Vlastnosti estimátorů

1. konzistence

• pro $n \to \infty$ konverguje $\hat{\vec{\theta}}$ k $\vec{\theta}$ $\lim_{n\to\infty} P\left(\left|\hat{\theta} - \theta\right| > \varepsilon\right) = 0$

2. předpojatost

• předpojatost $b \equiv E\left[\hat{\theta}\right] - \theta$

• $b = 0 \Rightarrow$ nepředpojatý (nevychýlený) odhad

3. efektivita

• MSE: $E\left[\left(\hat{\theta} - \theta\right)^2\right] = V\left[\hat{\theta}\right] + b^2$

statistická chyba estimátoru

systematická chyba estimátoru

Metoda maximální věrohodnosti

- sada n naměřených hodnot x_i : $\vec{x} = \left(x_1, x_2, \ldots, x_n\right)$ $\quad x_i \in f\left(x \mid \vec{\theta}\right)$ $\quad x_i$ nezávislé
- parametry: $\vec{\theta} = \left(\theta_1, \theta_2, \ldots, \theta_m\right)$
- cíl – najít nejlepší odhad $\hat{\vec{\theta}}$ parametrů $\vec{\theta}$
- odhad (estimátor) $\hat{\vec{\theta}} = \left(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_m\right)$

- pravděpodobnost $P\left\{x \in \left(x_i, x_i + \mathrm{d}x\right)\right\} = f\left(x_i \mid \vec{\theta}\right)\mathrm{d}x$
- pravděpodobnost, že naměříme hodnoty (x_1, x_2,…,x_n): $P = \prod_{i=1}^{n} f\left(x_i \mid \vec{\theta}\right)\mathrm{d}x$

věrohodnostní funkce $\quad L\left(\vec{\theta} \mid \vec{x}\right) \equiv \prod_{i=1}^{n} f\left(x_i \mid \vec{\theta}\right)$

- hledáme hodnoty $\hat{\vec{\theta}}$ pro které L nabývá maximum

$$\frac{\partial L}{\partial \theta_k} = 0, \qquad k = 1,2,\ldots m \qquad \Rightarrow \qquad \hat{\theta}_k$$

Odhad parametrů normálního rozdělení

- naměřené hodnoty: $\vec{x} = (x_1, x_2, \ldots, x_n)$

- hustota pravděpodobnosti: $f(x \mid \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[\frac{-(x-\mu)^2}{2\sigma^2}\right]$

- věrohodnostní funkce: $L(\mu, \sigma \mid \vec{x}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[\frac{-(x_i-\mu)^2}{2\sigma^2}\right]$

$$\ln L(\mu, \sigma \mid \vec{x}) = \sum_{i=1}^{n} \frac{-(x_i-\mu)^2}{2\sigma^2} - \sum_{i=1}^{n} \ln\left(\sqrt{2\pi}\sigma\right)$$

- odhad očekávané hodnoty: $\frac{\partial \ln L}{\partial \mu}\Big|_{\mu=\hat{\mu}} = 0 \quad \Rightarrow \quad \boxed{\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i}$

- odhad rozptylu: $\frac{\partial \ln L}{\partial \sigma}\Big|_{\sigma=\hat{\sigma}} = 0 \quad \Rightarrow \quad \boxed{\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} (x_i - \mu)^2}$

Musíme znát skutečnou očekávanou hodnotu

Odhad parametrů normálního rozdělení

- odhad očekávané hodnoty: $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i \equiv \bar{x}$ (aritmetický průměr)
- odhad rozptylu: $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 \approx \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 \equiv s_0^2$

- předpojatost?

$$E[\hat{\mu}] = \frac{1}{n}\sum_{i=1}^{n} E[x_i] = \mu \quad \Rightarrow \text{nepředpojatý odhad}$$

$$E[\hat{\sigma}^2] = \frac{1}{n}\sum_{i=1}^{n} E\left[(x_i - \mu)^2\right] = \sigma^2 \quad \Rightarrow \text{nepředpojatý odhad}$$

ale

$$E[s_0^2] = \frac{1}{n}\sum_{i=1}^{n} E\left[(x_i - \bar{x})^2\right] = \frac{n-1}{n}\sigma^2 \quad \Rightarrow \text{předpojatý odhad}$$

