

12. cvičení

Úloha 1

Máme náhodný výběr $X_1, \dots, X_n \sim \text{Bern}(\theta)$, kde $\theta \in [0, 1]$ je neznámý parametr. Hodnota $X_i = 1$ znamená, že v i -tém hodu padla panna. MLE odhad je $\hat{\theta}_{MLE} = 1/n \cdot \sum_{i=1}^n X_i$.

- Zopakujte si, jak k MLE odhadu dojít a ověřte, že tak skutečně vyjde. Může se vám hodit vyjádření $P(X_i = x_i) = \theta^{x_i}(1 - \theta)^{1-x_i}$ a pro méně psaní a jednodušší výpočty označení $Y = \sum_{i=1}^n X_i$ v průběhu výpočtu.
- Je $\hat{\theta}_{MLE}$ nestranný?
- Najděte alespoň dva různé nestranné estimátory parametru θ . Spočtete a porovnejte jejich varianci.

Úloha 2

V přednášce jsme odvodili, že lineární regrese metodou nejmenších čtverců (least squares) je MLE, pokud předpokládáme, že šum ε_i v modelu $Y_i = ax_i + b + \varepsilon_i$ je gaussovský. Předpokládejme teď místo toho, že šumy ε_i mají Laplaceovu hustotu $f(\varepsilon) = \frac{1}{2}e^{-|\varepsilon|}$. Ukažte, že MLE odhad (\hat{a}, \hat{b}) minimalizuje $\sum_{i=1}^n |y_i - ax_i - b|$. Jinými slovy: z čtverců reziduí se stanou absolutní hodnoty reziduí. (Vzorečku se říká *L1 regrese*.)

Úloha 3

Máme náhodný výběr $X_1, \dots, X_n \sim N(\mu, 1)$, kde μ je neznámé. Klasický MLE estimátor je $\hat{\mu}_{MLE} = \frac{1}{n} \sum_{i=1}^n X_i$. Představme si ale, že z nějakého důvodu věříme, že skutečné μ je malé, a tak chceme odhad trochu „stáhnout“ k nule. Uvažujme proto estimátor $\hat{\mu}' = \frac{1}{2} \hat{\mu}_{MLE}$.

Pro oba estimátory spočtete jejich bias, varianci, a MSE. Zkuste je porovnat.

Tahák

- Zkoumáme posloupnost n.n.v. se stejným rozdělením, např. $\text{Geom}(\theta)$, $U(0, \theta)$, kde θ je parametr.
- Zapisujeme $X_1, \dots, X_n \sim F_\theta$, tzv. náhodný výběr z F_θ (model s parametrem).
- Naměříme $X_1 = x_1, \dots$, chceme odhadnout θ .
- $\hat{\theta}$... nějaká metoda jak odhadnout θ pomocí naměřených dat (hodnot X_1, \dots, X_n). Angl. *estimator* – jeden získaný odhad je *estimate*, ten značíme $\hat{\theta}$.
- $L(\theta; x_1, \dots, x_n) = P[X_1 = x_1 \wedge \dots \wedge X_n = x_n]$... pravd. pozorovaných dat závislá na parametru θ .
- nebo $L(\dots) = f_{X_1, \dots, X_n}(x_1, \dots, x_n)$... hustota pravděpodobnosti ...
- $\ell(\theta; x_1, \dots, x_n) = \log L(\dots)$... pro snazší výpočty.
- Odhad metodou maximální věrohodnosti (Maximal Likelihood)* hledáme θ , pro které je maximální $L(\theta; x_1, \dots, x_n)$, resp. $\ell(\dots)$. Obvykle pomocí derivací funkce L , resp. ℓ .
- bias (vychýlení): $\mathbb{E}(\hat{\theta} - \theta)$... θ skutečný parametr, $\hat{\theta}$ náš odhad (náhodná veličina, protože závisí na naměřených datech)
- odhad je nevychýlený/nestranný/unbiased: bias = 0
- odhad je asymptoticky nevychýlený: bias konverguje k 0, neboli $\mathbb{E}(\hat{\theta}) \rightarrow \theta$
- odhad je konzistentní: $\hat{\theta}$ konverguje k 0 v pravděpodobnosti: pro všechna $\varepsilon > 0$: $P(|\hat{\theta} - \theta| > \varepsilon) \rightarrow 0$
- jeho *variance* je $\text{var}(\hat{\theta}) = \mathbb{E} \left[\left(\hat{\theta} - \mathbb{E}[\hat{\theta}] \right)^2 \right]$
- MSE (mean square error, střední kvadratická odchylka): $\mathbb{E}((\hat{\theta} - \theta)^2)$
- Věta: $\text{MSE} = \text{bias}^2 + \text{var}(\hat{\theta})$.