

M9750 Robustní a neparametrické statistické metody

cvičení 5 - robustní odhady parametru polohy, Jamesův - Steinův odhad

1. Pro stejná data jako z předchozího cvičení (náhodný výběr z normálního rozdělení $N(0, 1)$ o rozsahu $n = 30$) uvažujte následující R-odhady: Hodgesův-Lehmannův odhad se skóry $a_n(i) = i$, medián se skóry $a_n(i) = 1$ a van der Waerdenův odhad se skóry $a_n(i) = \Phi^{-1}\left(\frac{i}{n+1}\right)$.
 - (a) Nakreslete grafy funkcí $S_n(t)$ pro všechny odhady.
 - (b) Najděte příslušné odhady a porovnejte je.
 - (c) Odhady z bodu (b) porovnejte s jejich teoretickými formullemi.
2. Zkoumejte chování následujících odhadů: průměru, mediánu, Huberova odhadu, 5% a 10% useknutého a winsorizovaného průměru a Hodgesova-Lehmannova odhadu pro různá rozdělení na základě náhodného výběru o rozsahu $n = 30$. Uvažujte následující rozdělení (vždy s nulovou střední hodnotou a jednotkovým rozptylem; POZOR na parametrizaci v R): normální, logistické, t-rozdělení s 3 stupni volnosti a Cauchyho (s parametry 0,1).
 - (a) Nejprve vygenerujte náhodný výběr o rozsahu n z daného rozdělení a proveďte dané odhady. Celý postup opakujte 10 000 krát.
 - (b) Odhadněte střední hodnotu, rozptyl a střední čtvercovou chybu všech odhadů a výsledky porovnejte.
3. Zkoumejte chování vícerozměrného výběrového průměru a Jamesova - Steinova odhadu na základě náhodného výběru z p -rozměrného normálního rozdělení o rozsahu $n = 30$ se střední hodnotou $\boldsymbol{\theta} = (1, \dots, 1)'$ a varianční maticí $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}_p$. Pro jednodušost volte $p = 3$ a $\sigma^2 = 1$.
 - (a) Pro daný náhodný výběr spočítejte oba příslušné odhady a spočítejte jejich empirickou střední čtvercovou chybu.
 - (b) Celý postup opakujte 10 000 krát. Na základě těchto simulací odhadněte střední čtvercovou chybu obou odhadů (průměr empirických středních čtvercových chyb).
 - (c) Výsledky z bodu (b) porovnejte. Výsledek porovnejte i s teoretickou hodnotou střední čtvercové chyby pro výběrový průměr.

Aby všem vycházely stejné výsledky, nastavte před každým během simulace generátor náhodných čísel pomocí příkazu `set.seed(1234)`.

Funkce, které by se mohly hodit: `apply`, `hubers`, `mvrnorm` z knihovny MASS, `winsor.mean` z knihovny psych a `hl.loc` z knihovny ICSNP.