# Přednáška 7.

- Metoda maximální věrohodnosti (MLE)

# Metoda maximální věrohodnosti

**Připomenutí:**
PDF jako funkci neznámého parametru nazýváme **věrohodnostní funkce** (likelihood function):

$$p(\boldsymbol{\theta}|\boldsymbol{x}) = \prod_{i=1}^{N} p(\boldsymbol{\theta}|x_i)$$

**Princip maximálně věrohodného odhadu (Maximum Likelihood Estimation – MLE):**
Maximalizovat věrohodnost = odhadnout neznámý parametr tak, aby „co nejlépe“ odpovídal modelu dat

1. Často se opět využívá místo věrohodnostní funkce její logaritmus, tzv. **log-likelihood function** (označme symbolem $\mathcal{L}$).
2. Derivace $\mathcal{L} \rightarrow$ používá se pro nalezení maxima, tedy

$$\frac{\partial \ln p(\boldsymbol{\theta}|x)}{\partial \boldsymbol{\theta}} = 0$$

# MLE - vlastnosti

**Asymptotické rozdělení MLE odhadu:**

$$\boldsymbol{\hat{\theta}_{MLE} \sim N\left(\theta, \mathcal{J}^{-1}(\theta)\right)}$$

Kde $\boldsymbol{\mathcal{J}^{-1}(\theta)}$ je inverze Fisherovy informační matice.

**Důsledky:**

1. MLE odhady jsou asymptoticky nestranné.
2. MLE odhady jsou asymptoticky eficientní! Dosahují CRLB.
3. Přesnost MLE odhadů se dá (někdy) analyticky odvodit.

Příklad 1: Jak vypadá MLE odhad střední hodnoty exponenciálního rozdělení?

Příklad 2: Jak vypadá MLE odhad parametru $\mu$ pro rozdělení, které má hustotu pravděpodobnosti

$f_X(x) = \frac{\beta}{2\alpha\Gamma\left({}^{1}/_{\beta}\right)} e^{-\left(\frac{|x-\mu|}{\alpha}\right)^{\beta}}$. Pro jednoduchost zkusme uvažovat např. $\alpha = 1$.

# MLE – numerické řešení

Nelze nalézt analyticky taková $\theta$, že platí $\frac{\partial \ln p(\theta|x)}{\partial \theta} = 0$ . Proto označme

$$g(\theta) = \frac{\partial \ln p(\theta|x)}{\partial \theta}$$

A využijme rozvoj funkce $g(\theta)$ v okolí bodu $\theta_0$

$$g(\theta) \approx g(\theta_0) + \frac{\partial g(\theta)}{\partial \theta}|_{\theta=\theta_0}(\theta - \theta_0)$$

A z toho odvodíme vztah

$$\theta_1 = \theta_0 - \frac{g(\theta_0)}{\frac{\partial g(\theta)}{\partial \theta}|_{\theta=\theta_0}}$$

A následně rekurentní vztah

$$\theta_{k+1} = \theta_k - \frac{g(\theta_k)}{\frac{\partial g(\theta)}{\partial \theta}|_{\theta=\theta_k}}$$

Tato iterativní procedura se nazývá Newton-Raphsonův algoritmus.

Poznámky:

- Hrozí problémy s konvergencí (obzvláště pokud druhá derivace log-likelihood je malá)
- Velmi závisí na inicializaci (při špatné inicializaci hrozí konvergence k lokálním maximům, ale dokonce i minimům).

# Cvičení 7

1. Spočtěte MLE odhad střední hodnoty Normálního rozdělení.
2. Odhadněte parametr $\mu$ z dat, která mají rozdělení s hustotu pravděpodobnosti

$f_X(x) = \frac{\beta}{2\alpha\Gamma\left(1/\beta\right)}\, e^{-\left(\frac{|x-\mu|}{\alpha}\right)^{\beta}}$, kde $\alpha = 1$ a $\beta = 4$ .