1 R-Sweave

2 Kvantily a bodové odhady parametrů

- Příklad 2.1 ... $x > \log(1+x) > \frac{x}{1+x}$
- Příklad 2.2 ... Pro $n \to \infty$ konverguje Studentovo rozdělení o n stupních volnosti ke standardizovanému normálnímu rozdělení (uspokojivá shoda nastává cca pro $n \geq 50$)
- Příklad 2.3 ... Uvedené rovnosti platí, tj. $\chi^2_1(1-\alpha) = u^2_{1-\alpha/2}$; $F_{1,n-1}(1-\alpha) = t^2_{n-1}(1-\alpha/2)$; $\chi^2_n(1-\alpha) = \Gamma_{\frac{n}{2},2}(1-\alpha)$, a to pro libovolné n i α
- Příklad 2.4 ... $S^2_n \sim \Gamma(\frac{n}{2}, \frac{2\sigma^2}{n})$ exaktně; $S^2_n \sim N(\sigma^2, \frac{2\sigma^4}{n})$ asymptoticky (uspokojivá shoda nastává při zadaných parametrech cca pro $n \geq 500$); $S_n \sim \Gamma_G(\sqrt{\frac{2\sigma^2}{n}}, 2, \frac{n}{2})$ exaktně; $S_n \sim N(\sigma, \frac{\sigma^2}{2n})$ asymptoticky (uspokojivá shoda nastává při zadaných parametrech cca pro $n \geq 100$).
- Příklad 2.5
  - (a) Pro $X \sim N(\mu, \sigma^2)$, $\mu = 150$, $\sigma = 10^2$, pro $n = 5$, $n = 50$ i $n = 100$ je aktuální spolehlivost (pst pokrytí) $1 - \widehat{\alpha}$ blízká nominální spolehlivosti $1 - \alpha$;
  - (b) Pro $X \sim [pN(\mu, \sigma^2) + (1-p)N(\mu, \sigma_2^2)]$, $p = 0.9$, $\sigma_2^2 = 20^2$, pro $n = 5$, $n = 50$ i $n = 100$ je aktuální spolehlivost (pst pokrytí) $1 - \widehat{\alpha}$ blízká nominální spolehlivosti $1 - \alpha$; tj. 10% příměs lišící se pouze v rozptylu aktuální spolehlivost $1 - \widehat{\alpha}$ 95% DIS pro μ, když σ^2 neznáme, nijak závažným způsobem neovlivňuje.
- Příklad 2.6
  - (a) $X \sim N(5, 15)$, $n = 50$, ... Po opakovaném spuštění MC studie vidíme, že aktuální pravděpodobnost pokrytí $1 - \widehat{\alpha}$ nabývá rovnoměrně častokrát vyšší i nižší hodnoty než je nominální pravděpodobnost pokrytí $1 - \alpha = 0.95$. 95% DIS pro μ, když σ^2 známe, není v tomto případě konzervativní ani liberální. Aktuální spolehlivost je celkem výrazně vyšší / nižší než 0.95, což je způsobeno nízkým rozsahem náhodných výběrů.
  - (b) $X \sim N(5, 15)$, $n = 1000$ ... Po opakovaném spuštění MC studie vidíme, že aktuální pravděpodobnost pokrytí $1 - \widehat{\alpha}$ nabývá rovnoměrně častokrát vyšší i nižší hodnoty než je nominální pravděpodobnost pokrytí $1 - \alpha = 0.95$. DIS není ani konzervativní ani liberální.
  - (c) smíšené rozdělení s rozdílnými rozptyly, $n = 50$; n = 1000 ... Aktuální pravděpodobnost pokrytí DIS je systematicky podhodnocována (systematicky je aktuální pst pokrytí menší než nominální pst pokrytí), tj. v tomto případě je DIS liberální.
  - (d) smíšené rozdělení s rozdílnými středními hodnotami, $n = 50$ ... Aktuální pravděpodobnost pokrytí 95% DIS je systematicky podhodnocována (systematicky je aktuální pst pokrytí menší než nominální pst pokrytí), tj. v tomto případě je DIS liberální. Efekt liberálnosti zde není tak zřejmý jako v situaci (e), díky nižšímu rozsahu náhodných výběrů.
  - (e) smíšené rozdělení s rozdílnými středními hodnotami, $n = 1000$ ... Aktuální pravděpodobnost pokrytí 95% DIS je systematicky podhodnocována, tj. v tomto případě je DIS liberální. Efekt liberálnosti je velmi výrazný kvůli vysokým rozsahům náhodných výběrů.

3 Postačující statistika, monotónní poměr věrohodnosti a rovnoměrně nejsilnější test

- Příklad 3.3 ... Z animace vidíme, že poměr věrohodnosti pro parametr p je skutečně monotónní funkcí statistiky $X = \sum_{i=1}^{N} X_i$.
- Příklad 3.4 ...
  - Pro H_{02} rovnoměrně nejsilnější test existuje. Je jím například Waldův, či skóre test o pravděpodobnosti, neboť testovací statistika obou testů je založena na postačující statistice $X = \sum_{i=1}^{N} X_i$, která je pouze přičítána/odečítána/násobena/dělena konstantami p_0 a N.
  - Pro H_{03} rovnoměrně nejsilnější test existuje. Je jím například Waldův, či skóre test o pravděpodobnosti, neboť testovací statistika obou testů je založena na postačující statistice $X = \sum_{i=1}^{N} X_i$, která je pouze přičítána/odečítána/násobena/dělena konstantami p_0 a N.
  - Pro H_{01} rovnoměrně nejsilnější test neexistuje. Z přednášky z SI II víme, že rovnoměrně nejsilnější test neexistuje pro oboustrannou alternativu.
- Z animace vidíme, že poměr věrohodnosti pro parametr μ normálního rozdělení je skutečně monotónní funkcí statistiky $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$.
  - Pro H_{02} rovnoměrně nejsilnější test existuje. Je jím například Waldův test o parametru μ při známém rozptylu σ^2, neboť testovací statistika tohoto testu Z_W je založena na postačující statistice $\bar{X}$, která je pouze přičítána/odečítána/násobena/dělena konstantami μ_0, $\sqrt{n}$ a σ.
  - Pro H_{03} rovnoměrně nejsilnější test existuje. Je jím například Waldův test o parametru μ při známém rozptylu σ^2, neboť testovací statistika tohoto testu Z_W je založena na postačující statistice $\bar{X}$, která je pouze přičítána/odečítána/násobena/dělena konstantami μ_0, $\sqrt{n}$ a σ.
  - Pro H_{01} rovnoměrně nejsilnější test neexistuje. Z přednášky z SI II víme, že rovnoměrně nejsilnější test neexistuje pro oboustrannou alternativu.