- nepředpojatý odhad rozptylu: $s_1^2 \equiv \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$

Metoda nejmenších čtverců

$x_1, x_2, \dots x_N$

$y_1, y_2, \dots y_N$ $\quad y_i \in N(\lambda_i, \sigma_i)$ • experimentální data

$\lambda(x|\vec{\theta})$ $\quad \vec{\theta} = (\theta_1, \dots \theta_m)$ • modelová funkce

x – nezávislá proměnná

y – závislá proměnná

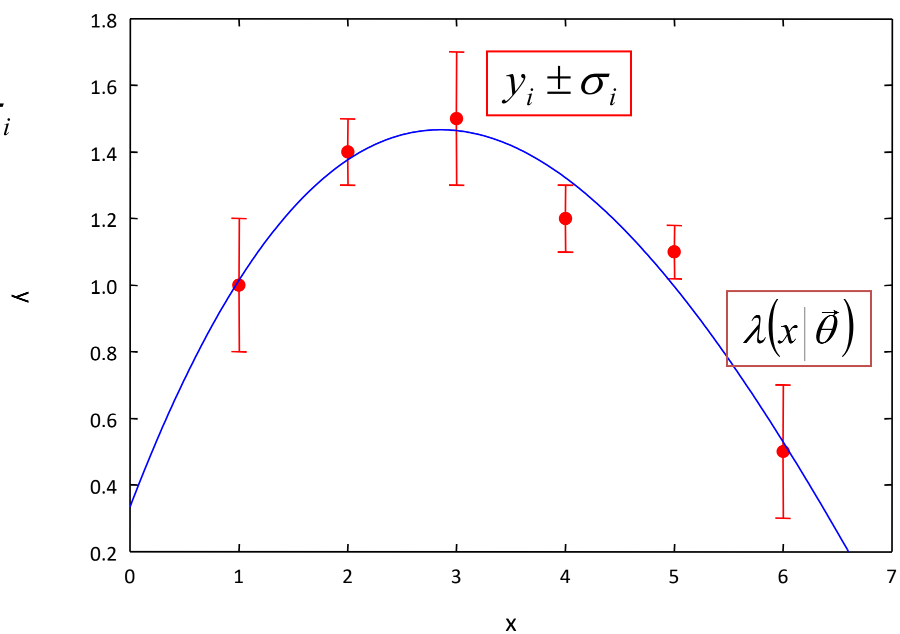
$$L(\vec{\theta}\,|\,\vec{y}) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left[-\frac{\left(y_i - \lambda(x_i\,|\,\vec{\theta})\right)^2}{2\sigma_i^2}\right]$$

• věrohodnostní funkce

$$\ln L(\vec{\theta}\,|\,\vec{y}) = -\sum_{i=1}^{N} \frac{\left(y_i - \lambda(x_i\,|\,\vec{\theta})\right)^2}{2\sigma_i^2} - \sum_{i=1}^{N} \ln \sqrt{2\pi}\sigma_i$$

$$\chi^2(\vec{\theta}\,|\,\vec{y}) = \sum_{i=1}^{N} \frac{\left(y_i - \lambda(x_i\,|\,\vec{\theta})\right)^2}{\sigma_i^2}$$

