Pr(S_E \geq 68) \\
 &= 1 - \Pr(S_E < 68) = 1 - \Pr(S_E \leq 67) \\
 &= 0.6611481 \doteq 0.6611
 \end{aligned}$$

```

644 1 - psignrank(SE - 1, m) # 0.6611481

```

- **Závěr testování**

Protože p -hodnota = 0.6611 je větší než $\alpha = 0.05$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

6. Interpretace výsledků

Na základě všech tří typů testování na hladině významnosti $\alpha = 0.05$ nezamítáme nulovou hypotézu H_0 . Délka lýtkové kosti z pravé strany tedy není statisticky významně větší než délka lýtkové kosti z levé strany. Ke stejnému závěru jsme dospěli taktéž znaménkovým testem v rámci řešení příkladu 9.8.

7. Grafická vizualizace výsledku testování

Porovnání délky lýtkové kosti s pravé a levé strany bychom provedli pomocí krabicového diagramu analogicky, jako je uvedeno v příkladu 9.8 (viz obrázek 16).

Poznámka: Ukázkou provedení Wilcoxonova exaktního párového testu provedeme také pomocí funkce `wilcox.test()`. Vstupními parametry budou vektor naměřených hodnot délek lýtkové kosti na pravé straně (`tibia.LR`), vektor naměřených hodnot délek lýtkové kosti na levé straně (`tibia.LL`), specifikace párového testu (`paired = T`), požadavek na vypočítání intervalu spolehlivosti zadaný argumentem `conf.int = T`, hodnota hladiny významnosti α zadaná prostřednictvím koeficientu spolehlivosti $1 - \alpha$ (`conf.level = 0.95`), pravostranný typ alternativní hypotézy (`alternative = 'greater'`) a argument `correct = F` zakazující provedení spojité korekce na výsledných hranicích intervalu spolehlivosti.

```
645 wilcox.test(tibia.LR, tibia.LL, paired = T, conf.int = T, conf.level = 0.95,  
646             alternative = 'greater', correct = F) # IS interpolovany
```

```
Wilcoxon signed rank test  
  
data: tibia.LR and tibia.LL  
V = 68, p-value = 0.6565  
alternative hypothesis: true location shift is greater than 0  
95 percent confidence interval:  
 -2.250057      Inf  
sample estimates:  
(pseudo)median  
 -0.499972
```

647
648
649
650
651
652
653
654
655
656
657

Součástí výstupu je hodnota testovací statistiky $V = 68$, interpolované hranice 95% empirického jednostranného intervalu spolehlivosti -2.25 a Inf pro rozdíl mezi délkami na pravé a levé straně a p -hodnota $p\text{-value} = 0.6565$. Jediné, co musíme stanovit zvlášť, je dolní hranice kritického oboru.

Poznámka: Výstup funkce `wilcox.test()` v tomto případě provádí čtyři varovné hlášky. První dvě hlášky ve znění: 1: In `wilcox.test.default(tibia.LR, tibia.LL, paired = T, conf.int = T, : cannot compute exact p-value (confidence interval) with ties`. nás upozorňuje na výskyt duplicitních rozdílů $|Z_1| = |Z_{15}| = 3$, $|Z_3| = |Z_{11}| = 4$, $|Z_6| = |Z_{10}| = 1$ a $|Z_7| = |Z_{12}| = 0.5$ a na použití modifikovaného postupu na výpočet p -hodnoty (resp. hranic intervalů spolehlivosti). Třetí a čtvrtá hláška ve znění: In `wilcox.test.default(tibia.LR, tibia.LL, paired = T, conf.int = T, : cannot compute exact p-value (confidence interval) with zeroes`. nás upozorňuje na nulové rozdílly $|Z_2 - 0| = |Z_5 - 0| = |Z_8 - 0| = 0$ a na použití modifikovaného (neexaktního) postupu na výpočet p -hodnoty (resp. hranic intervalů spolehlivosti). Pro nás však tyto varovné hlášky nemají žádný závažný důsledek.

★

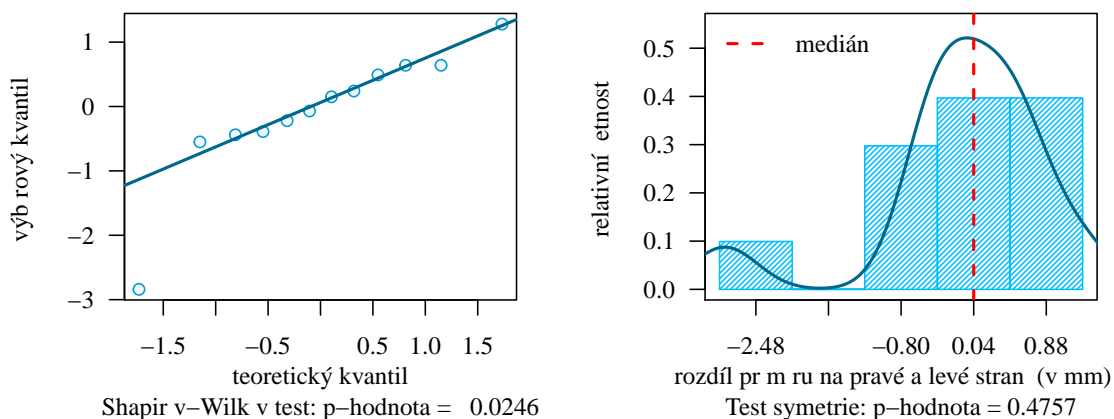
Příklad 9.21. Wilcoxonův párový exaktní test

Mějme datový soubor 21-goldman-measures.csv obsahující údaje o anteroposteriorním průměru hlavice stehenní kosti v mm na pravé straně (femur.HDR) a na levé straně (femur.HDL) u skeletů aljašské populace kmene Ipituaq (detaily viz sekce ??). Na hladině významnosti $\alpha = 0.10$ zjistěte, zda je u skeletů ženského pohlaví anteroposteriorní průměr hlavice stehenní kosti z levé strany menší než z pravé strany.

Řešení příkladu 9.21

Zadání příkladu je shodné se zadáním příkladu 9.9. Rozdíly naměřených anteriposteriorních průměrů hlavice stehenní kosti z levé a pravé strany u 12 ženských skeletů z aljašské populace kmene Ipituaq nepochází z normálního rozdělení (p -hodnota = 0.0246, viz obrázek 33). Na prozkoumání otázky ze zadání tedy použijeme Wilcoxonův neparametrický párový test. Nejprve ověříme nutný předpoklad symetrického rozdělení rozdílů okolo mediánu, a to jak testem symetrie, tak graficky (viz obrázek 33).

```
658 data <- read.delim('00-Data//21-goldman-measures.csv', sep = ';')
659 # head(data)
660 data.HDLR <- data[data$pop == 'Ipituaq' & data$sex == 'f', c('femur.HDL', 'femur.HDR')]
661 data.HDLR <- na.omit(data.HDLR)
662 femur.HDL <- data.HDLR$femur.HDL
663 femur.HDR <- data.HDLR$femur.HDR
664 femur.HDLR <- femur.HDL - femur.HDR
665 n <- length(femur.HDLR) # 12
```



Obrázek 33: Míra symetrie rozdílů délky těla anteroposteriorního průměru hlavice stehenní kosti u žen z aljašské populace kmene Ipituaq na levé a pravé straně

Na základě testu symetrie (p -hodnota = 0.4757) nezamítáme hypotézu o symetrickém rozdělení rozdílů okolo jejich mediánu. Ke stejnému závěru docházíme i po shlédnutí histogramu, ze kterého sice symetrie rozdílů není jednoznačně zjevná, nicméně nesmíme zapomenout na nízký rozsah náhodného výběru, kvůli kterému je test symetrie k mírnému porušení symetrie schovávavý. Předpoklad symetrie tedy považujeme za splněný, což znamená, že v použití Wilcoxonova testu nám nic nebrání. Nulová hypotéza a alternativní hypotéza i hladina významnosti α jsou analogické jako v příkladu 9.9.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : Medián rozdílů anteroposteriorního průměru hlavice stehenní kosti na levé a pravé straně u žen z kmene Ipituaq je větší nebo rovný nule.

H_1 : Medián rozdílů anteroposteriorního průměru hlavice stehenní kosti na levé a pravé straně u žen z kmene Ipituaq je menší než nula.

- **matematická formulace** nulové a alternativní hypotézy

$$H_0 : \tilde{z} \geq \tilde{z}_0, \text{ kde } \tilde{z}_0 = 0$$

$$H_1 : \tilde{z} < \tilde{z}_0, \text{ kde } \tilde{z}_0 = 0 \text{ (levostranná alternativa)}$$

2. Volba hladiny významnosti

- Hladina významnosti $\alpha = 0.10$.


3. Testování kritickým oborem

- **Testovací statistika**

V úvodu příkladu jsme vytvořili vektor femur.HDLR ($Z_i, i = 1, \dots, 26$) jako vektor rozdílů mezi anteroposteriorním průměrem hlavice stehenní kosti na levé a pravé straně. Nyní je potřeba stanovit rozdíly $Z_i - \tilde{z}_0$, které jsou stejné jako Z_i , neboť $\tilde{z}_0 = 0$, tj. $Z_i - \tilde{z}_0 = Z_i - 0 = Z_i$. V dalším kroku se zaměříme na stanovení počtu nenulových rozdílů a zjistíme, že žádný z rozdílů $Z_i - 0$ není nulový, tedy počet nenulových rozdílů ($m = n = 12$). Dále vytvoříme vektor femur.HDLR0 obsahující ty rozdíly mezi anteroposteriorními průměry na levé a pravé straně, pro něž rozdíl $Z_i - 0$ nebyl rovný nule. Délka tohoto vektoru bude v tomto případě $m = n = 12$. V dalších krocích testování budeme používat vektor femur.HDLR0. Nyní označíme znaménkem + kladné rozdíly $Z_i - 0 > 0$ a znaménkem - záporné rozdíly $Z_i - 0 < 0$ nebo nulu $Z_i - 0 = 0$, vypočítáme absolutní hodnoty všech rozdílů, tj. $|Z_i - 0| = |Z_i|$, $i = 1, \dots, m$ a stanovíme pořadí R_i všech rozdílů v absolutních hodnotách (viz tabulka 16).

Tabulka 16: Naměřené hodnoty X_i, Y_i , rozdíly $Z_i = X_i - Y_i$, rozdíly $Z_i - 0$, znaménka těchto rozdílů I , absolutní hodnoty rozdílů $|Z_i - 0|$ a pořadí rozdílů v absolutních hodnotách

měření	1	2	3	4	5	6	7	8	9	10	11	12
X_i	40.58	43.22	38.68	41.47	40.21	41.39	41.27	38.00	44.02	39.91	38.91	40.72
Y_i	41.13	41.94	38.04	41.86	40.43	40.75	41.03	37.85	44.09	42.75	39.35	40.23
Z_i	-0.55	1.28	0.64	-0.39	-0.22	0.64	0.24	0.15	-0.07	-2.84	-0.44	0.49
$Z_i - z_0$	-0.55	1.28	0.64	-0.39	-0.22	0.64	0.24	0.15	-0.07	-2.84	-0.44	0.49
I	-	+	+	-	-	+	+	+	-	-	-	+
$ Z_i - z_0 $	0.55	1.28	0.64	0.39	0.22	0.64	0.24	0.15	0.07	2.84	0.44	0.49
R_i	8.00	11.00	9.50	5.00	3.00	9.50	4.00	2.00	1.00	12.00	6.00	7.00

Tabulku analogickou tabulce 16 vytvoříme také pomocí softwaru .

```

666 z0 <- 0
667 n <- length(femur.HDLR)
668 m <- sum(femur.HDLR - z0 != 0)
669 femur.HDLR0 <- femur.HDLR
670 femur.HDLR0[femur.HDLR0 == 0] <- NA
671 I <- (femur.HDLR0 > z0)
672 Ri <- rank(abs(femur.HDLR0 - 0), na.last = "keep")
673 tab <- data.frame(rbind("Xi" = femur.HDL, "Yi" = femur.HDR,
674 "Zi" = femur.HDLR, "Zi-z0" = femur.HDLR - z0, "I" = I,
675 "|Zi-z0|" = abs(femur.HDLR0 - z0), "Ri" = Ri))
676 names(tab) <- 1 : n
677 SE <- sum(Ri * I) # 43

```

Z tabulky 16 vidíme, že ze všech 12 rozdílů $Z_i - 0$ je šest rozdílů kladných, přičemž pořadí těchto kladných rozdílů v posloupnosti absolutních hodnot rozdílů $|Z_i - 0|$ jsou 11, 9.5, 9.5, 4, 2 a 7. Hodnota testovací statistiky S_E bude tedy rovná 43.

$$S_E = \sum_{i=1}^n I(Z_i - \tilde{z}_0 > 0) = 11 + 9.5 + 9.5 + 4 + 2 + 7 = 43.$$

- **Kritický obor**

Horní hranici kritického oboru stanovíme pomocí kvantilové funkce `qsignrank()`.

$$\begin{aligned} W &= (-\infty; s_m(\alpha) - 1) \\ &= (-\infty; s_{12}(0.10) - 1) \\ &= (-\infty; 22 - 1) \\ &= (-\infty; 21) \end{aligned}$$

```
678 alpha <- 0.10
679 q <- qsignrank(alpha, m) - 1 # 21
```

- **Závěr testování**

Protože realizace testovací statistiky $s_E = 43$ nenáleží do kritického oboru, tj. $s_E \notin W$, H_0 nezamítáme na hladině významnosti $\alpha = 0.10$.