4 Test o střední hodnotě při známém rozptylu

- Příklad 4.1
  - (a) $X \sim N(20, 100)$, $n = 5$, či $n = 100$ . . . Po opakovaném spuštění MC studie vidíme, že aktuální pravděpodobnost pokrytí $1 - \widehat{\alpha}$ nabývá rovnoměrně častokrát vyšší i nižší hodnoty než je nominální pravděpodobnost pokrytí $1 - \alpha$. DIS pro μ, když σ^2 známe, není v tomto případě konzervativní ani liberální, stejně jako test založený na testovací statistice Z_W.
  - (b) $X \sim pN(20, 100) + (1 - p)N(28, 100)$, $n = 5$, či $n = 100$ . . . Po opakovaném spuštění MC studie vidíme, že aktuální pravděpodobnost pokrytí $1 - \widehat{\alpha}$ je soustavně nižší než nominální pravděpodobnost pokrytí $1 - \alpha$. DIS pro μ, když σ^2 známe, je v tomto případě liberální, stejně jako test založený na testovací statistice Z_W. Pro $n = 5$ není efekt liberálnosti tak zřejmý jako při $n = 100$, díky nižšímu rozsahu náhodných výběrů.
- Příklad 4.2
  - (a) $\mu = 146$ . . . Centrální rozdělení přísluší parametru $\mu_0 = 150$, proto se červená křivka centrálního rozdělení realizuje okolo hodnoty 0. Naproti tomu rozdělení testovací statistiky Z_W přísluší hodnotě $\mu = 146$ (modrý histogram superponovaný modrou křivkou). Hodnota $\mu = 146 < \mu_0 = 150$, proto parametr necentrality λ nabývá záporné hodnoty a histogram i s modrou křivkou necentrálního rozdělení se realizuje nad zápornou hodnotou (cca okolo hodnoty -3) nalevo od červené křivky centrálního rozdělení. Modrá křivka superponuje histogram přesně, protože náhodný výběr pochází z předpokládaného rozdělení $N(146, 10^2)$.
  - (b) $\mu = 155$ . . . Centrální rozdělení přísluší parametru $\mu_0 = 150$, proto se červená křivka centrálního rozdělení realizuje okolo hodnoty 0. Naproti tomu rozdělení testovací statistiky Z_W přísluší hodnotě $\mu = 155$ (modrý histogram superponovaný modrou křivkou). Hodnota $\mu = 155 > \mu_0 = 150$, proto parametr necentrality λ nabývá kladné hodnoty a histogram i s modrou křivkou necentrálního rozdělení se realizuje nad kladnou hodnotou (cca okolo hodnoty 3) napravo od červené křivky centrálního rozdělení. Modrá křivka superponuje histogram přesně, protože náhodný výběr pochází z předpokládaného rozdělení $N(155, 10^2)$.
  - (c) směs, $\mu = 146$ . . . Analogicky jako (a). Modrá křivka nesuperponuje histogram přesně, protože náhodný výběr, na jehož základě byly vypočítány testovací statistiky Z_W reprezentované histogramem, pochází ze směsi, zatímco my předpokládáme, že pochází z "nekontaminovaného"rozdělení $N(146, 10^2)$, kterému přísluší rozdělení testovací statistiky Z_W reprezentované modrou křivkou.
  - (d) směs, $\mu = 155$ . . . Analogicky jako (b). Modrá křivka nesuperponuje histogram přesně, protože náhodný výběr, na jehož základě byly vypočítány testovací statistiky Z_W reprezentované histogramem, pochází ze směsi, zatímco my předpokládáme, že pochází z "nekontaminovaného"rozdělení $N(155, 10^2)$, kterému přísluší rozdělení testovací statistiky Z_W reprezentované modrou křivkou.
- Příklad 4.3
  - (a) oboustranná alternativa: síla testu klesá s hodnotou $\mu \to \mu_0 = 150$ z obou stran. V hodnotě μ_0 je síla rovná hladině významnosti, tj. $1 - \beta = \alpha = 0.05$. Pro $n \to 0$ je pokles síly pozvolnější.
  - (b) pravostranná alternativa: síla testu klesá s $\mu \to \mu_0 = 150$ zprava. V hodnotě μ_0 je síla rovná hladině významnosti, tj. $1 - \beta = \alpha = 0.05$. Pro $n \to 0$ je pokles síly pozvolnější.
  - (c) levostranná alternativa: síla testu klesá s hodnotou $\mu \to \mu_0 = 150$ zleva. V hodnotě μ_0 je síla rovná hladině významnosti, tj. $1 - \beta = \alpha = 0.05$. Pro $n \to 0$ je pokles síly pozvolnější.
- Příklad 4.4 . . . Exaktní a empirická silofunkce jsou si tvarově velmi blízké všude s výjimkou intervalu cca $(149.2; 150.8)$. Exaktní sílu tedy můžeme aproximovat v případě, že μ je dostatečně vzdálená od μ_0, tj. cca pro $\mu \in (-\infty; 149.2\rangle \cup \langle 150.8; \infty)$.
- Příklad 4.5
  - (a) $X \sim N(\mu, 10^2)$ . . . V tomto případě je empirická silufunkce velmi blízká exaktní silufunkci, protože náhodný výběr pochází z předpokládaného rozdělení $N(\mu, \sigma^2)$.

- (b) $X \sim [pN(\mu, 10^2) + (1-p)N(\mu, 30^2)]$, kde $p = 0.9$ ... V tomto případě má empirická silofunkce mnohem pozvolnější pokles než silofunkce exaktní. Lokální minimum empirické silofunkce je větší než lokální minimum exaktní silofunkce a neodpovídá hodnotě α. Ze studie je tedy krásně patrné, jak moc může náhodný výběr pocházející ze směsi silofunkci poškodit.
- Příklad 4.6 ... Z uvedené animace krásně vidíme propojení polohy necentrálního rozdělení a hodnoty silofunkce. Pokud se s hodnotou μ blížíme k hodnotě μ_0 zleva, vidíme, že necentrální rozdělení se čím dál blíže posouvá k centrálnímu rozdělení a hodnota silofunkce klesá. V hodnotě $\mu = \mu_0 = 150$ centrální a necentrální rozdělení splývají a silofunkce dosahuje svého minima, které je rovné hladině významnosti $\alpha = 0.05$. Následně se s hodnotou μ vzdalujeme od hodnoty μ_0 směrem doprava, a vidíme, že necentrální rozdělení se čím dál více vzdaluje od centrálního rozdělení, zatímco hodnota silofunkce začíná růst k jedné.
- Příklad 4.7 ... Z grafů a souhrnné tabulky vidíme, že pro oboustrannou alternativu jsou rozsahy náhodných výběrů (pro pevně zvolené μ_0 a μ) vyšší než pro jednostranné alternativy. Pro pravostrannou alternativu (H_{12}: $\mu > \mu_0$) nemá smysl počítat rozsahy náhodných výběrů za předpokladu, že $\mu \leq \mu_0$. Pro levostrannou alternativu (H_{13}: $\mu < \mu_0$) nemá naopak smysl počítat rozsahy náhodných výběrů za předpokladu, že $\mu \geq \mu_0$. Dále, v případě, že by $\mu = \mu_0$, by minimální rozsah náhodného výběru musel být nekonečně velký (uvažujeme pouze teoreticky), a to pro všechny tři typy alternativní hypotézy.