• minimalizace χ^2



Metoda nejmenších čtverců – lineární fit $y = m\,x$

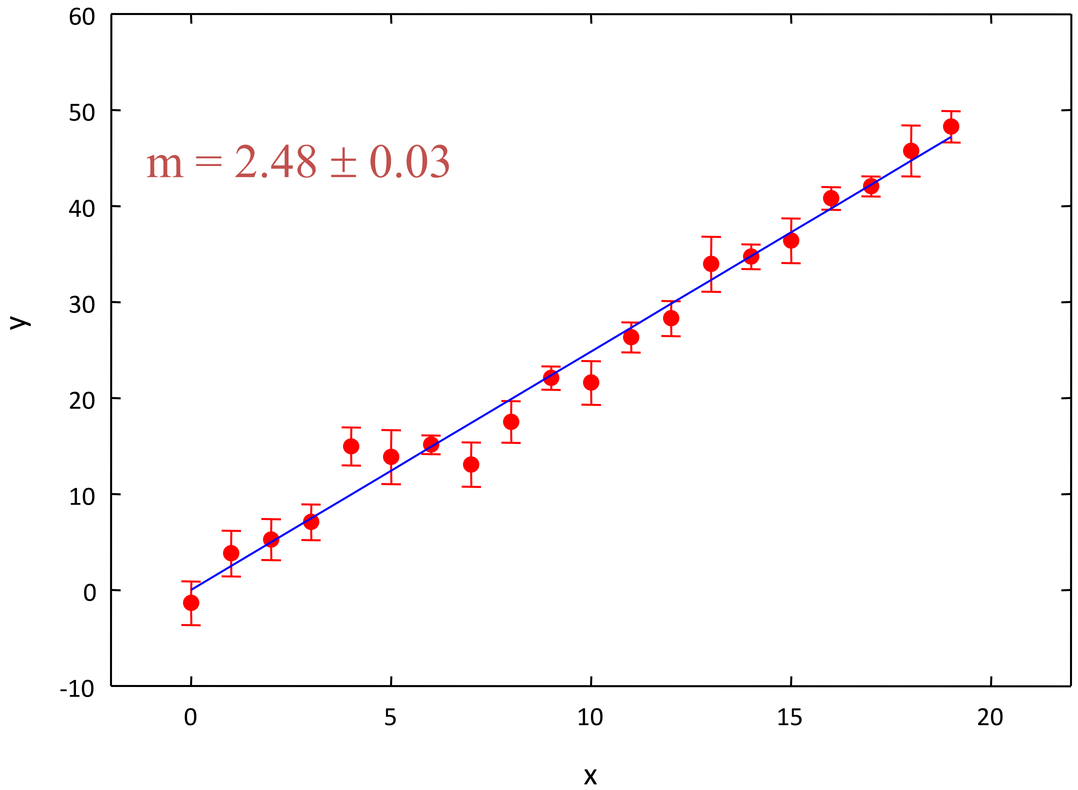
$$\lambda(x \mid m) = mx$$

$$\chi^2(m \mid \vec{y}) = \sum_{i=1}^{N} \frac{(y_i - mx_i)^2}{\sigma_i^2}$$

$$\widehat{m} = \frac{\sum_{i=1}^{N} \frac{y_i x_i}{\sigma_i^2}}{\sum_{i=1}^{N} \frac{x_i^2}{\sigma_i^2}} = \frac{\langle xy \rangle}{\langle x^2 \rangle}$$

$$V[\widehat{m}] = \frac{1}{\sum_{i=1}^{N} \frac{x_i^2}{\sigma_i^2}} = \frac{1}{\langle x^2 \rangle}$$

označení: $\langle a \rangle \equiv \sum_{i=1}^{N} \frac{a_i}{\sigma_i^2}$



Metoda nejmenších čtverců – vyjádření pomocí matic

$$x_1, x_2, \ldots x_N$$

$$y_1, y_2, \ldots y_N \qquad y_i \in N(\lambda_i, \sigma_i)$$

$$\lambda = \lambda(x \,|\, \vec{\theta}) \qquad \vec{\theta} = (\theta_1, \ldots \theta_m)$$

lineární model

$$\lambda(x_i \,|\, \vec{\theta}) = \sum_{j=1}^{m} a_j(x_i)\theta_j = \sum_{j=1}^{m} A_{ij}\theta_j$$

$$\boldsymbol{\lambda} = \mathbf{A\theta}$$

$$\chi^2 = \sum_{i=1}^{n} \frac{\left(y_i - \sum_{j=1}^{m} A_{ij}\theta_j\right)^2}{\sigma_i^2}$$

pokud $\sigma_i = \sigma$ $\downarrow$

$$\chi^2 = \frac{1}{\sigma^2}(\mathbf{y} - \mathbf{A\theta})^T(\mathbf{y} - \mathbf{A\theta})$$

obecně $\mathbf{V}$, $V_{ij} = \mathrm{cov}(y_i, y_j)$

$$\chi^2 = (\mathbf{y} - \mathbf{A\theta})^T \mathbf{V}^{-1}(\mathbf{y} - \mathbf{A\theta})$$