4. Testování intervalem spolehlivosti

- **Interval spolehlivosti**

Proti levostranné alternativní hypotéze postavíme pravostranný interval spolehlivosti. Abychom mohli stanovit hranice 90% pravostranného intervalu spolehlivosti, musíme nejprve vypočítat $\frac{m(m-1)}{2}$ Walshových průměrů, které následně seřadíme. Hranice intervalu spolehlivosti potom tvoří mínus nekonečno a hodnota, která se v seřazeném vektoru $\frac{m(m+1)}{2}$ Walshových průměrů nachází na $(s_m(1-\alpha)+1)$ -té pozici. Hodnotu $s_m(1-\alpha)$ stanovíme pomocí funkce `qsignrank()`.

Hranice intervalu spolehlivosti potom tvoří mínus nekonečno a hodnota, která se v seřazeném vektoru hodnot nachází na $(s_m(1-\alpha) + 1)$ -té pozici.

$$\begin{aligned} (d, h) &= \left(-\infty; V^{(s_m(1-\alpha)+1)}\right) \\ &= \left(-\infty; V^{(s_{12}(1-0.10)+1)}\right) \\ &= \left(-\infty; V^{(s_{12}(0.90)+1)}\right) \\ &= \left(-\infty; V^{(57)}\right) \\ &= \left(-\infty; 0.365\right) \end{aligned}$$

```
680 walsh <- NSM3::owa(rep(0, m), na.omit(femur.HDLR0))$owa
681 V <- sort(walsh)
682 qsignrank(1 - alpha, m) + 1 # 57
683 V[57] # 0.365
```

- **Závěr testování**

Protože $\tilde{z}_0 = 0$ náleží do 90% empirického pravostranného intervalu spolehlivosti, tj. $\tilde{x}_0 = 0 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0.10$.

5. Testování p -hodnotou

- **p -hodnota**

$$p\text{-hodnota} = \Pr(S_E \leq s_E) = \Pr(S_E \leq 43) = 0.6333008 \doteq 0.6333$$

- **Závěr testování**

Protože p -hodnota = 0.6333 je větší než $\alpha = 0.10$, H_0 nezamítáme na hladině významnosti $\alpha = 0.10$.

6. Interpretace výsledků

Na základě všech tří typů testování nezamítáme nulovou hypotézu na hladině významnosti $\alpha = 0.10$. Délka anteroposteriorního průměru hlavice stehenní kosti u žen z aljašské populace kmene Ipituaq na levé straně není statisticky významně menší než na pravé straně. Ke stejném závěru jsme dospěli také v řešení příkladu 9.9.

7. Grafická vizualizace výsledku testování

Rozdíl v anteroposteriorním průměru hlavice stehenní kosti na levé a pravé straně vizualizujeme pomocí krabicového diagramu, analogicky jako v příkladu 9.9 (viz obrázek 17).

Poznámka: Wilcoxonův párový test provedeme pomocí funkce `wilcox.test()`. Vstupními parametry budou vektor rozdílů naměřených hodnot na levé a pravé straně (`femur.HDLR`), argument `mu = 0` určující, že rozdíl porovnáme s konstantou $\tilde{z}_0 = 0$, požadavek na výpočet hranic intervalu spolehlivosti zadaný nastavením argumentu `conf.int = T`, hodnota hladiny významnosti α zadaná prostřednictvím koeficientu spolehlivosti $1 - \alpha$ nastavením hodnoty argumentu `conf.level = 0.90`, typ zvolené alternativní hypotézy (levostranná) zadaný pomocí argumentu `alternative = 'less'` a argument `correct = F`, který zakazuje aplikování spojité korekce na výsledné intervaly spolehlivosti a p -hodnotu.

```
685 wilcox.test(femur.HDLR, mu = 0, conf.int = T, conf.level = 0.90,
686             alternative = 'less', correct = F)
```

```

      Wilcoxon signed rank test
data:  femur.HDLR
V = 43, p-value = 0.6232
alternative hypothesis: true location is less than 0
90 percent confidence interval:
      -Inf 0.3199681
sample estimates:
(pseudo)median
      0.04502772
687
688
689
690
691
692
693
694
695
696
697
```

Součástí výstupu je hodnota testovací statistiky $V = 43$, interpolované hranice 90% empirického pravostranného intervalu spolehlivosti `-Inf` a `0.3199681` pro rozdíl mezi anteroposteriorními průměry na levé a pravé straně a p -hodnota `p-value = 0.6232`. Jediné, co musíme stanovit zvlášť, je dolní hranice kritického oboru. Hodnota pseudomediánu (`(pseudo)median = 0.04502772`) pro nás opět není zajímavá, narozdíl od hodnoty klasického mediánu (`median = 0.04`), kterou dopočítáme příkazem `median()`.



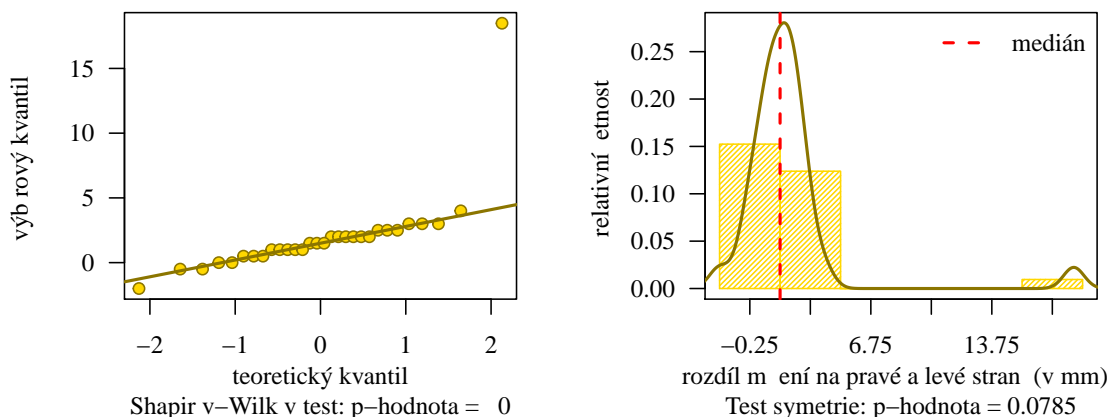
Příklad 9.22. Wilcoxonův párový asymptotický test

Máme datový soubor 21-goldman-measures.csv obsahující údaje o epikondylární šířce pažní kosti na pravé straně (humer.EBR) a na levé straně (humer.EBL) u mužů (sex = 'm' a sex == 'm?') a žen (sex = 'f' a sex == 'f?') pohřbených v oblasti Indian Knoll v Kentucky (viz sekce ??). Na hladině významnosti $\alpha = 0.01$ zjistíte, zda se střední hodnota epikondylární šířky pažní kosti u mužů na pravé straně a levé straně liší. K testování využijete všechna kompletní měření týkající se mužů (skupiny muž (sex == 'm') i pravděpodobně muž (sex == 'm?')).

Řešení příkladu 9.22

Zadání příkladu je shodné se zadáním příkladu 9.22. Rozdíly mezi hodnotami epikondylární šířky pažní kosti naměřenými na pravé a levé straně u 30 mužů z oblasti Indian Knoll v Kentucky nepocházejí z normálního rozdělení (p -hodnota < 0.0001 , viz obrázek 34). Na prozkoumání otázky ze zadání tentokrát použijeme asymptotickou variantu Wilcoxonova neparametrického párového testu. Nejprve ověříme nutný předpoklad symetrického rozdělení rozdílů okolo mediánu, a to jak testem symetrie, tak graficky (viz obrázek 34).

```
698 data <- read.delim('00-Data//21-goldman-measures.csv', sep = ';')
699 #head(data)
700 data.m <- data[data$sex %in% c('m', 'm?') & data$pop == 'Indian Knoll',
701               c('humer.EBL', 'humer.EBR')]
702 data.m <- na.omit(data.m)
703 humer.EBR <- data.m$humer.EBR
704 humer.EBL <- data.m$humer.EBL
705 humer.EBRL <- humer.EBR - humer.EBL
706 n <- length(humer.EBRL) # 30
```



Obrázek 34: Histogram a QQ-diagram rozdílů epikondylární šířky pažní kosti na pravé a levé straně

Na základě testu symetrie nezamítáme hypotézu o symetrickém rozdělení rozdílů hodnot naměřených na pravé a levé straně. Z histogramu na obrázku 34 vidíme, že odlehle pozorování má na symetrický charakter rozdílů výrazný vliv, podle testu však tento vliv není fatální. Předpoklad symetrie je tedy splněn, a proto můžeme asymptotickou variantu Wilcoxonova neparametrického testu použít. Otázka ze zadání je, analogicky jako v řešení příkladu 9.10, zněním alternativní hypotézy. Taktéž nulová hypotéza a hladina významnosti α jsou shodné jako v příkladu 9.10.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : Medián rozdílů epikondylární šířky u mužů o oblasti Indian Knoll na pravé a levé straně je rovný nule.

H_1 : Medián rozdílů epikondylární šířky u mužů o oblasti Indian Knoll na pravé a levé straně není rovný nule.

- **matematická formulace** nulové a alternativní hypotézy

$$H_0 : \tilde{z} = \tilde{z}_0, \text{ kde } \tilde{z}_0 = 0$$

$$H_1 : \tilde{z} \neq \tilde{z}_0, \text{ kde } \tilde{z}_0 = 0 \text{ (oboustranná alternativa)}$$

2. Volba hladiny významnosti

- Hladina významnosti $\alpha = 0.01$.