5 Test o střední hodnotě při neznámém rozptylu

- Příklad 5.1
  - (a) $X \sim N(20, 100)$, $n = 5$ i $n = 100$ ... Po opakovaném spuštění MC studie vidíme, že aktuální pravděpodobnost pokrytí $1-\widehat{\alpha}$ nabývá rovnoměrně častokrát vyšší i nižší hodnoty než je nominální pravděpodobnost pokrytí $1-\alpha$. DIS pro μ, když σ^2 neznáme, není v tomto případě konzervativní ani liberální, stejně jako test založený na testovací statistice T_W.
  - (b) $X \sim pN(20, 100) + (1-p)N(28, 100)$, kde $p = 0.9$, $n = 5$ ... Jelikož v tomto případě je náhodný výběr tvořen směsí dvou rozdělení s různými středními hodnotami, měli bychom být na pozoru. Po opakovaném spuštění MC studie vidíme, že aktuální pravděpodobnost pokrytí $1-\widehat{\alpha}$ bývá spíše nižší než nominální pravděpodobnost pokrytí $1-\alpha$, není to však pravidlem (bývá i vyšší). DIS pro μ, když σ^2 neznáme, je liberální, stejně jako test založený na testovací statistice T_W, nicméně díky nízkému rozsahu náhodného výběru ($n = 5$) není efekt liberálnosti ze simulace příliš zřejmý, DIS se chová, jako by nebyl ani konzervativní, ani liberální, a v praxi by efekt liberálnosti neměl příliš výrazný vliv. Z grafu však vidíme, že křivka Studentova rozdělení je oproti histogramu posunutá mírně doleva.
  - (c) $X \sim pN(20, 100) + (1-p)N(28, 100)$, kde $p = 0.9$, $n = 100$ ... Po opakovaném spuštění MC studie vidíme, že aktuální pravděpodobnost pokrytí $1-\widehat{\alpha}$ je soustavně nižší než nominální pravděpodobnost pokrytí $1-\alpha$. DIS pro μ, když σ^2 neznáme, je v tomto případě liberální, stejně jako test založený na testovací statistice T_W. Z grafu je zřejmé výrazné posunutí křivky Studentova rozdělení oproti histogramu směrem doleva.
- Příklad 5.2 ... Z animací vidíme, že oba vztahy platí, tj. $T_W \sim t_{n-1}$ a $T_W^2 \sim F_{1,n-1}$, a to asymptoticky. Příklad má spojitost s příkladem 2.3 a v něm uvedeným vztahem $F_{1,n-1}(1-\alpha) = t^2_{n-1}(1-\alpha/2)$. Z tohoto vztahu totiž vyplývá, že $t^2_{n-1} = F_{1,n-1}$ a proto pokud $T_W \sim t_{n-1}$, potom $T_W^2 \sim t^2_{n-1} = F_{1,n-1}$.
- Příklad 5.3 ... Z animací vidíme, že všechny tři uvedené vztahy platí, tj. $U_W \sim \chi^2_1$, $U_S \sim \chi^2_1$, $U_{LR} \sim \chi^2_1$ asymptoticky.
- Příklad 5.4 ... Z uvedené animace je zřejmé, že pro $n \to \infty$ se testovací statistiky U_W, U_S a U_{LR} vzájemně čím dál více blíží (uspokojivá shoda nastává cca od $n = 50$). Dále vidíme, že pro velmi nízké rozsahy náhodného výběru (cca $n \leq 8$) testovací statistika U_W selhává, a tedy není na testování H_0 o μ, když σ^2 neznáme, vůbec vhodná. Z obrázku 5 je potom patrné, že platí nerovnost $U_W > U_{LR} > U_S$. Tato nerovnost vyplývá z nerovnosti $x > \log(1+x) > \frac{x}{1+x}$, kterou jsme si ověřili v příkladu 2.1.
- Příklad 5.5 ... V tomto příkladě si vzájemně porovnáváme tvar centrálního Studentova rozdělení s tvarem necentrálních rozdělení s parametrem necentrality λ. Z grafů zobrazujících tvary hustot i distribučních funkcí je patrné, že s měnící se hodnotou parametru necentrality λ se mění nejen střední hodnota, ale i rozptyl Studentova necentrálního rozdělení. Čím je parametr necentrality λ (v absolutní hodnotě) větší, tím více se necentrální a centrální rozložení liší v (a) poloze; (b) v rozptylu, a to ve smyslu: čím vyšší λ, tím vyšší rozptyl.
- Příklad 5.6 ... Příklad je analogický příkladu 4.2.
  - (a) $\mu = 130$ ... Centrální rozdělení přísluší parametru $\mu_0 = 150$, proto se modrá křivka centrálního rozdělení realizuje okolo hodnoty 0. Naproti tomu rozdělení testovací statistiky T_W přísluší hodnotě $\mu = 130$ (červený histogram superponovaný červenou křivkou). Hodnota $\mu = 130 < \mu_0 = 150$, proto parametr necentrality λ nabývá záporné hodnoty a histogram i s červenou křivkou necentrálního rozdělení se realizuje nad zápornou hodnotou (cca okolo hodnoty -5) nalevo od modré křivky centrálního rozdělení. Červená křivka superponuje histogram přesně, protože náhodný výběr pochází z předpokládaného rozdělení $N(130, 30^2)$.
  - (b) $\mu = 170$ ... Centrální rozdělení přísluší parametru $\mu_0 = 150$, proto se modrá křivka centrálního rozdělení realizuje okolo hodnoty 0. Naproti tomu rozdělení testovací statistiky T_W přísluší hodnotě $\mu = 170$ (červený histogram superponovaný červenou křivkou). Hodnota $\mu = 170 > \mu_0 = 150$, proto parametr necentrality λ nabývá kladné hodnoty a histogram i s červenou křivkou necentrálního rozdělení se realizuje nad kladnou hodnotou (cca okolo hodnoty 5) napravo od modré křivky centrálního rozdělení. Červená křivka superponuje histogram přesně, protože náhodný výběr pochází z předpokládaného rozdělení $N(170, 30^2)$.

  * (c) směs, $\mu = 130$ ... Analogicky jako (a). Červená křivka nesuperponuje histogram přesně, protože náhodný výběr, na jehož základě byly vypočítány testovací statistiky Z_W reprezentované histogramem, pochází ze směsi, zatímco my předpokládáme, že pochází z "nekontaminovaného" rozdělení $N(130, 30^2)$, kterému přísluší rozdělení testovací statistiky Z_W reprezentované červenou křivkou.
  * (d) směs, $\mu = 170$ ... Analogicky jako (b). Červená křivka nesuperponuje histogram přesně, protože náhodný výběr, na jehož základě byly vypočítány testovací statistiky Z_W reprezentované histogramem, pochází ze směsi, zatímco my předpokládáme, že pochází z "nekontaminovaného" rozdělení $N(170, 30^2)$, kterému přísluší rozdělení testovací statistiky Z_W reprezentované červenou křivkou.

- Příklad 5.7 ... Příklad je analogický příkladu 4.6
  * Z uvedené animace krásně vidíme propojení polohy necentrálního rozdělení a hodnoty silofunkce. Pokud se s hodnotou μ blížíme k hodnotě μ_0 zleva, vidíme, že necentrální rozdělení se čím dál blíže posouvá k centrálnímu rozdělení a hodnota silofunkce klesá. V hodnotě $\mu = \mu_0 = 150$ centrální a necentrální rozdělení splývají a silofunkce dosahuje svého minima, které je rovné hladině významnosti $\alpha = 0.05$. Následně se s hodnotou μ vzdalujeme od hodnoty μ_0 směrem doprava, a vidíme, že necentrální rozdělení se čím dál více vzdaluje od centrálního rozdělení, zatímco hodnota silofunkce začíná růst k jedné.
- Příklad 5.8
  * Test normality: $n = 107 > 100 \to$ Lillieforsův test; p-hodnota $= 0.0989 > \alpha = 0.05 \to H_0$ nezamítáme na hladině významnosti $\alpha = 0.05$. Naměřené hodnoty basion-bregmatické výšky lebky žen starověké egyptské populace pochází z normálního rozdělení.
  * Test o μ, když σ^2 neznáme:
    · Pomocí testovací statistiky T_W: $t_W = -2.8004, W = (-\infty; -1.9826\rangle \cup \langle 1.9826; \infty)$, $t_W \in W \to H_0$ zamítáme; $\mu_0 = 126.942, IS = (124.7904; 126.5741)$, $\mu_0 \notin IS \to H_0$ zamítáme; p-hodnota $= 0.0061$, $\alpha = 0.05$, p-hodnota $< \alpha \to H_0$ zamítáme na hladině významnosti $\alpha = 0.05$.
    · Pomocí testovací statistiky U_W: $u_W = 7.9163$, $W = \langle 3.8415; \infty)$, $u_W \in W \to H_0$ zamítáme; $\mu_0 = 126.942$, $IS = (124.8078; 126.5596)$, $\mu_0 \notin IS \to H_0$ zamítáme; p-hodnota $= 0.0049$, $\alpha = 0.05$, p-hodnota $< \alpha \to H_0$ zamítáme na hladině významnosti $\alpha = 0.05$.
    · Pomocí testovací statistiky U_S: $u_S = 7.3710$, $W = \langle 3.8415; \infty)$, $u_S \in W \to H_0$ zamítáme; $\mu_0 = 126.942$, $IS = (124.7932; 126.5742)$, $\mu_0 \notin IS \to H_0$ zamítáme; p-hodnota $= 0.0066$, $\alpha = 0.05$, p-hodnota $< \alpha \to H_0$ zamítáme na hladině významnosti $\alpha = 0.05$.
    · Pomocí testovací statistiky U_{LR}: $u_{LR} = 7.6371$, $W = \langle 3.8415; \infty)$, $u_{LR} \in W \to H_0$ zamítáme; $\mu_0 = 126.942$, $IS = (124.7981; 126.5645)$, $\mu_0 \notin IS \to H_0$ zamítáme; p-hodnota $= 0.0057$, $\alpha = 0.05$, p-hodnota $< \alpha \to H_0$ zamítáme na hladině významnosti $\alpha = 0.05$.
    · Interpretace: Mezi basion–bregmatickou výškou lebky žen starověké a novověké egyptské populace existuje statisticky významný rozdíl ($|\mu - \mu_0| = 1.2598\,\text{mm}$).

