10. cvičení z PSt

Centrální limitní věta

1. Označme X_i dobu stahování i-tého souboru. Platí

$$E(X_i) = 5, \qquad \mathrm{Var}(X_i) = 4.$$

Pro součet $S_{50} = \sum_{i=1}^{50} X_i$ máme

$$E(S_{50}) = 250, \qquad \mathrm{Var}(S_{50}) = 200.$$

Podle CLT

$$\frac{S_{50} - 250}{\sqrt{200}} \approx N(0,1).$$

(a)

$$P(S_{50} > 270) \approx 1 - \Phi\left(\frac{270 - 250}{\sqrt{200}}\right) = 1 - \Phi(\sqrt{2}) \approx 1 - 0.93 = 0.07.$$

(b) Pro výběrový průměr $\overline{X}$ platí

$$E(\overline{X}) = 5, \qquad \mathrm{Var}(\overline{X}) = \frac{4}{50}.$$

Tedy

$$P(\overline{X} < 4.5) \approx \Phi\left(\frac{4.5 - 5}{2/\sqrt{50}}\right) = \Phi\left(-\frac{5\sqrt{2}}{4}\right) \approx \Phi(-1.77) \approx 0.04.$$

MLE a nestrannost

4. Nechť $X_1, \ldots, X_n \sim N(\mu, 1)$.

(a) Hustota jednoho pozorování je

$$f_\mu(x_i) = \frac{1}{\sqrt{2\pi}} e^{-(x_i - \mu)^2/2}.$$

Věrohodnost:

$$L(\mu) = \prod_{i=1}^{n} f_\mu(x_i) = (2\pi)^{-n/2} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}(x_i - \mu)^2\right).$$

Log-likelihood:

$$\ell(\mu) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{n}(x_i - \mu)^2.$$

(b)

$$\ell'(\mu) = \sum_{i=1}^{n}(x_i - \mu).$$

Z podmínky $\ell'(\mu) = 0$ dostáváme

$$\widehat{\mu}_{\mathrm{MLE}} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

(c)

$$E(\widehat{\mu}_{\mathrm{MLE}}) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \mu.$$

Estimátor je tedy nestranný.

(d) Pro obecné známé σ^2:

$$\ell(\mu) = -n \log \sigma - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2 + \text{konst.}$$

Derivace podle μ vede opět na

$$\widehat{\mu}_{\text{MLE}} = \frac{1}{n} \sum_{i=1}^{n} x_i.$$

6. Například estimátor

$$\widehat{\mu} = X_1$$

pro odhad střední hodnoty μ v modelu $X_i \sim N(\mu, 1)$. Platí

$$E(\widehat{\mu}) = \mu,$$

tedy je nestranný, ale ignoruje všechna ostatní pozorování, takže má zbytečně velký rozptyl:

$$\text{Var}(\widehat{\mu}) = 1, \qquad \text{Var}(\overline{X}) = \frac{1}{n}.$$