3. Testování kritickým oborem

- **Testovací statistika**

V úvodu příkladu jsme vytvořili vektor `humer.EBRL` ($Z_i, i = 1, \dots, 40$) jako vektor rozdílů hodnot epikondalárních šířek pažní kosti naměřených na pravé a levé straně. Dále jsme stanovili rozdíl $Z_i - \tilde{z}_0$. Protože ale $\tilde{z}_0 = 0$, jsou rozdíl $Z_i - \tilde{z}_0$ rovny původnímu vektoru rozdílů Z_i , tj. $Z_i - \tilde{z}_0 = Z_i - 0 = Z_i$. V dalším kroku se zaměříme na stanovení počtu nenulových rozdílů a zjistíme, že dva rozdíl $Z_i - 0$ jsou nulové (počet nenulových rozdílů $m = 28$). Nyní vytvoříme vektor `humer.EBRL0` obsahující ty rozdíl mezi pravou a levou stranou, pro něž rozdíl $Z_i - \tilde{z}_0$ nebyl rovný nule. Délka tohoto vektoru bude v tomto případě $m = 28$. V dalších krocích testování budeme používat právě vektor `humer.EBRL0`. Dále si označíme znaménkem $+$ rozdíl $Z_i - 0$, které jsou kladné a znaménkem $-$ rozdíl $Z_i - 0$, které jsou záporné, nebo 0. Nakonec vypočítáme absolutní hodnoty rozdílů $|Z_i - 0| = |Z_i|, i = 1, \dots, m$ a stanovíme pořadí těchto rozdílů v absolutních hodnotách.

```

707 z0 <- 0
708 n <- length(humer.EBRL)
709 m <- sum(humer.EBRL - z0 != 0)
710 humer.EBRL0 <- humer.EBRL
711 humer.EBRL0[humer.EBRL0 == 0] <- NA
712
713 I <- (humer.EBRL0 > z0)
714 Ri <- rank(abs(humer.EBRL0 - 0), na.las = "keep")
715 tab <- data.frame(rbind("Xi" = humer.EBR, "Yi" = humer.EBL,
716                       "Zi" = humer.EBRL, "Zi-z0" = humer.EBRL - z0, "I" = I,
717                       "Zi-z0" = abs(humer.EBRL0 - z0), "Ri" = Ri))
718 names(tab) <- 1 : n
719 SE <- sum(Ri * I, na.rm = T) # 383

```

Celkem 28 rozdílů $Z_i - 0, i = 1, \dots, 28$ je kladných. Pořadí těchto rozdílů v absolutních hodnotách jsou 12, 17, 17, 27, 22, 25, 3, 22, 8, 8, 8, 17, 12, 17, 3, 25, 8, 17, 25, 17, 3, 22, 12, 28, 8. Hodnota testovací statistiky S_E , která je definovaná jako součet těchto pořadí, bude tedy rovná 383.

$$S_E = \sum_{i=1}^m I(Z_i - 0 > 0) = 12 + 17 + 17 + \dots + 12 + 28 + 8 = 383.$$

Nyní dopočítáme hodnotu testovací statistiky S_A .

$$\begin{aligned}
S_A &= \frac{S_E - \frac{m(m+1)}{4}}{\sqrt{\frac{m(m+1)(2m+1)}{24}}} = \frac{383 - \frac{28(28+1)}{4}}{\sqrt{\frac{28(28+1)(2 \times 28+1)}{24}}} = \frac{383 - \frac{28 \times 29}{4}}{\sqrt{\frac{28 \times 29 \times 57}{24}}} = \frac{383 - 203}{\sqrt{1928.5}} = \frac{180}{43.91469} \\
&= 4.098856 \doteq 4.0989
\end{aligned}$$

```
720 SE <- sum(Ri * I, na.rm = T) # 383
721 SA <- (SE - m * (m + 1) / 4) / sqrt(m * (m + 1) * (2 * m + 1) / 24) # 4.098856
```

- **Kritický obor**

Protože $S_A \sim N(0, 1)$, horní a dolní hranici kritického oboru nalezneme pomocí funkce `qnorm()` jako $\alpha/2$ kvantil a $1 - \alpha/2$ kvantil standardizovaného normálního rozdělení $N(0, 1)$.

$$\begin{aligned} W &= (-\infty; u_{\alpha/2}) \cup (u_{1-\alpha/2}; \infty) \\ &= (-\infty; u_{0.01/2}) \cup (u_{1-0.01/2}; \infty) \\ &= (-\infty; u_{0.005}) \cup (u_{0.995}; \infty) \\ &= (-\infty; -2.575829) \cup (2.575829; \infty) \\ &= (-\infty; -2.5758) \cup (2.5758; \infty) \end{aligned}$$

```
722 alpha <- 0.01
723 qnorm(alpha / 2) # -2.575829
724 qnorm(1 - alpha / 2) # 2.575829
```

- **Závěr testování**

Protože realizace testovací statistiky $s_A = 4.0989$ náleží do kritického oboru, tj. $s_A \in W$, H_0 zamítáme na hladině významnosti $\alpha = 0.01$.

4. Testování intervalem spolehlivosti

- **Interval spolehlivosti**

Abychom mohli stanovit hranice 99% intervalu spolehlivosti, musíme nejprve vypočítat $\frac{m(m+1)}{2}$ Walshových průměrů $\frac{Z_i + Z_j}{2}$, $i = 1, \dots, m$, $j = 1, \dots, m$, $j \leq i$ a ty následně vzestupně seřadit. Walshovy průměry získáme pomocí funkce `owa` z knihovny `NSM3`. Prvním argumentem této funkce bude vektor nul o délce $m = 28$, druhým argumentem bude vektor `humer.EBRL0`, obsahující ty rozdíly mezi pravou a levou stranou, pro něž rozdíl $Z_i - 0$ nebyl rovný nule. Z tohoto vektoru však nejprve příkazem `na.omit()` odstraníme NA hodnoty, aby délka vektoru byla taktéž $m = 28$. Získané Walshovy průměry seřadíme příkazem `sort()`. Následně vypočítáme pozicové souřadnice $C_{1-\alpha/2}$ a $C_{\alpha/2}$.

$$\begin{aligned} C_{1-\alpha/2} &= \frac{m(m+1)}{4} - u_{1-\alpha/2} \sqrt{\frac{m(m+1)(2m+1)}{24}} \\ &= \frac{28(28+1)}{4} - u_{1-0.01/2} \sqrt{\frac{28(28+1)(2 \times 28+1)}{24}} \\ &= \frac{28 \times 29}{4} - u_{0.995} \sqrt{\frac{28 \times 29 \times 57}{24}} \\ &= 203 - 2.575829 \sqrt{1928.5} \\ &= 203 - 2.575829 \times 43.91469 = 89.88327 \doteq 90 \end{aligned}$$

$$\begin{aligned}
C_{\alpha/2} &= \frac{m(m+1)}{4} - u_{\alpha/2} \sqrt{\frac{m(m+1)(2m+1)}{24}} \\
&= \frac{28(28+1)}{4} - u_{0.01/2} \sqrt{\frac{28(28+1)(2 \times 28+1)}{24}} \\
&= \frac{28 \times 29}{4} - u_{0.005} \sqrt{\frac{28 \times 29 \times 57}{24}} \\
&= 203 + 2.575829 \sqrt{1928.5} \\
&= 203 + 2.575829 \times 43.91469 = 316.1167 \doteq 316
\end{aligned}$$

Hranice intervalu spolehlivosti potom tvoří ty hodnoty, které se v seřazeném vektoru Walshových průměrů V nachází na 90. pozici a na 316. pozici.

$$\begin{aligned}
(d, h) &= \left(V^{(C_{1-\alpha/2})}; V^{(C_{\alpha/2})} \right) \\
&= \left(V^{(90)}; V^{(316)} \right) \\
&= (1; 2.25)
\end{aligned}$$

```

725 walsh <- NSM3::owa(rep(0, m), na.omit(humer.EBRL0))$owa
726 V <- sort(walsh)
727 C1 <- m * (m + 1) / 4 - qnorm(1 - alpha / 2) * sqrt(m * (m + 1) * (2 * m + 1) / 24) #
      89.88325
728 C2 <- m * (m + 1) / 4 - qnorm(alpha / 2) * sqrt(m * (m + 1) * (2 * m + 1) / 24) #
      316.1167
729 V[90] # 1
730 V[316] # 2.25

```

- **Závěr testování**

Protože $\tilde{x}_0 = 0$ nenáleží do 99% empirického oboustranného intervalu spolehlivosti, tj. $\tilde{x}_0 = 0 \notin IS$, H_0 zamítáme na hladině významnosti $\alpha = 0.01$.

5. **Testování p -hodnotou** Příslušou p -hodnotu vypočítáme pomocí vzorce $2 \min\{\Pr(S_A \leq s_A), \Pr(S_A > s_A)\}$, kde S_A je spojitá náhodná veličina pocházející ze standardizovaného normálního rozdělení, tj. $S_A \sim N(0, 1)$ a $s_A = 4.0989$. $\Pr(S_A \leq s_A)$ je tedy distribuční funkce standardizovaného normálního rozdělení $N(0, 1)$ v hodnotě 4.0989. Její hodnotu spočítáme příkazem `pnorm()`.

- **p -hodnota**

$$\begin{aligned}
p\text{-hodnota} &= 2 \min\{\Pr(S_A \leq s_A), \Pr(S_A \geq s_A)\} \\
&= 2 \min\{\Pr(S_A \leq 4.0989), \Pr(S_A \geq 4.0989)\} \\
&= 2 \min\{\Pr(S_A \leq 4.0989), 1 - \Pr(S_A \leq 4.0989)\} \\
&= 2 \min\{0.9999792, 2.075984 \times 10^{-5}\} \\
&= 2 \times 2.075984 \times 10^{-5} = 4.151969 \times 10^{-5} \doteq 4.1520 \times 10^{-5}
\end{aligned}$$

```

731 p.hodnota <- 2 * min(pnorm(SA), 1 - pnorm(SA)) # 4.151969e-05

```

- **Závěr testování**

Protože p -hodnota 4.1520×10^{-5} je menší než $\alpha = 0.01$, H_0 zamítáme na hladině významnosti $\alpha = 0.01$.

6. Interpretace výsledků

Za základě všech tří typů testování zamítáme nulovou hypotézu na hladině významnosti $\alpha = 0.01$. Mezi epikondylární šířkou pažní kosti na pravé a levé straně u mužů z oblasti Indian Knoll v Kentucky existuje statisticky významný rozdíl.

7. Grafická vizualizace výsledku testování

Porovnání epikondylárních šířek naměřených na pravé a levé straně můžeme zobrazit pomocí krabicového diagramu analogicky jako v příkladu 9.10 (viz obrázek 20).

Poznámka: Asymptotickou variantu Wilcoxonova párového testu můžeme provést pomocí funkce `wilcox.test()`. Vstupními parametry budou vektor epikondylárních šířek naměřených na pravé straně (`humer.EBR`), vektor epikondylárních šířek naměřených na levé straně (`humer.EBL`), argument `paired = T` určující, že oba vektory považujeme za párová pozorování, argument povolující výpočet hranic intervalu spolehlivosti `conf.int = T`, hodnota hladiny významnosti α zadaná prostřednictvím koeficientu spolehlivosti $1 - \alpha$ nastavením hodnoty argumentu `conf.level = 0.99`, typ zvolené alternativní hypotézy (oboustranná) zadaný argumentem `alternative = 'two.sided'` a argument `correct = T`, který zvolí k výpočtu právě asymptotickou variantu Wilcoxonova neparametrického testu.

```
732 wilcox.test(humer.EBR, humer.EBL, paired = T, conf.int = T, conf.level = 0.99,  
733           alternative = 'two.sided', correct = T) # interpolovany IS i p-hodnota
```

```
Wilcoxon signed rank test with continuity correction  
  
data: humer.EBR and humer.EBL  
V = 383, p-value = 4.099e-05  
alternative hypothesis: true location shift is not equal to 0  
99 percent confidence interval:  
 0.9999622 2.2500614  
sample estimates:  
(pseudo)median  
 1.749982
```

734
735
736
737
738
739
740
741
742
743
744

Součástí výstupu je hodnota testovací statistiky $S_E V = 383$, interpolované hranice 99% asymptotického empirického oboustranného intervalu spolehlivosti 0.9999622 a 2.2500614 pro rozdíl hodnot naměřených při prvním a druhém měření (zpřesnění hranic bylo provedeno procesem nazývaným *interpolace*), a p -hodnota `p-value = 4.099e-05`. Jediné, co musíme stanovit zvlášť, jsou dolní a horní hranice kritického oboru. Všimněme si, že součástí výstupu je také hodnota pseudomediánu (`pseudo)median = 1.749982` rozdílů hodnot naměřených na pravé a levé straně. Tento pseudomedián je vypočítaný jako interpolovaná prostřední hodnota z posloupnosti seřazených Walshových průměrů. My však preferujeme hodnotu klasického mediánu rozdílů (`median = 1.5`), kterou bychom získali příkazem `median()`, nebo příkazem `quantile()`, analogicky, jako jsme si ukazovali v kapitole ??.