6 Test o rozptylu

- Příklad 6.1
  - (a) $X \sim N(0,1)$ ... předpoklad asymptotického χ^2_{n-1} rozdělení testovací statistiky F_W v tomto případě platí.
  - (b) $X \sim pN(0,1) + (1-p)N(0,4)$, $p = 0.9$ ... předpoklad asymptotického χ^2_{n-1} rozdělení testovací statistiky F_W v tomto případě neplatí.
  - (c) $X \sim \mathrm{Exp}(1)$, ... předpoklad asymptotického χ^2_{n-1} rozdělení testovací statistiky F_W v tomto případě neplatí.
- Příklad 6.2 ... Vztah $F_{W,n-1} \sim \chi^2_{n-1}$ platí exaktně, vztah $F_{W,n} \sim \chi^2_n$ platí pouze asymptoticky (pro $n \to \infty$). Je to jednak proto, že $F_{W,n-1} = F_{W,n}$, a jednak proto, že pro $n \to \infty$ se rozdělení χ^2_{n-1} a χ^2_n k sobě vzájemně blíží (viz následující příklad 6.3). Z rovnosti $F_{W,n-1} = F_{W,n}$ potom dále vyplývá, že $F_{W,n-1} \sim \chi^2_n$ asymptoticky a $F_{W,n} \sim \chi^2_{n-1}$ exaktně.
- Příklad 6.3 ... Z animace je zřejmé, že skutečně pro $n \to \infty$ se rozdělení χ^2_{n-1} a χ^2_n k sobě vzájemně blíží (uspokojivá shoda nastává cca pro $n \geq 100$).
- Příklad 6.4
  - T_W ... $T_W \sim N(0,1)$ asymptoticky (uspokojivá shoda nastává pro cca $n \geq 25$)
  - U_W ... $U_W \sim \chi^2_1$ asymptoticky (uspokojivá shoda nastává pro cca $n \geq 20$)
  - U_S ... $U_S \sim \chi^2_1$ asymptoticky (uspokojivá shoda nastává pro cca $n \geq 10$)
  - U_{LR} ... $U_{LR} \sim \chi^2_1$ asymptoticky (uspokojivá shoda nastává pro cca $n \geq 5$)
- Příklad 6.5 ... Z animace je zřejmé, že testovací statistiky U_W, U_S i U_{LR} konvergují pro $n \to \infty$ k χ^2_1 rozdělení. Nejrychleji konverguje testovací statistika U_{LR} (dobré výsledky již pro $n \geq 5$), následně U_S (cca pro $n \geq 10$) a nakonec U_W (cca pro $n \geq 30$). Naprosto nevhodná pro test nulové hypotézy o rozptylu při rozsahu náhodného výběru $n \leq 5$ je statistika U_W. Z obrázku 6 zobrazujícího průměrné hodnoty testovacích statistik U_W, U_S a U_{LR} je potom patrné, že pro test o rozptylu nerovnost $U_W > U_{LR} > U_S$ neplatí.
- Příklad 6.6
  - Test normality: $n = 107 > 100 \to$ Lillieforsův test; p-hodnota $= 0.0989 > \alpha = 0.05 \to H_0$ nezamítáme na hladině významnosti $\alpha = 0.05$. Naměřené hodnoty basion-bregmatické výšky lebky žen starověké egyptské populace pochází z normálního rozdělení. (viz příklad 5.8)
  - Test o rozptylu σ^2, když μ neznáme:
    - Pomocí testovací statistiky $F_{W,n-1}$: $f_{W,n-1} = 116.9533$, $W = (-\infty; 79.4013\rangle \cup \langle 136.3822; \infty)$, $f_{W,n-1} \notin W \to H_0$ nezamítáme; $\sigma_0^2 = 19.6249, IS = (16.8292; 28.9063)$, $\sigma_0^2 \in IS \to H_0$ nezamítáme; p-hodnota $= 0.4394$, $\alpha = 0.05$, p-hodnota $> \alpha \to H_0$ nezamítáme na hladině významnosti $\alpha = 0.05$.
    - Pomocí testovací statistiky U_W: $u_W = 0.3875$, $W = \langle 3.8415; \infty)$, $u_W \notin W \to H_0$ nezamítáme; $\sigma_0^2 = 19.6249$, $IS = (15.7030; 27.1973)$, $\sigma_0^2 \in IS \to H_0$ nezamítáme; p-hodnota $= 0.5336$, $\alpha = 0.05$, p-hodnota $> \alpha \to H_0$ nezamítáme na hladině významnosti $\alpha = 0.05$.
    - Pomocí testovací statistiky U_S: $u_S = 0.4629$, $W = \langle 3.8415; \infty)$, $u_S \notin W \to H_0$ nezamítáme; $\sigma_0^2 = 19.6249$, $IS = (16.9176; 29.3017)$, $\sigma_0^2 \in IS \to H_0$ nezamítáme; p-hodnota $= 0.4963$, $\alpha = 0.05$, p-hodnota $> \alpha \to H_0$ nezamítáme na hladině významnosti $\alpha = 0.05$.
    - Pomocí testovací statistiky U_{LR}: $u_{LR} = 0.4361$, $W = \langle 3.8415; \infty)$, $u_{LR} \notin W \to H_0$ nezamítáme; $\sigma_0^2 = 19.6249$, $IS = (16.5975; 28.3950)$, $\sigma_0^2 \in IS \to H_0$ nezamítáme; p-hodnota $= 0.5090$, $\alpha = 0.05$, p-hodnota $> \alpha \to H_0$ nezamítáme na hladině významnosti $\alpha = 0.05$.
    - Interpretace: Mezi rozptylem basion–bregmatické výšky lebky žen starověké a novověké egyptské populace neexistuje statisticky významný rozdíl ($|\sigma^2 - \sigma_0^2| = 2.0279\,\mathrm{mm}^2$).

