

M7988 Modely ztrát v neživotním pojištění

cvičení 7 - Model selection

V datovém souboru `insurance.RData` jsou uloženy výše pojistných plnění 227 klientů jedné pojišťovny (proměnná `data`). Naším úkolem bude porovnat dva modely: exponenciální a lognormální s parametry odhadnutými metodou maximální věrohodnosti. MLE parametru λ je $\hat{\lambda} = \frac{1}{\bar{X}}$ a pro lognormální rozdělení $\hat{\mu} = \frac{1}{n} \sum_{i=1}^n \log(X_i)$ a

$$\hat{\sigma} = \left\{ \frac{1}{n} \sum_{i=1}^n (\log(X_i) - \hat{\mu})^2 \right\}^{1/2}$$

1. Posuzujte vhodnost obou modelů pomocí grafických metod.
 - (a) Do jednoho grafu nakreslete histogram a odhadnuté hustoty obou modelů.
 - (b) Do jednoho grafu nakreslete jádrový odhad hustoty a odhadnuté hustoty obou modelů.
 - (c) Do jednoho grafu nakreslete empirickou distribuční funkci a odhadnuté distribuční funkce obou modelů.
 - (d) Do jednoho grafu nakreslete odchylku empirické distribuční funkce od distribuční funkce odhadnutých modelů.
 - (e) Nakreslete Q-Q plot a P-P plot pro oba modely.
2. Posuzujte vhodnost obou modelů pomocí statistických testů.
 - (a) Proveďte Kolmogorovův - Smirnovův test. Příslušnou p -hodnotu určete pomocí parametrického bootstrapu.
 - (b) Proveďte Pearsonův χ^2 test.
3. Který z obou modelů je vhodnější?
 - (a) Pro oba modely spočítejte hodnotu logaritmické věrohodnosti.
 - (b) Pro oba modely spočítejte hodnotu AIC.
 - (c) Pro oba modely spočítejte hodnotu BIC.

Funkce, které by se mohly hodit: `fitdistr` z knihovny `MASS` a funkce `chisq.test`, `ecdf` a `ks.test`.