Druhá možnost provedení Wilcoxonova párového testu je opět použití funkce `wilcox.test()`, kde vstupními parametry budou vektor rozdílů epikondylárních šířek naměřených na pravé a levé straně (`humer.EBRL`), argument `mu = 0` určující, že rozdílů porovnávané s konstantou $\tilde{z}_0 = 0$, požadavek na výpočet hranic intervalu spolehlivosti (`conf.int = T`), hodnota hladiny významnosti α (`conf.level = 0.99`), typ zvolené alternativní hypotézy (`alternative = 'two.sided'`) a argument `correct = T`, který umožňuje použití právě asymptotické varianty Wilcoxonova párového testu. Argument `paired = T` tentokrát vynecháváme, neboť jde o zápis, kde jsme párový test porovnáující hodnoty neměřené na prvním a druhém měření převedli na jednovývěrový neparametrický test porovnáující rozdílů Z_i s konstantou $\tilde{z}_0 = 0$.

Výstup tohoto příkazu je totožný s výše uvedeným výstupem. Záleží tedy na nás, jakou syntaxi k zadání asymptotického Wilcoxonova párového testu použijeme.

```
745 wilcox.test(humer.EBRL, mu = 0, conf.int = T, conf.level = 0.99,  
746 alternative = 'two.sided', correct = T) # interpolovany IS i p-hodnota
```

```
Wilcoxon signed rank test with continuity correction  
  
data: humer.EBRL  
V = 383, p-value = 4.099e-05  
alternative hypothesis: true location is not equal to 0  
99 percent confidence interval:  
 0.9999622 2.2500614  
sample estimates:  
(pseudo)median  
 1.749982
```

747
748
749
750
751
752
753
754
755
756
757

Poznámka: Výstup funkce `wilcox.test()` v tomto případě provází čtyři varovné hlášky. První dvě hlášky ve znění: 1: In `wilcox.test.default(humer.EBR, humer.EBL, paired = T, conf.int = T, : cannot compute exact p-value (confidence interval) with ties`. nás upozorňuje na výskyt duplicitních rozdílů $|Z_1| = |Z_{17}| = |Z_{27}| = 1.5$, $|Z_2| = |Z_3| = |Z_4| = |Z_{16}| = |Z_{18}| = |Z_{22}| = |Z_{24}| = 2$, $|Z_7| = |Z_{12}| = |Z_{26}| = 2.5$, $|Z_8| = |Z_{20}| = |Z_{23}| = 3$, $|Z_9| = |Z_{10}| = |Z_{19}| = |Z_{25}| = |Z_{28}| = 0.5$ a $|Z_{13}| = |Z_{14}| = |Z_{15}| = |Z_{21}| = |Z_{30}| = 1$ a na použití modifikovaného postupu na výpočet p -hodnoty (resp. hranic intervalů spolehlivosti). Třetí a čtvrtá hláška ve znění: In `wilcox.test.default(humer.EBR, humer.EBL, paired = T, conf.int = T, : cannot compute exact p-value (confidence interval) with zeroes`. nás upozorňuje na výskyt nulových rozdílů $|Z_6 - 0| = |Z_{11}| = 0$ a na použití modifikovaného postupu na výpočet p -hodnoty (resp. hranic intervalů spolehlivosti). Pro nás však tyto varovné hlášky nemají žádný závažný důsledek.

★

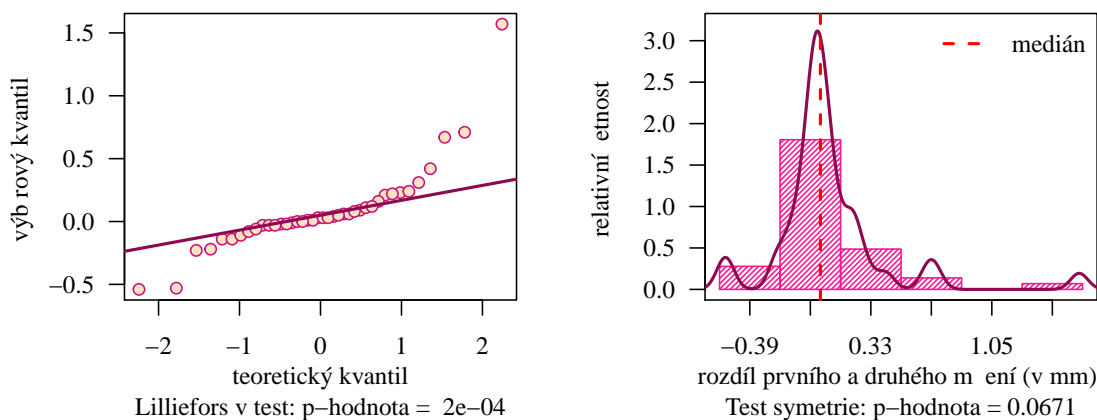
Příklad 9.23. Wilcoxonův párový asymptotický test

Máme datový soubor 02-paired-means-clavicle.txt obsahující údaje o hodnotách vertikálního průměru středu délky těla klíční kosti z pravé a levé strany (*clavicula*) z pohřebiště u Sv. Jakuba v Brně, převážně z období středověku, naměřené jedním výzkumníkem ve dvou opakovaných měřeních (viz sekce ??). Hodnoty naměřené při prvním opakování jsou uloženy v proměnné `simd.1`, hodnoty naměřené při druhém opakování jsou uloženy v proměnné `simd.2`. Můžeme zjistit, že aritmetický průměr hodnot získaných v rámci prvního měření je větší než aritmetický průměr hodnot získaných v rámci druhého měření. Na hladině významnosti $\alpha = 0.05$ zjistíte, zda je střední hodnota prvního měření větší než střední hodnota druhého měření vertikálního průměru délky těla klíční kosti na levé straně provedené tímto výzkumníkem.

Řešení příkladu 9.23

Zadání příkladu je shodné se zadáním příkladu 9.11. Rozdíly ve vertikálních průměrech středu délky těla 39 klíčnicích kostí z levé strany naměřených při prvním a druhém měření týmž výzkumníkem nepochází z normálního rozdělení (p -hodnota = 0.0002, viz obrázek 35). K testování tedy opět použijeme asymptotickou variantu Wilcoxonova neparametrického párového testu. Nejprve ale ověříme předpoklad symetrického rozdělení rozdílů okolo mediánu testem symetrie.

```
758 data <- read.delim('00-Data//02-paired-means-clavicle.txt')
759 # head(data)
760 data.12L <- data[data$side == 'L', c('simd.1', 'simd.2')]
761 data.12L <- na.omit(data.12L)
762 simd.1L <- data.12L$simd.1
763 simd.2L <- data.12L$simd.2
764 simd.12L <- simd.1L - simd.2L
765 n <- length(simd.12L) # 40
```



Obrázek 35: Míra symetrie rozdílů prvního a druhého měření vertikálního průměru ve středu délky těla klíční kosti na levé straně získaných jedním výzkumníkem po odstranění odlehlé hodnoty

Protože p -hodnota testu symetrie vyšla 0.0671, hypotézu o symetrickém rozdělení rozdílů okolo mediánu nezamítáme na hladině významnosti $\alpha = 0.05$, a tedy předpoklad o symetrickém rozdělení rozdílů okolo mediánu je splněn (grafické vizualizace viz obrázek ??). Wilcoxonův neparametrický párový test tedy můžeme použít. Jelikož je zadání příkladu shodné se zadáním příkladu 9.11, bude i znění nulové hypotézy, alternativní hypotézy i hodnota hladiny významnosti α stejná jako v příkladu 9.11.

1. Stanovení hypotéz

- slovní formulace nulové a alternativní hypotézy

H_0 : Medián rozdílů prvního a druhého měření vertikálního průměru těla klíční kosti na levé straně

měřených jedním výzkumníkem je menší než nula.

H_1 : Medián rozdílů prvního a druhého měření vertikálního průměru těla klíční kosti na levé straně měřených jedním výzkumníkem je větší než nula.

- **matematická formulace** nulové a alternativní hypotézy

$H_0 : \tilde{z} \leq \tilde{z}_0$, kde $\tilde{z}_0 = 0$

$H_1 : \tilde{z} > \tilde{z}_0$, kde $\tilde{z}_0 = 0$ (pravostranná alternativa)

2. Volba hladiny významnosti

- Hladina významnosti $\alpha = 0.05$.

3. Testování kritickým oborem

- **Testovací statistika**

V úvodu příkladu jsme vytvořili vektor `simd.12L` ($Z_i, i = 1, \dots, 40$) jako vektor rozdílů prvního a druhého měření vertikálního průměru těla klíční kosti na levé straně měřených jedním výzkumníkem. Dále jsme stanovili rozdílů $Z_i - \tilde{z}_0$. Protože ale $\tilde{z}_0 = 0$, jsou rozdílů $Z_i - \tilde{z}_0$ rovny původnímu vektoru rozdílů Z_i , tj. $Z_i - \tilde{z}_0 = Z_i - 0 = Z_i$. V dalším kroku se zaměříme na stanovení počtu nenulových rozdílů a zjistíme, že dva rozdílů $Z_i - 0$ jsou nulové (počet nenulových rozdílů $m = 38$). Nyní vytvoříme vektor `simd.12L0` obsahující ty rozdílů mezi prvním a druhým měřením, pro něž rozdíl $Z_i - \tilde{z}_0$ nebyl rovný nule. Délka tohoto vektoru bude v tomto případě $m = 38$. V dalších krocích testování budeme používat právě vektor `simd.12L0`. Dále si označíme znaménkem $+$ rozdílů $Z_i - 0$, které jsou kladné a znaménkem $-$ rozdílů $Z_i - 0$, které jsou záporné, nebo 0. Nakonec vypočítáme absolutní hodnoty rozdílů $|Z_i - 0| = |Z_i|, i = 1, \dots, m$ a stanovíme pořadí těchto rozdílů v absolutních hodnotách.

```
766 z0 <- 0
767 n <- length(simd.12L)
768 m <- sum(simd.12L - z0 != 0)
769 simd.12L0 <- simd.12L
770 simd.12L0[ simd.12L0 == 0 ] <- NA
771
772 I <- (simd.12L0 > z0)
773 Ri <- rank(abs(simd.12L0 - 0), na.las = "keep")
774 tab <- data.frame(rbind("Xi" = simd.1L, "Yi" = simd.2L,
775                        "Zi" = simd.12L, "Zi-z0" = simd.12L - z0, "I" = I,
776                        "Zi-z0" = abs(simd.12L0 - z0), "Ri" = Ri))
777 names(tab) <- 1 : n
```

Celkem 23 rozdílů $Z_i - 0, i = 1, \dots, 38$ je kladných. Pořadí těchto rozdílů v absolutních hodnotách jsou 32, 37, 15, 36, 26, 7.5, 30, 12, 15, 1.5, 33, 28, 3, 19, 25, 7.5, 20, 31, 18, 10.5, 38, 13 a 22. Hodnota testovací statistiky S_E , která je definovaná jako součet těchto pořadí, bude tedy rovná 480.

$$S_E = \sum_{i=1}^m I(Z_i - 0 > 0) = 32 + 37 + 15 + \dots + 38 + 13 + 22 = 480.$$

Nyní dopočítáme hodnotu testovací statistiky S_A .

$$\begin{aligned} S_A &= \frac{S_E - \frac{m(m+1)}{4}}{\sqrt{\frac{m(m+1)(2m+1)}{24}}} = \frac{480 - \frac{38(38+1)}{4}}{\sqrt{\frac{38(38+1)(2 \times 38+1)}{24}}} = \frac{480 - \frac{38 \times 39}{4}}{\sqrt{\frac{38 \times 39 \times 77}{24}}} = \frac{480 - 370.5}{\sqrt{4754.75}} = \frac{109.5}{68.9547} \\ &= 1.587999 \doteq 1.5880 \end{aligned}$$

```
778 SE <- sum(Ri * I, na.rm = T) # 480
779 SA <- (SE - m * (m + 1) / 4) / sqrt(m * (m + 1) * (2 * m + 1) / 24) # 1.587999
```