7 Test o korelačním koeficientu

- Příklad 7.1 …Z uvedené animace vidíme, že Fisherova Z-transformace Z_R konverguje k normálnímu rozdělení $N\left(\frac{1}{2}\ln\frac{1+\rho}{1-\rho}, \frac{1}{n-3}\right)$ rychleji (uspokojivá shoda pro situaci, kdy $\rho = 0.8$, je cca $n \geq 5$) než výběrový korelační koeficient R k rozdělení $N\left(\rho, \frac{(1-\rho^2)^2}{n-1}\right)$ (uspokojivá shoda pro situaci, kdy $\rho = 0.8$, je cca $n \geq 35$).
- Příklad 7.2
  * pro $n = 5$ …Z uvedené animace vidíme, že Fisherova Z-transformace Z_R konverguje k normálnímu rozdělení $N\left(\frac{1}{2}\ln\frac{1+\rho}{1-\rho}, \frac{1}{n-3}\right)$ nejlépe pro ρ blízké nule. S $\rho \to 1$ se kvalita shody rozdělení Z_R a normálního rozdělení zhoršuje (situace se začíná zhoršovat cca od $\rho > 0.5$ a neuspokojivá začíná být cca od $\rho > 0.8$). V případě výběrového korelačního koeficientu R je situace neuspokojivá pro libovolné ρ.
  * pro $n = 50$ …Z uvedené animace vidíme, že pro $n = 50$ je kvalita konvergence Fisherovy Z-transformace Z_R i výběrového korelačního koeficientu R k příslušnému rozdělení uspokojivá pro libovolnou hodnotu korelačního koeficientu ρ.
- Příklad 7.3
  * Test dvourozměrné normality: Mardiův test: (a) Test o nulovém koeficientu šikmosti: p-hodnota $= 0.3853 > \alpha = 0.05 \to H_0$ nezamítáme na hladině významnosti $\alpha = 0.05$. (b) Test o nulovém koeficientu špičatosti: p-hodnota $= 0.1998 > \alpha = 0.05 \to H_0$ nezamítáme na hladině významnosti $\alpha = 0.05$. Naměřené hodnoty největší výšky mozkovny a morfologické výšky tváře mužů starověké egyptské populace nevykazují statisticky významné zešikmení ani zešpičatění, z čehož vyplývá, že naměřené hodnoty největší výšky mozkovny a morfologické výšky tváře mužů starověké egyptské populace pochází z dvourozměrného normálního rozdělení.
  * Test o korelačním koeficientu ρ:
    · $H_0 : \rho = 0.251$; $H_1 : \rho \neq 0.251$
    · Pomocí testovací statistiky Z_W: $z_W = 1.1048$, $W = (-\infty; -1.9600\rangle \cup \langle 1.9600; \infty)$, $z_W \notin W \to H_0$ nezamítáme; $\rho_0 = 0.251$, $IS = (0.1869; 0.4606)$, $\rho_0 \in IS \to H_0$ nezamítáme; p-hodnota $= 0.2692$, $\alpha = 0.05$, p-hodnota $> \alpha \to H_0$ nezamítáme na hladině významnosti $\alpha = 0.05$.
    · Pomocí testovací statistiky U_{LR}: $u_{LR} = 1.2418$, $W = \langle 3.8415; \infty)$, $u_{LR} \notin W \to H_0$ nezamítáme; $\rho_0 = 0.251$, $IS = (0.1880; 0.4596)$, $\rho_0 \in IS \to H_0$ nezamítáme; p-hodnota $= 0.2651$, $\alpha = 0.05$, p-hodnota $> \alpha \to H_0$ nezamítáme na hladině významnosti $\alpha = 0.05$.
    · Interpretace: Mezi korelačním koeficientem největší výšky mozkovny a morfologické výšky tváře mužů starověké a novověké egyptské populace neexistuje statisticky významný rozdíl. Mezi největší výškou mozkovny a morfologickou výškou tváře mužů starověké egyptské populace existuje mírný stupeň přímé lineární závislosti ($\rho = 0.3306$). Mezi největší výškou mozkovny a morfologickou výškou tváře mužů novověké egyptské populace existuje nízký stupeň přímé lineární závislosti ($\rho = 0.251$).
- Příklad 7.4 …příklad je analogický příkladu 4.7.
  * Z grafů je patrné, že pro oboustrannou alternativu jsou rozsahy náhodných výběrů (pro pevně zvolené ρ_0 a ρ) vyšší než pro jednostranné alternativy. Pro pravostrannou alternativu (H_{12}: $\rho > \rho_0$) nemá smysl počítat rozsahy náhodných výběrů za předpokladu, že $\rho \leq \rho_0$. Pro levostrannou alternativu (H_{13}: $\rho < \rho_0$) nemá naopak smysl počítat rozsahy náhodných výběrů za předpokladu, že $\rho \geq \rho_0$. Dále, v případě, že by $\rho = \rho_0$, by minimální rozsah náhodného výběru musel být nekonečně velký (uvažujeme pouze teoreticky), a to pro všechny tři typy alternativní hypotézy.
  * (a) Chyba v zadání. V situaci (a) raději uvažujme, že $H_0 : \rho \geq \rho_0$, a tedy $H_1 : \rho < \rho_0$. Za předpokladu hodnot parametrů uvedených v zadání je minálmí rozsah náhodného výběru $n_{\min} = 20$. Poznámka: Minimální rozsah náhodného výběru lze spočítat i pro stávající zadání, kde $H_0 : \rho \leq \rho_0$, nicméně z praktického hlediska toto zadání nedává smysl.
  * (b) Za předpokladu hodnot parametrů uvedených v zadání je pro $H_0 : \rho = \rho_0$ minálmí rozsah náhodného výběru $n_{\min} = 438$.

* (c) Chyba v zadání. V situaci (c) raději uvažujme, že $H_0 : \rho \leq \rho_0$, a tedy $H_1 : \rho > \rho_0$. Za předpokladu hodnot parametrů uvedených v zadání je minálmí rozsah náhodného výběru $n_{\min} = 84$. Poznámka: Minimální rozsah náhodného výběru lze spočítat i pro původní zadání, kde $H_0 : \rho \geq \rho_0$, nicméně z praktického hlediska toto zadání nedává smysl.

8 Test o pravděpodobnosti

- Příklad 8.1 ... *Poznámka:* Uspokojivá shoda rozdělení testovací statistiky Z_W se standardizovaným normálním rozdělením $N(0,1)$ je více patrná, pokud v histogramu zvolíme hrubší dělení realizací z_W do třídicích intervalů, a to s ekvidistantní šířkou $d = 1$, namísto $d = 0.5$, jak je uvedeno v materiálech. Před pročtením komentářů si proto, prosím, toto třídění ve svých kódech opravte.
  - (a) $p = 0.1$, $p = 0.9$ ... Z animace vidíme, že Haldova podmínka je dobrým ukazatelem situací, kdy je možné rozdělení testovací statistiky Z_W aproximovat normálním rozdělením $N(0,1)$. Podle této podmínky je pro dobrou aproximaci (a tedy spolehlivé testování) při volbě $p = 0.1$, resp. $p = 0.9$ potřebné mít náhodný výběr o rozsahu $N > 100$. Z animace vidíme, že pro nižší rozsahy není aproximace upokojivá. Dále vidíme, že minimální rozsahy náhodných výběrů jsou symetrické pro hodnoty parametru p okolo hodnoty $p = 0.5$ (vysvětlení, proč pro $p = 0.1$ i pro $p = 0.9$ je minimální rozsah náhodného výběru N stejný. Můžete si ověřit, že toto platí pro libovolnou dvojici p a $1 - p$.
  - (b) $p = 0.5$ ... Z animace vidíme, že Haldova podmínka je dobrým ukazatelem situací, kdy je možné rozdělení testovací statistiky Z_W aproximovat normálním rozdělením $N(0,1)$. Podle této podmínky je pro dobrou aproximaci (a tedy spolehlivé testování) při volbě $p = 0.5$ potřebné mít náhodný výběr o rozsahu $N > 36$ (nutno z Haldovy podmínky dopočítat). Pro nižší rozsahy není aproximace dostatečně kvalitní, jak je vidět z animace. Pro $p = 0.5$, potřebujeme nejmenší minimální rozsah náhodného výběru. Čím více se s p blížíme do 0, či 1, tím vyšší minimální rozsah náhodného výběru potřebujeme, aby byla Haldova podmínka dobré aproximace splněna a aproximace rozdělení testovací statistiky Z_W normálním rozdělením byla dostatečně kvalitní.

