10 Neparametrické testy o mediánech

- Předpoklady k použití parametrických testů:
  - normalita dat
  - homogenita rozptylů
- Závažné porušení předpokladů → **Neparametrické testy**
  - předpoklad pouze o spojitém rozdělení dat
  - slabší než parametrické testy → nepravdivou hypotézu zamítají s menší pstí
  - **pořadové testy** ... stanovíme pořadí dat a s tímto pořadím dále pracujeme
  - střední hodnotu nahradíme mediánem → **hypotézy o mediánech**.

10.1 Jednovýběrové testy

- Nechť $X_1, \ldots, X_n$ je náhodný výběr ze spojitého rozdělení s mediánem $x_{0.50}$ a c je konstanta.
- $H_0 : x_{0.50} = c$
- $H_1 : x_{0.50} \neq c$ ($H_{12} : x_{0.05} < c$, nebo $H_{13} : x_{0.05} > c$)

Znaménkový test

- SIGN.test(x, md=c, alternative='two.sided'), knihovna PASWR
- argument alternative může nabývat variant 'two.sided' (H_{11}), 'less' (H_{12}), 'greater' (H_{13})
- výstupem: statistika S^+, IS, p-hodnota
- kritický obor $W = \langle 0, k_1 \rangle \cup \langle k_2, n \rangle$
- n je počet nenulových hodnot $x - c$; k_1 a k_2 → **statistické tabulky** pro n

Wilcoxonův test

- wilcox.test(x, mu=c, alternative='two.sided', correct=F, exact=F)
- argument alternative může nabývat variant 'two.sided' (H_{11}), 'less' (H_{12}), 'greater' (H_{13})
- výstupem: statistika S^+, IS, p-hodnota
- kritický obor $(-\infty; k\rangle$
- n je počet nenulových hodnot $x - c$; k → **statistické tabulky** pro n

10.2 Párové testy

- $(X_1, Y_1), \ldots (X_n, Y_n)$ je náhodný výběr z dvourozměrného spojitého rozdělení
- vytvoříme rozdíly $Z_1 = X_1 - Y_1, \ldots, Z_n = X_n - Y_n$
- nový náhodný výběr $Z_1, \ldots, Z_n$ je ze spojitého rozdělení s mediánem $z_{0.50}$ a c je konstanta.
- $H_0 : z_{0.05} = 0$
- $H_1 : z_{0.05} \neq 0$ ($H_{12} : x_{0.05} < 0$, nebo $H_{13} : x_{0.05} > 0$)

Znaménkový test

- SIGN.test(z, md=0, alternative='two.sided'), knihovna PASWR
- argument alternative může nabývat variant 'two.sided', 'less', 'greater', podle tvaru H_1
- výstupem: statistika S^+, IS, p-hodnota
- kritický obor $W = \langle 0, k_1 \rangle \cup \langle k_2, n \rangle$
- n je počet nenulových hodnot z (resp. $x - y$); k_1 a k_2 → **statistické tabulky** pro n

Wilcoxonův test

- wilcox.test(z, mu=0, alternative='two.sided', correct=F, exact=F)
- argument alternative může nabývat variant 'two.sided', 'less', 'greater', podle tvaru H_1
- výstupem: statistika S^+, IS, p-hodnota
- kritický obor $(-\infty; k\rangle$
- n je počet nenulových hodnot z (resp. $x - y$); $k \to$ **statistické tabulky** pro n

10.3 Dvouvýběrové testy

- Nechť $X_1, \dots X_n$ a $Y_1, \dots Y_m$ jsou dva nezávislé náhodné výběry ze dvou spojitých rozdělení.
- $x_{0.50}$ ... medián prvního rozdělení; $y_{0.5}$ ... medián druhého rozdělení
- n je rozsah prvního výběru, m je rozsah druhého výběru
- $H_0 : x_{0.5} - y_{0.5} = 0$
- $H_{11} : x_{0.5} - y_{0.5} \neq 0$ ($H_{12} : x_{0.5} - y_{0.5} < 0$; $H_{13} : x_{0.5} - y_{0.5} > 0$)

Wilcoxonův test

- wilcox.test(x, y, alternative = 'two.sided', correct=F, exact=F)
- výstupem: statistika S^+, IS, p-hodnota
- kritický obor $(-\infty; k\rangle$; $k \to$ **statistické tabulky** pro n a m

10.4 Vícevýběrové testy

- Nechť máme $r \geq 2$ nezávislých náhodných výběrů, každý ze spojitého rozdělení
- H_0 : *Všechny výběry pochází z téhož rozdělení*
- H_1 : *Alespoň jeden výběr pochází z jiného rozdělení*

Kruskal-Wallisův test

- kruskal.test(x, g)
- g ... vektor skupin, příslušných danému pozorování

10.5 Metody mnohonásobného porovnávání

- Zamítneme-li H_0, že všechny výběry jsou z téhož rozdělení, zajímá nás, která dvojice výběrů se od sebe významně liší
- H_0 : k-tý a l-tý výběr pochází z téhož rozdělení
- H_1 : k-tý a l-tý výběr nepochází z téhož rozdělení

Neményiova metoda

- posthoc.kruskal.nemenyi.test(x, group, method='Chisquare') knihovna PMCMR