- **Kritický obor**

Protože $S_A \sim N(0, 1)$, horní a dolní hranici kritického oboru nalezneme pomocí funkce `qnorm()` jako $\alpha/2$ kvantil a $1 - \alpha/2$ kvantil standardizovaného normálního rozdělení $N(0, 1)$.

$$\begin{aligned} W &= \langle u_{1-\alpha}; \infty \rangle \\ &= \langle u_{1-0.05}; \infty \rangle \\ &= \langle u_{0.95}; \infty \rangle \\ &= \langle 1.644854; \infty \rangle \\ &= \langle 1.6449; \infty \rangle \end{aligned}$$

```
780 alpha <- 0.05
781 qnorm(1 - alpha) # 1.644854
```

- **Závěr testování**

Protože realizace testovací statistiky $s_A = 1.5880$ nenáleží do kritického oboru, tj. $s_A \notin W$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

4. Testování intervalem spolehlivosti

- **Interval spolehlivosti**

Proti pravostranné alternativní hypotéze postavíme levostranný interval spolehlivosti. Abychom mohli stanovit hranice 95% levostranného intervalu spolehlivosti, musíme nejprve vypočítat $\frac{m(m+1)}{2}$ Walshových průměrů $\frac{Z_i + Z_j}{2}$, $i = 1, \dots, m$, $j = 1, \dots, m$, $j \leq i$ a ty následně vzestupně seřadit. Walshovy průměry získáme pomocí funkce `owa` z knihovny `NSM3` a seřadíme je příkazem `sort()`. Následně vypočítáme pozicovou souřadnici $C_{1-\alpha}$.

$$\begin{aligned} C_{1-\alpha} &= \frac{m(m+1)}{4} - u_{1-\alpha} \sqrt{\frac{m(m+1)(2m+1)}{24}} \\ &= \frac{38(38+1)}{4} - u_{1-0.05} \sqrt{\frac{38(38+1)(2 \times 38+1)}{24}} \\ &= \frac{38 \times 39}{4} - u_{0.95} \sqrt{\frac{38 \times 39 \times 77}{24}} \\ &= 370.5 - 1.644854 \sqrt{4754.75} \\ &= 370.5 - 1.644854 \times 68.9547 = 257.0796 \doteq 257 \end{aligned}$$

Hranice intervalu spolehlivosti potom tvoří hodnoty, které se v seřazeném vektoru Walshových průměrů V nachází na 257. pozici a nekonečno.

$$\begin{aligned} (d, h) &= \left(V^{(C_{1-\alpha})}; \infty \right) \\ &= \left(V^{(257)}; \infty \right) \\ &= \left(-4.44 \times 10^{-16}; \infty \right) \end{aligned}$$

```

782 walsh <- NSM3::owa(rep(0, m), na.omit(simd.12L0))$owa
783 V <- sort(walsh)
784 C1 <- m * (m + 1) / 4 - qnorm(1 - alpha) * sqrt(m * (m + 1) * (2 * m + 1) / 24) #
      257.0796
785 V[257] # -4.440892e-16

```

- **Závěr testování**

Protože $\tilde{x}_0 = 0$ náleží do 95% empirického levostranného intervalu spolehlivosti, tj. $\tilde{x}_0 = 0 \in IS$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

5. **Testování p -hodnotou** Příslušou p -hodnotu vypočítáme pomocí vzorce $\Pr(S_A > s_A)$, kde S_A je spojitá náhodná veličina pocházející ze standardizovaného normálního rozdělení, tj. $S_A \sim N(0, 1)$ a $s_A = 1.587999$. K výpočtu použijeme příkaz `pnorm()`.

- **p -hodnota**

$$p\text{-hodnota} = \Pr(S_A \geq s_A) = \Pr(S_A \geq 1.587999) = 1 - \Pr(S_A \leq 1.587999) = 0.05614326 \doteq 0.05614$$

```

786 p.hodnota <- 1 - pnorm(SA) # 0.05614326

```

- **Závěr testování**

Protože p -hodnota = 0.05614 je větší než $\alpha = 0.05$, H_0 nezamítáme na hladině významnosti $\alpha = 0.05$.

6. **Interpretace výsledků**

Za základě všech tří typů testování nezamítáme nulovou hypotézu na hladině významnosti $\alpha = 0.05$. Střední hodnota prvního měření vertikálního průměru délky těla klíční kosti na levé straně není statisticky významně větší než střední hodnota druhého měření.

7. **Grafická vizualizace výsledku testování**

Porovnání naměřených hodnot získaných v rámci prvního a druhého měření můžeme zobrazit pomocí krabicového diagramu analogicky jako v příkladu 9.11 (viz obrázek ??).

Poznámka: Asymptotickou variantu Wilcoxonova párového testu můžeme provést pomocí funkce `wilcox.test()`. Vstupními parametry budou vektor rozdílů vertikálních průměrů středu délky těla klíční kosti na levé straně naměřených na prvním a druhém měření (`simd.12L`), argument `mu = 0` určující, že rozdíl porovnáme s konstantou $\tilde{z}_0 = 0$, požadavek na výpočet hranic intervalu spolehlivosti (`conf.int = T`), hodnota hladiny významnosti α (`conf.level = 0.95`), typ zvolené alternativní hypotézy (`alternative = 'greater'`) a argument `correct = T`, který umožňuje použití právě asymptotické varianty Wilcoxonova párového testu.

```

787 wilcox.test(simd.12L, mu = 0, conf.int = T, conf.level = 0.95,
788             alternative = 'greater', correct = T) # interpolovany IS i p-hodnota

```

Součástí výstupu je hodnota testovací statistiky $S_E V = 480$, interpolované hranice 95% asymptotického empirického oboustranného intervalu spolehlivosti $-5.685398e-05$ a `lnf` pro rozdíl hodnot naměřených při prvním a druhém měření (zpřesnění hranic bylo provedeno procesem nazývaným *interpolace*), a p -hodnota `p-value = 0.05693`. Jediné, co musíme stanovit zvlášť, jsou dolní a horní hranice kritického oboru. Všimněme si, že součástí výstupu je také hodnota pseudomediánu (`pseudo`)`median = 0.04497987` rozdílů prvního a druhého měření. Tento pseudomedián je vypočítaný jako interpolovaná prostřední hodnota z posloupnosti seřazených Walshových průměrů. My však preferujeme hodnotu klasického mediánu rozdílů (`median = 0.03`), kterou bychom získali příkazem `median()`.

Poznámka: Výstup funkce `wilcox.test()` v tomto případě provází čtyři varovné hlášky. První dvě hlášky ve znění: `ln wilcox.test.default(simd.12L, mu = 0, conf.int = T, conf.level = 0.95): cannot compute exact p-value (confidence`

Wilcoxon signed rank test with continuity correction	789
	790
data: simd.12L	791
V = 480, p-value = 0.05693	792
alternative hypothesis: true location is greater than 0	793
95 percent confidence interval:	794
-5.685398e-05 Inf	795
sample estimates:	796
(pseudo)median	797
0.04497987	798
	799

interval) with ties. nás upozorňuje na výskyt duplicitních rozdílů $|Z_1 - 0| = |Z_{11} - 0| = |Z_{13} - 0| = |Z_{29} - 0| = |Z_{35} - 0| = |Z_{40} - 0| = 0.03$, $|Z_5 - 0| = |Z_{16} - 0| = |Z_{25} - 0| = 0.06$, $|Z_{17} - 0| = |Z_{18} - 0| = 0.01$ a na použití modifikovaného postupu na výpočet p -hodnoty (resp. hranic intervalů spolehlivosti). Třetí a čtvrtá hláška ve znění: In wilcox.test.default(simd.12L, mu = 0, conf.int = T, conf.level = 0.95, : cannot compute exact p-value (confidence interval) with zeroes. nás upozorňuje na výskyt nulových rozdílů $|Z_{20} - 0| = |Z_{33} - 0| = 0$ a na použití modifikovaného postupu na výpočet p -hodnoty (resp. hranic intervalů spolehlivosti). Pro nás však tyto varovné hlášky nemají žádný závažný důsledek.



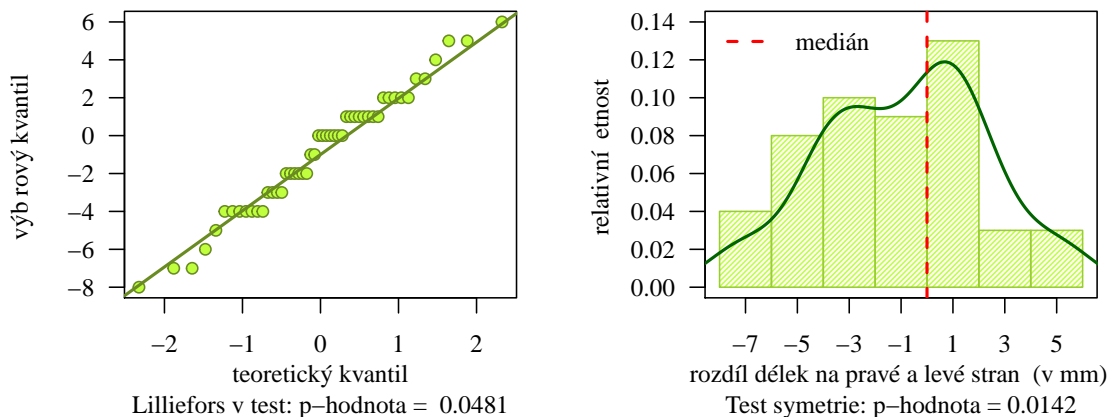
Příklad 9.24. Wilcoxonův párový asymptotický test

Máme datový soubor 03-paired-means-clavicle2.txt obsahující údaje o délkách klíční kosti (*clavicula*) z pravé strany (*length.R*) a levé strany (*length.L*) z anglického souboru dokumentovaných skeletů (Parsons, 1916, viz soubor D-03-paired-means-clavicle2). Na hladině významnosti $\alpha = 0.05$ testujte hypotézu, že střední hodnota délky klíční kosti u žen z pravé strany je větší než střední hodnota délky klíční kosti u žen z levé strany.

Řešení příkladu 9.24

Zadání příkladu je shodné se zadáním příkladu 9.12. Rozdíly délky těla klíčních kostí z levé a pravé strany u žen nepochází z normálního rozdělení (p -hodnota = 0.0481, viz obrázek ??). Bude-li splněn předpoklad symetrického rozdělení rozdílů okolo mediánu, použijeme k ověření otázky ze zadání opět asymptotickou variantu Wilcoxonova neparametrického párového testu.

```
800 data <- read.delim('00-Data//03-paired-means-clavicle2.txt')
801 # head(data)
802 data.RL <- data[data$sex == 'f', c('length.R', 'length.L')]
803 data.RL <- na.omit(data.RL)
804 length.RLF <- data.RL$length.R - data.RL$length.L
805 n <- length(length.RLF) # 32
```



Obrázek 36: Míra symetrie rozdílu délky těla klíčních kostí na levé a pravé straně žen

Protože na základě testu symetrie hypotézu o symetrickém rozdělení rozdílů okolo mediánu zamítáme na hladině významnosti $\alpha = 0.05$ (p -hodnota = 0.0142), není možné otázku ze zadání ověřit pomocí Wilcoxonova neparametrického testu. K otestování můžeme použít pouze znaménkový test (viz řešení příkladu 9.12). ★

Příklad 9.25. Wilcoxonův párový asymptotický test

Máme datový soubor 21-goldman-measures.csv obsahující údaje o anteroposteriorním průměru diafýzy pažní kosti na pravé straně (humer.ADR) a na levé straně (humer.ADL) u mužů (sex = 'm' a sex == 'm?') a žen (sex = 'f' a sex == 'f?') pohřbených v oblasti Indian Knoll v Kentucky (viz sekce ??). Na hladině významnosti $\alpha = 0.05$ testujte hypotézu, že střední hodnota anteroposteriorního průměru diafýzy pažní kosti u mužů na levé straně je větší než na pravé straně. K testování využijte všechna kompletní měření týkající se mužů (skupiny muž (sex == 'm') i pravděpodobně muž (sex == 'm?')).