- Příklad 8.2 ... Příklady 8.2, 8.3 a 8.4 ukazují jediný správný výpočet aktuální pravděpodobnosti pokrytí pro $(1-\alpha)\%$ empirický IS pro parametr p. Konkrétně příklad 8.2 nám ukazuje úskalí aktuální pravděpodobnosti pokrytí 95% Waldova empirického DIS. Pokud náhodná veličina $X \sim \text{Bin}(30, 0.8)$ a při experimentu dojde k zaznamenání 24 úspěchů ze 30, je vše v pořádku a aktuální pravděpodobnost pokrytí $(1-\widehat{\alpha}) = 0.9463 \doteq 0.95$. Nicméně, pokud náhodná veličina $X \sim \text{Bin}(30, 0.79)$ a při experimentu dojde k zanamenání 24 úspěchů ze 30 (což by reálně s vysokou pravděpodobností skutečně nastalo), klesne aktuální pravděpodobnost pokrytí $(1-\widehat{\alpha})$ na hodnotu $0.8876 \doteq 0.89$, tj. aktuální pravděpodobnost pokrytí je o 0.06 nižší než nominální pravděpodobnost pokrytí, což v pořádku není.

- Příklad 8.3 ... v tomto příkladu si vytváříme náhled na aktuální pravděpodobnost pokrytí 95% Waldova, skóre a věrohodnostního empirického DIS pro parametr p a pro různé rozsahy náhodného výběru N. Díky tomuto náhledu můžeme snadno zhodnotit, které DIS jsou pro určitá p a N naprosto nevhodná, jelikož jejich aktuální pravděpodbnost prokrytí se od nominální pravděpodobnosti pokrytí propastně liší.
  - Waldův 95% DIS ... pro $N = 5$ DIS ve většině případů z hlediska pravděpodobnosti pokrytí naprosto selhává, aktuální pst pokrytí je nižší než nominální pst pokrytí a DIS je silně liberální. Pro rostoucí N se situace lepší, DIS nicméně povětšinou zůstává liberální, navíc výrazně selhává při nízkých hodnotách parametru p, zejména pro $p < 0.1$ nebo $p > 0.9$.
  - Skóre 95% DIS ... pro $N = 5$ dává DIS dobré výsledky (není ani konzervativní ani liberální) pro $p \in (0.2; 0.8)$. Pro cca $p < 0.2$ a $p > 0.8$ je liberální. S rostoucím N však jeho kvalita velmi rychle roste, pro $N = 100$ funguje skvěle pro cca $p \in (0.05; 0.95)$, pro $N = 500$ funguje skvěle de facto na celém intervalu $(0; 1)$. Ze všech DIS je z hlediska pravděpodobnosti pokrytí suverénně nejlepší.
  - Věrohodnostní 95% DIS ... pro $N = 5$ dává DIS dobré výsledky (není ani konzervativní ani liberální) pro cca $p \in (0.2; 0.8)$. Pro $p < 0.2$ a $p > 0.8$ je liberální. S rostoucím N však jeho kvalita velmi rychle roste, pro $N = 100$ funguje skvěle pro cca $p \in (0.1; 0.9)$, pro $N = 500$ funguje skvěle pro cca $p \in (0.02; 0.98)$.

- Příklad 8.4 ... v tomto příkladu si vytváříme náhled na aktuální pravděpodobnost pokrytí zpětně transformovaných 95% empirických DIS pro parametr p a pro různé rozsahy náhodného výběru N.
  - 95% zpětně transformovaný DIS s transformací $p/(1-p)$ ... pro $N = 5$ DIS nevychází vůbec dobře, na celém intervalu (0; 1) je liberální a od hodnoty $p > 0.8$ naprosto selhává; pro $N = 100$ se situace lepší, nicméně DIS je stále liberální a selhává pro $p > 0.9$. S rostoucím N se jeho vlastnosti zlepšují, nicméně v porovnání s následujícími alternativami i s alternativami z příkladu 8.3 vychází suverénně nejhůře.

  - 95% zpětně transformovaný DIS s transformací $\ln(p/(1-p))$ ... pro $N = 5$ DIS vychází dobře (není ani konzervativní ani liberální); pro $p \in (0.2; 0.8)$. Pro $p < 0.2$ nebo $p > 0.8$ je liberální, pro $p < 0.1$, či $p > 0.9$ selhává. S rostoucím N se však jeho kvalita rapidně zlepšuje a s výjimkou intervalu cca $(0; 0.05)$ a $(0.95; 1)$ není ani konzervativní ani liberální. V úzkém okolí hodnot 0 a 1 je liberální, v intervalech $(0; 0.05)$, $(0.95; 1)$ je spíše konzervativní.
  - 95% zpětně transformovaný DIS s transformací $\arcsin(\sqrt{p})$ ... pro $N = 5$ vychází dobře (není ani konzervativní ani liberální) pro $p \in (0.15; 0.85)$. Pro cca $p < 0.15$ nebo $p > 0.85$ je liberální a následně selhává. Stejné vlastnosti si drží i pro větší rozsahy náhodných výběrů; pro $N = 100$ je liberální pro cca $p < 0.1$ a $p > 0.9$; pro $N = 500$ je liberální pro cca $p < 0.05$ a $p > 0.95$, atd. Jeho nespornou výhodou navíc je, že rozptyl transformace je pevný (neměnný), a tedy není závislý na odhadu parametru p.

- Seřazení 95% empirických DIS pro parametr p podle kvality (z hlediska pravděpodobnosti pokrytí) od nejlepšího po nejhorší: (1) skóre, (2–3) věrohodnostní, (2–3) transformace $\arcsin(\sqrt{p})$, (4) transformace $\ln\left(\frac{p}{1-p}\right)$, (5) Waldův, (6) transformace $\frac{p}{1-p}$. Nejvhodnější je na analýzy zvolit skóre DIS a tedy také skóre test.