Řešení příkladu 9.25

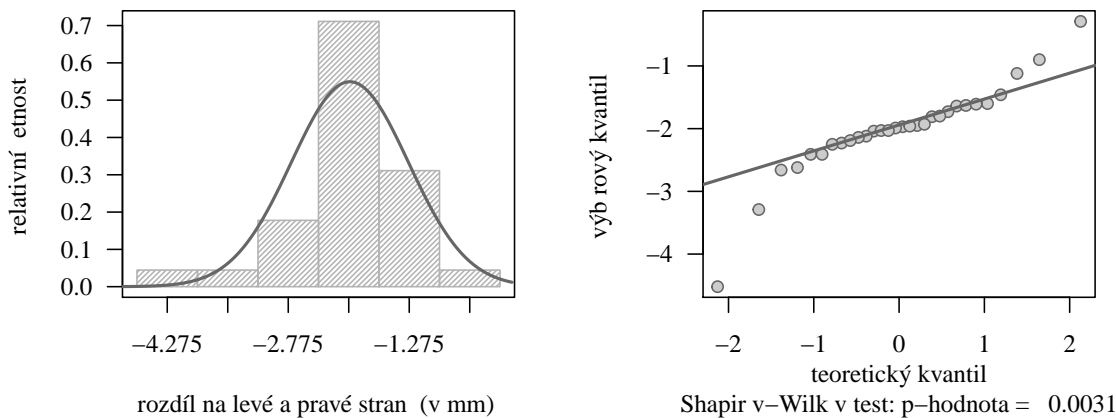
Načteme datový soubor a pomocí operátoru [] z něj vybereme pouze údaje o anteroposteriorním průměru diafýzy pažní kosti na levé straně (humer.ADL) a na pravé straně (humer.ADR) týkající se pouze mužů (v souladu se zadáním uvažujeme obě skupiny mužů (sex %in% c('m', 'm?'))) z pohřebiště Indian Knoll v Kentucky (pop == 'Indian Knoll'). Údaje vložíme do proměnné data.m. Z datové tabulky následně odstraníme chybějící údaje a zjistíme rozsah náhodného výběru.

```
806 data <- read.delim('00-Data//21-goldman-measures.csv', sep = ';')
807 #head(data)
808 data.m <- data[data$sex %in% c('m', 'm?') & data$pop == 'Indian Knoll',
809               c('humer.ADL', 'humer.ADR')]
810 data.m <- na.omit(data.m)
811 dim(data.m) # 30x2
```

Datový soubor obsahuje kompletní údaje o anteroposteriorním průměru diafýzy pažní kosti na pravé i levé straně u 30 mužů z populace z oblasti Indian Knoll. Jelikož máme za úkol porovnat párový znak, použijeme k tomuto porovnání párový test. Prvním krokem k provedení tohoto testu je vytvoření rozdílů hodnot naměřených na pravé a levé straně. V druhém kroku je potřeba ověřit předpoklad normality těchto rozdílů.

```
812 humer.ADL <- data.m$humer.ADL
813 humer.ADR <- data.m$humer.ADR
814 humer.ADLR <- humer.ADL - humer.ADR
```

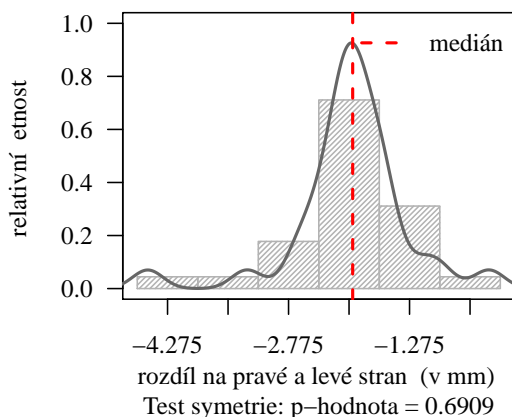
Rozsah náhodného výběru ($n = 30$) je přesně na horní hranici použití Shapirova-Wilkova testu ($\alpha = 0.05$) k ověření předpokladu normality rozdílů. Graficky zhodnotíme normalitu QQ-diagramem a histogramem. Naměřené hodnoty rozdělíme do šesti ekvidistantních intervalů se šířkou 3.5 mm prostřednictvím stanovených hranic $-2, 1.5, \dots, 19$ (viz obrázek 37).



Obrázek 37: Histogram a QQ-diagram rozdílů epikondylární šířky pažní kosti na pravé a levé straně

Protože p -hodnota vyšla < 0.0031 , což je menší než 0.05 , nulovou hypotézu o normalitě rozdílů zamítáme na hladině významnosti $\alpha = 0.05$. Náhodný výběr rozdílů tedy nepochází z normálního rozdělení.

Protože náhodný výběr rozdílů nepochází z normálního rozdělení, nemůžeme na otestování hypotézy ze zadání použít parametrický párový test. Vzhledem k rozsahu náhodného výběru ($n = 30$) použijeme k otestování nulové hypotézy neparametrickou alternativu párového testu, a sice Wilcoxonův asymptotický neparametrický párový test. Nejprve ověříme nezbytný předpoklad – symetrii rozdílů okolo mediánu v kombinaci s histogramem superponovaným křivkou jádrového odhadu hustoty.



Obrázek 38: Histogram a QQ-diagram rozdílů anteroposteriorního průměru diafýzy pažní kosti na pravé a levé straně

Na základě testu symetrie nezamítáme hypotézu o symetrickém rozdělení rozdílů hodnot naměřených na pravé a levé straně okolo jejich mediánu (p -hodnota = '0.6909'). Z histogramu na obrázku 38 vidíme, že data jsou okolo mediánu skutečně rozložena symetricky. Předpoklad symetrie je tedy splněn, a proto můžeme asymptotickou variantu Wilcoxonova neparametrického testu použít. V souladu se zadáním budeme testovat (nulovou) hypotézu, že střední hodnota anteroposteriorního průměru diafýzy pažní kosti u mužů na levé straně je větší než na pravé straně. Zbývá tedy dodefinovat alternativní hypotézu tak, aby byla doplňkem k hypotéze nulové.

1. Stanovení hypotéz

- **slovní formulace** nulové a alternativní hypotézy

H_0 : Medián rozdílů anteroposteriorního průměru diafýzy pažní kosti na levé a pravé straně je větší nebo roven nule.

H_1 : Medián rozdílů anteroposteriorního průměru diafýzy pažní kosti na levé a pravé straně je menší než nula.

- **matematická formulace** nulové a alternativní hypotézy

H_0 : $\tilde{z} \geq \tilde{z}_0$, kde $\tilde{z}_0 = 0$

H_1 : $\tilde{z} < \tilde{z}_0$, kde $\tilde{z}_0 = 0$ (levostranná alternativa)

2. Volba hladiny významnosti

- Hladina významnosti $\alpha = 0.05$.

3. Testování kritickým oborem

- **Testovací statistika**

V úvodu příkladu jsme vytvořili vektor humer.ADLR (Z_i , $i = 1, \dots, 30$) jako vektor rozdílů hodnot anteroposteriorních průměrů diafýzy pažní kosti naměřených na levé a pravé straně. Dále jsme stanovili

rozdíly $Z_i - \tilde{z}_0$. Protože ale $\tilde{z}_0 = 0$, jsou rozdíly $Z_i - \tilde{z}_0$ rovny původnímu vektoru rozdílů Z_i , tj. $Z_i - \tilde{z}_0 = Z_i - 0 = Z_i$. V dalším kroku se zaměříme na stanovení počtu nenulových rozdílů a zjistíme, že žádný rozdíl $Z_i - 0$ není nulový (počet nenulových rozdílů $m = 30$). Nyní vytvoříme vektor `humer.ADLR0` obsahující ty rozdíly mezi pravou a levou stranou, pro něž rozdíl $Z_i - \tilde{z}_0$ nebyl rovný nule. Délka tohoto vektoru bude v tomto případě $m = n = 30$. Kvůli zachování univerzality postupu budeme v dalších krocích testování používat právě vektor `humer.ADLR0`, který je jinak totožný s vektorem `humer.ADLR`. Dále si označíme znaménkem $+$ rozdíly $Z_i - 0$, které jsou kladné a znaménkem $-$ rozdíly $Z_i - 0$, které jsou záporné, nebo 0. Nakonec vypočítáme absolutní hodnoty rozdílů $|Z_i - 0| = |Z_i|$, $i = 1, \dots, m$ a stanovíme pořadí těchto rozdílů v absolutních hodnotách.

```

815 z0 <- 0
816 n <- length(humer.ADLR)
817 m <- sum(humer.ADLR - z0 != 0)
818 humer.ADLR0 <- humer.ADLR
819 humer.ADLR0[humer.ADLR0 == 0] <- NA
820
821 I <- (humer.ADLR0 > z0)
822 Ri <- rank(abs(humer.ADLR0 - 0), na.las = "keep")
823 tab <- data.frame(rbind("Xi" = humer.ADL, "Yi" = humer.ADR,
824                       "Zi" = humer.ADLR, "Zi-z0" = humer.ADLR - z0, "I" = I,
825                       "Zi-z0" = abs(humer.ADLR0 - z0), "Ri" = Ri))
826 names(tab) <- 1 : n
827 SE <- sum(Ri * I, na.rm = T) # 383

```

Žádný z rozdílů $Z_i - 0$, $i = 1, \dots, 28$ není kladný, a tedy nás nezajímá pořadí žádného rozdílu. Hodnota testovací statistiky S_E , která je definovaná jako součet těchto pořadí, bude v tomto případě rovná 0.

$$S_E = \sum_{i=1}^m I(Z_i - 0 > 0) = 0.$$

Nyní dopočítáme hodnotu testovací statistiky S_A .

$$\begin{aligned}
S_A &= \frac{S_E - \frac{m(m+1)}{4}}{\sqrt{\frac{m(m+1)(2m+1)}{24}}} = \frac{0 - \frac{30(30+1)}{4}}{\sqrt{\frac{30(30+1)(2 \times 30+1)}{24}}} = \frac{0 - \frac{30 \times 31}{4}}{\sqrt{\frac{30 \times 31 \times 61}{24}}} = \frac{0 - 232.5}{\sqrt{2363.75}} = \frac{-232.5}{48.61841} \\
&= -4.782139 \doteq -4.7821
\end{aligned}$$

```

828 SE <- sum(Ri * I, na.rm = T) # 0
829 SA <- (SE - m * (m + 1) / 4) / sqrt(m * (m + 1) * (2 * m + 1) / 24) # -4.782139

```

• Kritický obor

Protože $S_A \sim N(0, 1)$, horní hranici kritického oboru nalezneme pomocí funkce `qnorm()` jako α kvantil standardizovaného normálního rozdělení $N(0, 1)$. Dolní hranicí bude mínus nekonečno.

$$\begin{aligned}
W &= (-\infty; u_\alpha) \\
&= (-\infty; u_{0.05}) \\
&= (-\infty; -1.644854) \\
&= (-\infty; -1.6449)
\end{aligned}$$


```
830 alpha <- 0.05
831 qnorm(alpha) # -1.644854
```

- **Závěr testování**

Protože realizace testovací statistiky $s_A = -4.7821$ náleží do kritického oboru, tj. $s_A \in W$, H_0 zamítáme na hladině významnosti $\alpha = 0.05$.