9 Test o rozdílu středních hodnot

- Příklad 9.1
  - (a) $X_1 \sim N(4, 2^2)$, $X_2 \sim N(2.5, 2^2)$ ... Centrální rozdělení přísluší parametru $\mu_0 = 0$, proto se červená křivka centrálního rozdělení realizuje okolo hodnoty 0. Naproti tomu rozdělení testovací statistiky T_W (zelený histogram superponovaný zelenou křivkou) přísluší hodnotě rozdílu $\mu_1 - \mu_2$, který pro $\mu_1 = 4$ a $\mu_2 > 4$ nabývá záporných hodnot, pro $\mu_2 = 4$ nabývá nuly a pro $\mu_2 < 4$ nabývá kladných hodnot. Zelený histogram i se zelenou křivkou necentrálního rozdělení se tedy nejprve realizuje nad zápornými hodnotami a s klesající hodnotou parametru μ_2 se postupně posouvá přes nulu do kladných čísel. Zelená křivka superponuje histogram přesně, protože oba náhodné výběry X_1, resp. X_2 pochází z předpokládaných rozdělení $N(4, 2^2)$, resp. $N(2.5, 2^2)$.
  - (b) $X_1 \sim$ směs, $X_2 \sim N(2.5, 2^2)$ ... Analogicky jako (a). Zelená křivka nesuperponuje histogram přesně, protože náhodný výběr X_1 pochází ze směsi, a tedy testovací statistiky T_W reprezentované histogramem jsou touto směsí kontaminovány. Rozdělení testovací statistiky T_W reprezentované zelenou křivkou nicméně odpovídá předpokladu, že X_1, resp. X_2 pochází z "nekontaminovaného"normálního rozdělení $N(4, 2^2)$, resp. $N(2.5, 2^2)$.
- Příklad 9.2
  - (1–a, b, c, d) $n = 5$, $n = 50$, $n = 100$ ... ani konzervativní ani liberální
  - (2–a, b, c, d) $n = 5$ ... konzervativní
  - (2–a, b, c, d) $n = 50$, $n = 100$ ... ani konzervativní ani liberální
  - (3–a, b, c, d), (4–a, b, c, d), $n = 5$ ... liberální
  - (3–a, b, c, d), (4–a, b, c, d), $n = 50$ ... mírně liberální
  - (3–a, b, c, d), (4–a, b, c, d), $n = 100$ ... ani konzervativní, ani liberální
  - Z hlediska pravděpodobnosti pokrytí vychází pro všechny čtyři situace (a)–(d) nejlépe Waldovy DIS pro shodné rozptyly obou náhodných výběrů. Tyto DIS nejsou konzervativní ani liberální již pro nízké rozsahy ($n \geq 5$) náhodných výběrů. Pro vyšší rozsahy ($n \geq 50$) vychází dobře také Waldovy DIS pro různé rozptyly, které jsou pro nižší rozsahy $n < 50$ konzervativní. Věrohodnostní DIS vychází z hlediska pravděpodobnosti pokrytí dobře pro náhodné výběry s rozsahy $n \geq 100$. Pro $n \leq 100$ jsou věrohodnostní DIS mírně liberální, pro nízké rozsahy ($n \leq 5$) jsou nevhodné, jelikož jsou velmi liberální. Všimněme si, že z hlediska pravděpodobnosti pokrytí není příliš výrazný rozdíl mezi intervaly spolehlivosti sestrojenými za předpokladu shodných rozptylů a za předpokladu různých rozptylů. Závěrem poznamenejme, že pravděpodobnost prokrytí je pouze jednou z vlastností intervalu spolehlivosti, kterou bychom před jeho použitím v praxi měli znát. Finální volba intervalu spolehlivosti by však měla záviset i na dalších aspektech (například, pokud DIS používáme k testování, záleží také na síle testu, který DIS reprezentuje :)).
- Příklad 9.3
  - Test normality muži: $n = 50 \rightarrow$ Andersonův-Darlingův test; p-hodnota $= 0.4316 > \alpha = 0.05 \rightarrow H_0$ nezamítáme na hladině významnosti $\alpha = 0.05$. Naměřené hodnoty délky klíční kosti z pravé strany u mužů pochází z normálního rozdělení.
  - Test normality ženy: $n = 50 \rightarrow$ Andersonův-Darlingův test; p-hodnota $= 0.0191 < \alpha = 0.05 \rightarrow H_0$ zamítáme na hladině významnosti $\alpha = 0.05$. Naměřené hodnoty největší výšky mozkovny žen starověké egyptské populace nepochází z normálního rozdělení. $\rightarrow$ Nesprávně zvolená data na tento typ příkladu z mé strany. Správný postup by nyní byl použít na ověření zadané hypotézy neparametrický test. Pro účely procvičení se však pro tentokrát tvařme, že normalita náhodného výběru pro ženy je též v pořádku.
  - Test o shodě rozptylů σ_1^2/σ_2^2:
    * Jelikož různé testy o podílu rozptylů probíráme až v následující sekci, otestujeme hypotézu o shodě rozptylů pouze pomocí testu, který známe z bakalářského studia, tj. pomocí F-testu, který provedeme příkazem `var.test()`, a to pouze na základě p-hodnoty: p-hodnota $= 0.0981 > \alpha = 0.05$, tj. H_0 nezamítáme na hladině významnosti $\alpha = 0.05$. Mezi rozptylem délky klíční kosti z pravé strany u mužů a žen neexistuje statisticky významný rozdíl (předpokládáme, že rozptyly jsou shodné).

- Test o rozdílu středních hodnot:
  - Pomocí testovací statistiky T_W: $t_W = 7.1019, W = (-\infty; -1.9845\rangle \cup \langle 1.9845; \infty)$, $t_W \in W \rightarrow H_0$ zamítáme; $\mu_0 = 0, IS = (10.1313; 17.9887)$, $\mu_0 \notin IS \rightarrow H_0$ zamítáme; p-hodnota < 0.0001, $\alpha = 0.05$, p-hodnota $< \alpha \rightarrow H_0$ zamítáme na hladině významnosti $\alpha = 0.05$.
  - Pomocí testovací statistiky T_W s Welchovou aprox. st. volnosti: $t_W = 7.1019$, $W = (-\infty; -1.9858\rangle \cup \langle 1.9858; \infty)$, $t_W \notin W \rightarrow H_0$ zamítáme; $\mu_0 = 0, IS = (10.1286; 17.9914)$, $\mu_0 \notin IS \rightarrow H_0$ zamítáme; p-hodnota < 0.0001, $\alpha = 0.05$, p-hodnota $< \alpha \rightarrow H_0$ zamítáme na hladině významnosti $\alpha = 0.05$.
  - Pomocí testovací statistiky U_{LR}, $\sigma_1^2 = \sigma_2^2$: $u_{LR} = 41.5196$, $W = \langle 3.8415; \infty)$, $u_{LR} \in W \rightarrow H_0$ zamítáme; $\mu_0 = 0$, $IS = (10.1816; 17.9384)$, $\mu_0 \notin IS \rightarrow H_0$ zamítáme; p-hodnota < 0.0001, $\alpha = 0.05$, p-hodnota $< \alpha \rightarrow H_0$ zamítáme na hladině významnosti $\alpha = 0.05$.
  - Pomocí testovací statistiky U_{LR}, $\sigma_1^2 \neq \sigma_2^2$: $u_{LR} = 51.4667$, $W = \langle 3.8415; \infty)$, $u_{LR} \in W \rightarrow H_0$ zamítáme; $\mu_0 = 0$, $IS = (10.2188; 17.9012)$, $\mu_0 \notin IS \rightarrow H_0$ zamítáme; p-hodnota < 0.0001, $\alpha = 0.05$, p-hodnota $< \alpha \rightarrow H_0$ zamítáme na hladině významnosti $\alpha = 0.05$.
  - Interpretace: Mezi délku klíční kosti z pravé strany u mužů a žen existuje statisticky významný rozdíl. ($|\mu_1 - \mu_2| = |151.74 - 137.68| = 14.06\,\text{mm}$).
  - *Poznámka:* Vzhledem k výsledku testu o podílu rozptylů bychom v praxi pro otestování H_0 o rozdílu středních hodnot použili buď (a) klasický t-test, nebo (c) věrohodnostní test, když $\sigma_1^2 = \sigma_2^2$.