4. Testování intervalem spolehlivosti

- **Interval spolehlivosti**

Proti levostranné alternativě postavíme pravostranný interval spolehlivosti. Abychom mohli stanovit hranice 95 % intervalu spolehlivosti, musíme nejprve vypočítat $\frac{m(m+1)}{2}$ Walshových průměrů $\frac{Z_i+Z_j}{2}$, $i = 1, \dots, m$, $j = 1, \dots, m$, $j \leq i$ a ty vzestupně seřadit. Následně vypočítáme pozicovou souřadnici C_α .

$$\begin{aligned} C_{1-\alpha} &= \frac{m(m+1)}{4} - u_\alpha \sqrt{\frac{m(m+1)(2m+1)}{24}} \\ &= \frac{30(30+1)}{4} - u_{0.05} \sqrt{\frac{30(30+1)(2 \times 30+1)}{24}} \\ &= \frac{30 \times 31}{4} - u_{0.95} \sqrt{\frac{30 \times 31 \times 61}{24}} \\ &= 232.5 + 1.644854 \sqrt{2363.75} \\ &= 232.5 + 1.644854 \times 48.61841 = 312.4702 \doteq 312 \end{aligned}$$

Hranice intervalu spolehlivosti potom tvoří mínus nekonečno a hodnota, které se v seřazeném vektoru Walshových průměrů V nachází na 312. pozici.

$$\begin{aligned} (d, h) &= \left(-\infty; V^{(C_{\alpha/2})}\right) \\ &= \left(-\infty; V^{(312)}\right) \\ &= \left(-\infty; -1.83\right) \end{aligned}$$

```
832 walsh <- NSM3::owa(rep(0, m), na.omit(humer.ADLR0))$owa
833 V <- sort(walsh)
834 C2 <- m * (m + 1) / 4 - qnorm(alpha) * sqrt(m * (m + 1) * (2 * m + 1) / 24) # 312.4702
835 V[C2] # -1.83
```

- **Závěr testování**

Protože $\tilde{x}_0 = 0$ nenáleží do 95% empirického oboustranného intervalu spolehlivosti, tj. $\tilde{x}_0 = 0 \notin IS$, H_0 zamítáme na hladině významnosti $\alpha = 0.05$.

5. **Testování p -hodnotou** Příslušou p -hodnotu vypočítáme pomocí vzorce $2\Pr(S_A \leq s_A)$, kde S_A je spojitá náhodná veličina pocházející ze standardizovaného normálního rozdělení, tj. $S_A \sim N(0, 1)$ a $s_A = -4.782139$. $\Pr(S_A \leq s_A)$ je tedy distribuční funkce standardizovaného normálního rozdělení $N(0, 1)$ v hodnotě -4.782139 . Její hodnotu spočítáme příkazem `pnorm()`.

- **p -hodnota**

$$p\text{-hodnota} = \Pr(S_A \leq s_A) = \Pr(S_A \leq -4.782139) = 8.671988 \times 10^{-7} \doteq 8.6720 \times 10^{-7}$$

- **Závěr testování**

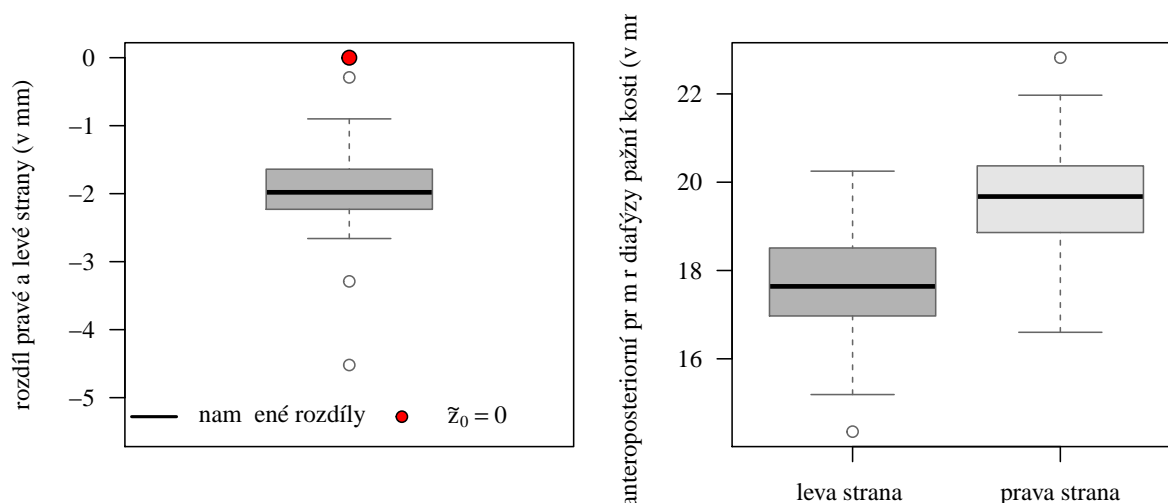
Protože p -hodnota 8.6720×10^{-7} je menší než $\alpha = 0.05$, H_0 zamítáme na hladině významnosti $\alpha = 0.05$.

6. Interpretace výsledků

Za základě všech tří typů testování zamítáme nulovou hypotézu na hladině významnosti $\alpha = 0.05$. Anteroposteriorní průměr diafýzy pažní kosti na levé straně u muže z oblasti Indian Knoll je statisticky významně menší než na pravé straně.

7. Grafická vizualizace výsledku testování

Porovnání anteroposteriorních průměrů diafýzy pažní kosti naměřených na pravé a levé straně můžeme zobrazit pomocí krabicového diagramu (viz obrázek 39).



Obrázek 39: Krabicový diagram rozdílů anteroposteriorního průměru diafýzy pažní kosti na pravé a levé straně

Poznámka: Asymptotickou variantu Wilcoxonova párového testu můžeme provést pomocí funkce `wilcox.test()`. Vstupními parametry budou vektor hodnot anteroposteriorního průměru pažní kosti naměřených na levé straně (`humer.ADL`), vektor hodnot anteroposteriorního průměru pažní kosti naměřených na pravé straně (`humer.ADR`), argument `paired = T` určující, že oba vektory považujeme za párová pozorování, argument povolující výpočet hranic intervalu spolehlivosti `conf.int = T`, hodnota hladiny významnosti α zadaná prostřednictvím koeficientu spolehlivosti $1 - \alpha$ nastavením hodnoty argumentu `conf.level = 0.95`, typ zvolené alternativní hypotézy (levostranná) zadaný argumentem `alternative = 'less'` a argument `correct = T`, který zvolí k výpočtu právě asymptotickou variantu Wilcoxonova neparametrického testu.

```
837 wilcox.test(humer.ADL, humer.ADR, paired = T, conf.int = T, conf.level = 0.95,
838             alternative = 'less', correct = T) # interpolovany IS i p-hodnota
```

Součástí výstupu je hodnota testovací statistiky $S_E V = 0$, interpolované hranice 95% asymptotického empirického levostranného intervalu spolehlivosti $-\ln f$ a -1.825024 pro rozdíl hodnot naměřených na levé a pravé straně (zpřesnění hranic bylo provedeno procesem nazývaným *interpolace*), a p -hodnota $p\text{-value} = 9.115e-07$. Jediné, co musíme stanovit zvlášť, je horní hranice kritického oboru (dolní hranicí je mínus nekonečno). Součástí výstupu je také hodnota pseudomedianu (`pseudo`)`median = -1.980009` rozdílů hodnot naměřených na pravé a levé straně. My však preferujeme hodnotu klasického medianu rozdílů (`median = -1.98`), kterou bychom získali kupříkladu příkazem `median()`.

```

Wilcoxon signed rank test with continuity correction

data: humer.ADL and humer.ADR
V = 0, p-value = 9.115e-07
alternative hypothesis: true location shift is less than 0
95 percent confidence interval:
 -Inf -1.825024
sample estimates:
(pseudo)median
 -1.980009

```

839
840
841
842
843
844
845
846
847
848
849

Poznámka: Výstup funkce `wilcox.test()` provází dvě varovné hlášky ve znění: `In wilcox.test.default(humer.ADL, humer.ADR, paired = T, conf.int = T,): cannot compute exact p-value (confidence interval) with ties.` nás upozorňuje na výskyt duplicitních rozdílů $|Z_8| = |Z_{12}| = 2.41$, $|Z_{16}| = |Z_{23}| = 2.03$ a na použití modifikovaného postupu na výpočet p -hodnoty (resp. hranic intervalů spolehlivosti). Hlášky ve znění: `In wilcox.test.default(humer.ADL, humer.ADR, paired = T, conf.int = T,): cannot compute exact p-value (confidence interval) with zeroes.`, které by nás upozorňovaly na výskyt nulových rozdílů se tentokrát nevypisují, protože žádný z rozdílů $|Z_i - 0|$, $i = 1, \dots, 30$ není rovný nule.

★

9.7 Pořadový exaktní test o nezávislosti

Nechť $(X_1, Y_1)^T \dots (X_n, Y_n)^T$ je náhodný výběr z dvourozměrného rozdělení. Na hladině významnosti α testujeme jednu z následujících tří hypotéz oproti příslušné alternativní hypotéze.

$$\begin{array}{lll} H_{01} : \rho = 0 & \text{oproti} & H_{11} : \rho \neq 0 \quad (\text{oboustranná alt.}) \\ H_{02} : \rho \leq 0 & \text{oproti} & H_{12} : \rho > 0 \quad (\text{pravostranná alt.}) \\ H_{03} : \rho \geq 0 & \text{oproti} & H_{13} : \rho < 0 \quad (\text{levostranná alt.}) \end{array}$$


Test nazýváme pořadovým testem o nezávislosti. Testovací statistika má tvar

$$S_E = R_S = 1 - \frac{6}{n(n^2 - 1)} \sum_{i=1}^n (R_i - Q_i)^2, \quad (9.6)$$


kde n je rozsah náhodného výběru, R_i je pořadí náhodné veličiny X_i a Q_i je pořadí náhodné veličiny Y_i . Všimněme si, že statistika S_E je současně Spearmanovým korelačním koeficientem R_S , resp. Spearmanův korelační koeficient je používán jako testovací statistika pořadového testu o nezávislosti.

Kritický obor podle zvolené alternativní hypotézy má tvar

$$\begin{array}{ll} H_{11} : \rho \neq 0 & W = \langle -1; r_n(\alpha/2) \rangle \cup \langle r_n(1 - \alpha/2); 1 \rangle \\ H_{12} : \rho > 0 & W = \langle -1; r_n(\alpha) \rangle \\ H_{13} : \rho < 0 & W = \langle r_n(1 - \alpha); 1 \rangle \end{array}$$

kde $r_n(\alpha/2)$, $r_n(1 - \alpha/2)$, $r_n(1 - \alpha)$ a $r_n(1 - \alpha/2)$ jsou kvantily pro pořadový test o nezávislosti, jejichž hodnoty získáme pomocí softwaru  a funkce `qSpearman()` z knihovny `SuppDists`. Tvary intervalů spolehlivosti si pro pořadový test o nezávislosti neuvádíme. p -hodnota má v závislosti na zvolené alternativní hypotéze jeden z následujících tvarů

$$\begin{array}{ll} H_{11} : \tilde{x} \neq \tilde{x}_0 & p\text{-hodnota} = 2 \min\{\Pr(S_E \leq s_E), \Pr(S_E \geq s_E)\} \\ H_{12} : \tilde{x} > \tilde{x}_0 & p\text{-hodnota} = \Pr(S_E \geq s_E) \\ H_{13} : \tilde{x} < \tilde{x}_0 & p\text{-hodnota} = \Pr(S_E \leq s_E) \end{array}$$

kde S_E je náhodná veličina, s_E je realizace testovací statistiky S_E (viz vzorec 9.6), tedy konkrétní číslo, $\Pr(S_E \geq s_E) = 1 - \Pr(S_E < s_E) = 1 - \Pr(S_E \leq s_E - 1)$, což vyplývá z faktu, že náhodná veličina S_E se řídí tabelovaným (diskrétním) rozdělením pro pořadový test o nezávislosti, a $\Pr(S_E \leq s_E)$ je distribuční funkce tabelovaného rozdělení pro pořadový test o nezávislosti, jejíž hodnotu získáme pomocí  a implementované funkce `pSpearman()` z knihovny `SuppDists`.