10 Test o podílu rozptylů

- Příklad 10.1 (a) . . . Podíl rozptylů největší výšky mozkovny mužů a žen
  - Test normality muži: $n = 164 > 100 \rightarrow$ Lillieforsův test; p-hodnota $= 0.0710 > \alpha = 0.05 \rightarrow H_0$ nezamítáme na hladině významnosti $\alpha = 0.05$. Naměřené hodnoty největší výšky mozkovny mužů starověké egyptské populace pochází z normálního rozdělení.
  - Test normality ženy: $n = 78$; $50 < n < 100 \rightarrow$ Andersonův-Darlingův test; p-hodnota $= 0.4501 > \alpha = 0.05 \rightarrow H_0$ nezamítáme na hladině významnosti $\alpha = 0.05$. Naměřené hodnoty největší výšky mozkovny žen starověké egyptské populace pochází z normálního rozdělení.
  - Test o podílu rozptylů σ_1^2/σ_2^2:
    - Pomocí testovací statistiky F_W: $f_W = 0.9005, W = (-\infty; 0.6893\rangle \cup \langle 1.4904; \infty)$, $f_W \notin W \rightarrow H_0$ nezamítáme; $\sigma_0^2 = 1, IS = (0.6042; 1.3065)$, $\sigma_0^2 \in IS \rightarrow H_0$ nezamítáme; p-hodnota $= 0.5747$, $\alpha = 0.05$, p-hodnota $> \alpha \rightarrow H_0$ nezamítáme na hladině významnosti $\alpha = 0.05$.
    - Pomocí testovací statistiky U_{LR}: $u_{LR} = 0.2936$, $W = \langle 3.8415; \infty)$, $u_{LR} \notin W \rightarrow H_0$ nezamítáme; $\sigma_0^2 = 1$, $IS = (0.6091; 1.3087)$, $\sigma_0^2 \in IS \rightarrow H_0$ nezamítáme; p-hodnota $= 0.5879$, $\alpha = 0.05$, p-hodnota $> \alpha \rightarrow H_0$ nezamítáme na hladině významnosti $\alpha = 0.05$.
    - Interpretace: Mezi rozptylem největší výšky mozkovny mužů a žen starověké egyptské populace neexistuje statisticky významný rozdíl. ($|\sigma_1^2 - \sigma_2^2| = |21.1387 - 23.4734| = 2.3347\,\mathrm{mm}^2$).

  Poznámka: Graf zobrazený na obrázku 1 vlevo ukazuje, že hodnota parametru $\sigma_0^2 = 1$ skutečně náleží do 95% věrohodnostního empirického intervalu spolehlivosti.
- Příklad 10.1 (b) . . . Podíl rozptylů morfologické výšky tváře mužů a žen
  - Test normality muži: $n = 164 > 100 \rightarrow$ Lillieforsův test; p-hodnota $= 0.0450 < \alpha = 0.05 \rightarrow H_0$ zamítáme na hladině významnosti $\alpha = 0.05$. Při pohledu na p-hodnotu však vidíme, že k zamítnutí H_0 o normalitě došlo na hladině významnosti $\alpha = 0.05$ hraničně. Pokud bychom chtěli být naprosto korektní, museli bychom H_0 o normalitě zamítnout, udělat závěr, že data nepochází z normálního rozdělení a H_0 o podílu rozptylů otestovat vhodným neparametrickým testem. Nicméně víme, že parametrické testy jsou mnohem silnější, než testy neparametrické, proto v praxi při hraničním zamítnutí H_0 přivíráme oči a i přes hraniční zamítnutí H_0 o normalitě považujeme naměřené hodnoty za normálně rozložené. Tak tomu bude i v tomto příkladu. V reálné situaci by však záleželo pouze na vás, zda byste se striktně řídili výsledkem testu normality, či zda byste také přivřeli oči :). *Poznámka: Na rozdíl od příkladu 9.3 zde byla data s hraniční normalitou použita záměrně, abyste si vyzkoušeli řešení při sporné a hraniční situaci.*
  - Test normality ženy: $n = 78$; $50 < n < 100 \rightarrow$ Andersonův-Darlingův test; p-hodnota $= 0.4053 > \alpha = 0.05 \rightarrow H_0$ nezamítáme na hladině významnosti $\alpha = 0.05$. Naměřené hodnoty morfologické výšky tváře žen starověké egyptské populace pochází z normálního rozdělení.
  - Test o podílu rozptylů σ_1^2/σ_2^2:
    - Pomocí testovací statistiky F_W: $f_W = 1.5207, W = (-\infty; 0.6893\rangle \cup \langle 1.4904; \infty)$, $f_W \in W \rightarrow H_0$ zamítáme; $\sigma_0^2 = 1, IS = (1.0203; 2.2063)$, $\sigma_0^2 \notin IS \rightarrow H_0$ zamítáme; p-hodnota $= 0.0397$, $\alpha = 0.05$, p-hodnota $< \alpha \rightarrow H_0$ zamítáme na hladině významnosti $\alpha = 0.05$.
    - Pomocí testovací statistiky U_{LR}: $u_{LR} = 4.3956$, $W = \langle 3.8415; \infty)$, $u_{LR} \in W \rightarrow H_0$ zamítáme; $\sigma_0^2 = 1$, $IS = (1.0285; 2.2100)$, $\sigma_0^2 \notin IS \rightarrow H_0$ zamítáme; p-hodnota $= 0.0360$, $\alpha = 0.05$, p-hodnota $< \alpha \rightarrow H_0$ zamítáme na hladině významnosti $\alpha = 0.05$.
    - Interpretace: Mezi rozptylem morfologické výšky tváře mužů a žen starověké egyptské populace existuje statisticky významný rozdíl. ($|\sigma_1^2 - \sigma_2^2| = |51.4368 - 33.8242| = 17.6126\,\mathrm{mm}^2$).
      Poznámka: Jelikož v tomto případě došlo k zamítnutí H_0 o podílu rozptylů opět více méně hraničně, bylo by vhodné provést také neparametrický test o shodě rozptylů obou náhodných výběrů a prověřit, zda by závěrem neparametrického testu bylo rovněž zamítnutí H_0 o shodě rozptylů.
- Graf zobrazený na obrázku 1 vpravo ukazuje, že hodnota parametru $\sigma_0^2 = 1$ skutečně nenáleží do 95% věrohodnostního empirického intervalu spolehlivosti.

11 Test o parametru λ Poissonova rozdělení

- Příklad 11.1
  * Pomocí testovací statistiky Z_W: $z_W = 0.1811, W = (-\infty; -1.9600\rangle \cup \langle 1.9600; \infty)$, $z_W \notin W \to H_0$ nezamítáme; $\lambda_0 = 0.6, IS = (0.5018; 0.7182)$, $\lambda_0 \in IS \to H_0$ nezamítáme; p-hodnota $= 0.8563$, $\alpha = 0.05$, p-hodnota $> \alpha \to H_0$ nezamítáme na hladině významnosti $\alpha = 0.05$.
  * Pomocí testovací statistiky U_S: $u_S = 0.0333, W = \langle 3.8415; \infty)$, $u_S \notin W \to H_0$ nezamítáme; $\lambda_0 = 0.6$, $IS = (0.5109; 0.7283)$, $\lambda_0 \in IS \to H_0$ nezamítáme; p-hodnota $= 0.8551$, $\alpha = 0.05$, p-hodnota $> \alpha$ $\to H_0$ nezamítáme na hladině významnosti $\alpha = 0.05$.
  * Pomocí testovací statistiky U_{LR}: $u_{LR} = 0.0331$, $W = \langle 3.8415; \infty)$, $u_{LR} \notin W \to H_0$ nezamítáme; $\lambda_0 = 0.6$, $IS = (0.5081; 0.7247)$, $\lambda_0 \in IS \to H_0$ nezamítáme; p-hodnota $= 0.8555$, $\alpha = 0.05$, p-hodnota $> \alpha \to H_0$ nezamítáme na hladině významnosti $\alpha = 0.05$.
  * Interpretace: Střední počet úmrtí v důsledku kopnutí koněm v Pruských armádních jednotkách během jednoho roku není statisticky významně odlišný od hodnoty 0.6. (Náhodná veličina X popisující počet úmrtí v důsledku kopnutí koněm v Pruských armádních jednotkách pochází z Poissonova rozdělení (viz Statistická inference I) s parametrem λ, který není statisticky významně odlišný od hodnoty 0.6.)