Příklad 9.26. Pořadový exaktní test o nezávislosti

Řešení příkladu 9.26



```
850 data <- read.delim('00-Data//21-goldman-measures.csv', sep = ';')
851 head(data)
852 data <- data[data$sex == 'm' & data$pop == 'Hawikuh', c('acetab.HR', 'iblade.LR')]
853 data <- na.omit(data)
854
855 MVN::mvn(data, mvnTest = 'mardia')$multi
856 # sikkmost: 0.041166 # spicatost: 0.9761901
857 MVN::mvn(data, mvnTest = 'hz')$multi # 0.0436649
858 MVN::mvn(data, mvnTest = 'royston')$multi # 0.2507807
859
860 acetab.HRM <- data$acetab.HR
861 iblade.LRM <- data$iblade.LR
862 n <- length(iblade.LRM) # 14
863 alpha <- 0.05
864 rS <- cor(acetab.HRM, iblade.LRM, method = 'spearman') # 0.4856559
865
866 library(SuppDists)
867 qSpearman(alpha / 2, n) # -0.5252747
868 qSpearman(1 - alpha / 2, n) # 0.5384615
869
870 2 * min(pSpearman(rS, n), 1 - pSpearman(rS, n)) # 0.08218591
```

Příklad 9.27. Pořadový exaktní test o nezávislosti

Řešení příkladu 9.27



```
871 data <- read.delim('00-Data//21-goldman-measures.csv', sep = ';')
872 head(data)
873 data <- data[data$sex == 'm' & data$pop == 'Cliff Dweller', c('tibia.LL', 'radius.LL')]
874 data <- na.omit(data)
875 MVN::mvn(data, mvnTest = 'mardia')$multi
876 # sikmost: 0.000229396 # spicatost: 0.0056161
877 MVN::mvn(data, mvnTest = 'hz')$multi # 0.009711723
878 MVN::mvn(data, mvnTest = 'royston')$multi # 0.07266684
879
880 radius.LLM <- data$radius.LL
881 tibia.LLM <- data$tibia.LL
882 n <- length(tibia.LLM) # 17
883 alpha <- 0.05
884 rS <- cor(radius.LLM, tibia.LLM, method = 'spearman') # 0.8725396
885
886 library(SuppDists)
887 qSpearman(1 - alpha, n) # 0.4166667
888 1 - pSpearman(rS, n) # 2.981788e-06
```

Příklad 9.28. Pořadový exaktní test o nezávislosti

Řešení příkladu 9.28



```
889 data <- read.delim('00-Data//21-goldman-measures.csv', sep = ';')
890 head(data)
891 data <- data[data$sex == 'f' & data$pop == 'Tigara', c('RHMLD', 'humer.DR')]
892 data <- na.omit(data)
893 MVN::mvn(data, mvnTest = 'mardia')$multi
894 # sikmost: 0.1269080 # spicatost: 0.38712602
895 MVN::mvn(data, mvnTest = 'hz')$multi # 0.01409578
896 MVN::mvn(data, mvnTest = 'royston')$multi # 0.006155462
897
898 humer.LRF <- data$RHML
899 humer.HDRF <- data$RHHD
900 n <- length(humer.LRF) # 24
901 alpha <- 0.05
902 rS <- cor(humer.LRF, humer.HDRF, method = 'spearman') # 0.4881252
```

```
Error in cor(humer.LRF, humer.HDRF, method = "spearman"): supply both 'x' and 'y' or a matrix-like 'x'
```

903

```
904 library(SuppDists)
905 qSpearman(alpha, n) # -0.3426087
906 pSpearman(rS, n) # 0.9920688
```

9.8 Pořadový asymptotický test o nezávislosti

Pro náhodný výběr o rozsahu $n > 30$ máme možnost použít k otestování nulové hypotézy asymptotickou variantu testu. Tuto variantu nazýváme pořadový asymptotický test o nezávislosti. Stejně jako u exaktního testu si pro asymptotický test uvedeme pouze postup testování pomocí kritického oboru a p -hodnoty. Testovací statistika asymptotického testu má tvar

$$S_A = r_S \sqrt{n-1} \quad (9.7)$$

kde r_S je testovací statistika definovaná vztahem 9.6 a n je rozsah náhodného výběru. Za platnosti nulové hypotézy pochází statistika S_A ze standardizovaného normálního rozdělení, tj.

$$S_A = r_S \sqrt{n-1} \stackrel{H_0}{\sim} N(0, 1).$$

Kritický obor podle zvolené alternativní hypotézy má tvar

$$\begin{array}{ll} H_{11} : \tilde{x} \neq \tilde{x}_0 & W = (-\infty; u_{\alpha/2}) \cup (u_{1-\alpha/2}; \infty) \\ H_{12} : \tilde{x} > \tilde{x}_0 & W = (u_{1-\alpha}; \infty) \\ H_{13} : \tilde{x} < \tilde{x}_0 & W = (-\infty; u_{\alpha}) \end{array}$$

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

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

$$\begin{array}{ll} H_{11} : \tilde{x} \neq \tilde{x}_0 & p\text{-hodnota} = 2 \min\{\Pr(S_A \leq s_A), \Pr(S_A \geq s_A)\} \\ H_{12} : \tilde{x} > \tilde{x}_0 & p\text{-hodnota} = \Pr(S_A \geq s_A) \\ H_{13} : \tilde{x} < \tilde{x}_0 & p\text{-hodnota} = \Pr(S_A \leq s_A) \end{array}$$

kde S_A je náhodná veličina, s_A je realizace testovací statistiky S_A (viz vzorec 9.7), tedy konkrétní číslo, $\Pr(S_A \geq s_A) = 1 - \Pr(S_A \leq s_A)$, což vyplývá z faktu, že náhodná veličina S_A pochází z normálního (spojitého) rozdělení (viz kapitola ??), a $\Pr(S_A \leq s_A)$ je distribuční funkce standardizovaného normálního rozdělení, jejíž hodnotu získáme pomocí \mathbb{R} a implementované funkce `pnorm()`.

Příklad 9.29. Pořadový asymptotický test o nezávislosti

Řešení příkladu 9.29



```
907 data <- read.delim('00-Data//16-anova-head.txt')
908 head(data)
909 data <- data[data$sex == 'f', c('head.L', 'bizyg.W')]
910 data <- na.omit(data)
911 MVN::mvn(data, mvnTest = 'mardia')$multi
912 # sikkmost: 0.0427523 # spicatost: 0.5455480
913 MVN::mvn(data, mvnTest = 'hz')$multi # 0.09943109
914 MVN::mvn(data, mvnTest = 'royston')$multi # 0.04040737
915
916 head.LF <- data$head.L
917 bizyg.WF <- data$bizyg.W
918 n <- length(head.LF) # 100
919 alpha <- 0.05
920
921 rS <- cor(head.LF, bizyg.WF, method = 'spearman') # 0.1712287
922 sA <- rS * sqrt(n - 1) # 1.703704
923 qnorm(alpha / 2) # -1.959964
924 qnorm(1 - alpha / 2) # 1.959964
925 2 * min(pnorm(sA), 1 - pnorm(sA)) # 0.08843632
```

Příklad 9.30. Pořadový asymptotický test o nezávislosti

Řešení příkladu 9.31

```
926 data <- read.delim('00-Data//16-anova-head.txt')
927 head(data)
928 data <- data[data$sex == 'm', c('bigo.W', 'bizyg.W')]
929 data <- na.omit(data)
930 MVN::mvn(data, mvnTest = 'mardia')$multi
931 # skimost: 0.002218132 # spicatost: 0.05247948
932 MVN::mvn(data, mvnTest = 'hz')$multi # 0.006343478
933 MVN::mvn(data, mvnTest = 'royston')$multi # 0.00644503
934
935 bigo.WM <- data$bigo.W
936 bizyg.WM <- data$bizyg.W
937 n <- length(bigo.WM) # 75
938 alpha <- 0.05
939 rS <- cor(bigo.WM, bizyg.WM, method = 'spearman') # 0.4648327
940 sA <- rS * sqrt(n - 1) # 3.998642
941
942 qnorm(1 - alpha) # 1.644854
943 1 - pnorm(sA) # 3.185348e-05
```



Příklad 9.31. Pořadový asymptotický test o nezávislosti

Řešení příkladu 9.31

```
944 dataA <- read.delim('00-Data//28-one-sample-mean-skull-m5.txt')
945 head(dataA)
946 data <- dataA[dataA$year == '150', c('skull.V', 'skull.NH')] # B H V NH # -3300:B-H:YYN
# -3300:B-NH:YYN # 150:B-H:YNY
947 data <- na.omit(data)
948 MVN::mvn(data, mvnTest = 'mardia')$multi
949 # skimost: 0.002218132 # spicatost: 0.05247948
950 MVN::mvn(data, mvnTest = 'hz')$multi # 0.006343478
951 MVN::mvn(data, mvnTest = 'royston')$multi # 0.00644503
952
953 bigo.WM <- data$bigo.W
954 bizyg.WM <- data$bizyg.W
955 n <- length(bigo.WM) # 75
956 alpha <- 0.05
957 rS <- cor(bigo.WM, bizyg.WM, method = 'spearman') # 0.4648327
```

```
Error in cor(bigo.WM, bizyg.WM, method = "spearman"): supply both 'x' and 'y' or a matrix-like 'x' 958
```

```
959 sA <- rS * sqrt(n - 1) # 3.998642
```



```
Warning in sqrt(n - 1): NaNs produced
```

960

```
961 qnorm(1 - alpha) # 1.644854  
962 1 - pnorm(sA) # 3.185348e-05
```

```

963 dataA <- read.delim('00-Data//21-goldman-measures.csv', sep = ';')
964 head(dataA)
965 jmena <- c('femur.LL', 'LFBL', 'LFAB', 'femur.HDL', 'LFMLD', 'LFAPD',
'acetab.HL', 'acetab.HR')
966 for (i in 1:length(jmena)){
967   for(j in 1:length(jmena)){
968     if(i != j){
969       print(paste(jmena[i], '-', jmena[j]))
970 data <- dataA[dataA$sex == 'f' & dataA$pop == 'Indian Knoll', c(jmena[i], jmena[j])] #
971
972 data <- na.omit(data)
973 print(MVN::mvn(data, mvnTest = 'mardia')$multi )
974 # skimost: 0.002218132 # spicatost: 0.05247948
975 print(MVN::mvn(data, mvnTest = 'hz')$multi) # 0.006343478
976 print(MVN::mvn(data, mvnTest = 'royston')$multi) # 0.00644503
977   }
978 }
979 }
980
981
982 dataA <- read.delim('00-Data//21-goldman-measures.csv', sep = ';')
983 head(dataA)
984 jmena <- c('humer.LL', 'humer.EBL', 'humer.DL', 'LHMLD', 'humer.ADL',
'humer.LR',
985 'humer.EBR', 'humer.DR', 'RHMLD',
986 'humer.ADR', 'femur.LL', 'LFBL', 'LFEB', 'LFAB', 'femur.HDL', 'LFMLD',
987 'LFAPD', 'femur.HDR', 'tibia.LL', 'LTMLD', 'LTAPD', 'BIB',
988 'acetab.HL', 'acetab.HR')
989 jmena <- names(dataA)[3:45]
990 for (i in 1:length(jmena)){
991   for(j in 2){
992     if(i != j){
993       print(paste(jmena[i], '-', jmena[j]))
994 data <- dataA[dataA$sex == 'm' & dataA$pop == 'Indian Knoll', c(jmena[i], jmena[j])] #
995
996 data <- na.omit(data)
997 #print(MVN::mvn(data, mvnTest = 'mardia')$multi )
998 # skimost: 0.002218132 # spicatost: 0.05247948
999 #print(MVN::mvn(data, mvnTest = 'hz')$multi) # 0.006343478
1000 #print(MVN::mvn(data, mvnTest = 'royston')$multi) # 0.00644503
1001
1002 humer.EBL <- data$humer.EBL
1003 y <- data[, jmena[i]]
1004 print(cor(humer.EBL, y, method = 'spearman'))
1005   }
1006 }
1007 }
1008
1009 #humer.EBL + humer.LL
1010 #humer.EBL + humer.DL
1011
1012 bigo.WM <- data$bigo.W
1013 bizyg.WM <- data$bizyg.W
1014 n <- length(bigo.WM) # 75
1015 alpha <- 0.05
1016 rS <- cor(bigo.WM, bizyg.WM, method = 'spearman') # 0.4648327

```

```
Error in cor(bigo.WM, bizyg.WM, method = "spearman"): supply both 'x' and 'y' or a matrix-like 'x'
```

1017

```
1018 sA <- rS * sqrt(n - 1) # 3.998642
```

```
Warning in sqrt(n - 1): NaNs produced
```

1019

```
1020 qnorm(1 - alpha) # 1.644854  
1021 1 - pnorm(sA) # 3.185348e-05
```