$$\frac{\partial \chi^2}{\partial \theta_k} = -\frac{2}{\sigma^2}\sum_{i=1}^{N}\left(y_i - \sum_{j=1}^{m} A_{ij}\theta_j\right)A_{ik}$$

$$\frac{\partial \chi^2}{\partial \theta_k} = 0 \quad \downarrow$$

$$\sum_{i=1}^{N} A_{ik} y_i = \sum_{i=1}^{N}\sum_{j=1}^{m} A_{ik}A_{ij}\theta_j$$

$$\mathbf{A}^T\mathbf{y} = \mathbf{A}^T\mathbf{A\theta}$$

$$\hat{\boldsymbol{\theta}} = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{y} \equiv \mathbf{B}\,\mathbf{y}$$

$$\hat{\boldsymbol{\theta}} = \left(\mathbf{A}^T\mathbf{V}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{V}^{-1}\mathbf{y} \equiv \mathbf{B}\,\mathbf{y}$$

$\mathbf{U}$, $U_{ij} = \mathrm{cov}(\hat{\theta}_i, \hat{\theta}_j)$

$$\mathbf{U} = \mathbf{B\,VB}^T$$

$$\left(U^{-1}\right)_{ij} = \frac{1}{2}\left[\frac{\partial^2 \chi^2}{\partial \theta_i \partial \theta_j}\right]_{\vec{\theta} = \hat{\vec{\theta}}}$$

Metoda nejmenších čtverců – fit polynomu

$$\lambda\left(x_i \,\middle|\, \vec{\theta}\right) = \sum_{j=1}^{m} x_i^j \theta_j = \sum_{j=1}^{m} A_{ij}\theta_j$$

$$\vec{\theta} = \left(\theta_1, \ldots \theta_m\right) \qquad \sigma_i \equiv \sigma$$

$$A_{ij} = x_i^j$$

$$\mathbf{A} = \begin{pmatrix} 1 & x_1 & x_1^2 & \ldots x_1^m \\ 1 & x_2 & x_2^2 & \ldots x_2^m \\ \ldots & & & \\ 1 & x_N & x_N^2 & \ldots x_N^m \end{pmatrix} \qquad N \times m$$

$$\mathbf{A}^T\mathbf{A} = \begin{pmatrix} N & \sum_{i=1}^{N} x_i & \sum_{i=1}^{N} x_i^2 & \ldots \sum_{i=1}^{N} x_i^m \\ \sum_{i=1}^{N} x_i & \sum_{i=1}^{N} x_i^2 & \sum_{i=1}^{N} x_i^3 & \ldots \sum_{i=1}^{N} x_i^{m+1} \\ \ldots & & & \\ \sum_{i=1}^{N} x_i^m & \sum_{i=1}^{N} x_i^{m+1} & \sum_{i=1}^{N} x_i^{m+2} & \ldots \sum_{i=1}^{N} x_i^{m+m} \end{pmatrix} \qquad m \times m$$

$$\hat{\theta} = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{y}$$

Metoda nejmenších čtverců – fit paraboly

$$\mathbf{A}^T\mathbf{V}^{-1}\mathbf{A} = \begin{pmatrix} \sum_{i=1}^{N}\frac{1}{\sigma_i^2} & \sum_{i=1}^{N}\frac{x_i}{\sigma_i^2} & \sum_{i=1}^{N}\frac{x_i^2}{\sigma_i^2} \\ \sum_{i=1}^{N}\frac{x_i}{\sigma_i^2} & \sum_{i=1}^{N}\frac{x_i^2}{\sigma_i^2} & \sum_{i=1}^{N}\frac{x_i^3}{\sigma_i^2} \\ \sum_{i=1}^{N}\frac{x_i^2}{\sigma_i^2} & \sum_{i=1}^{N}\frac{x_i^3}{\sigma_i^2} & \sum_{i=1}^{N}\frac{x_i^4}{\sigma_i^2} \end{pmatrix}$$

$$\hat{\boldsymbol{\theta}} = \left(\mathbf{A}^T\mathbf{V}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{V}^{-1}\mathbf{y} \equiv \mathbf{B}\,\mathbf{y}$$

$$\mathbf{U},\quad U_{ij} = \mathrm{cov}\left(\hat{\theta}_i, \hat{\theta}_j\right)$$

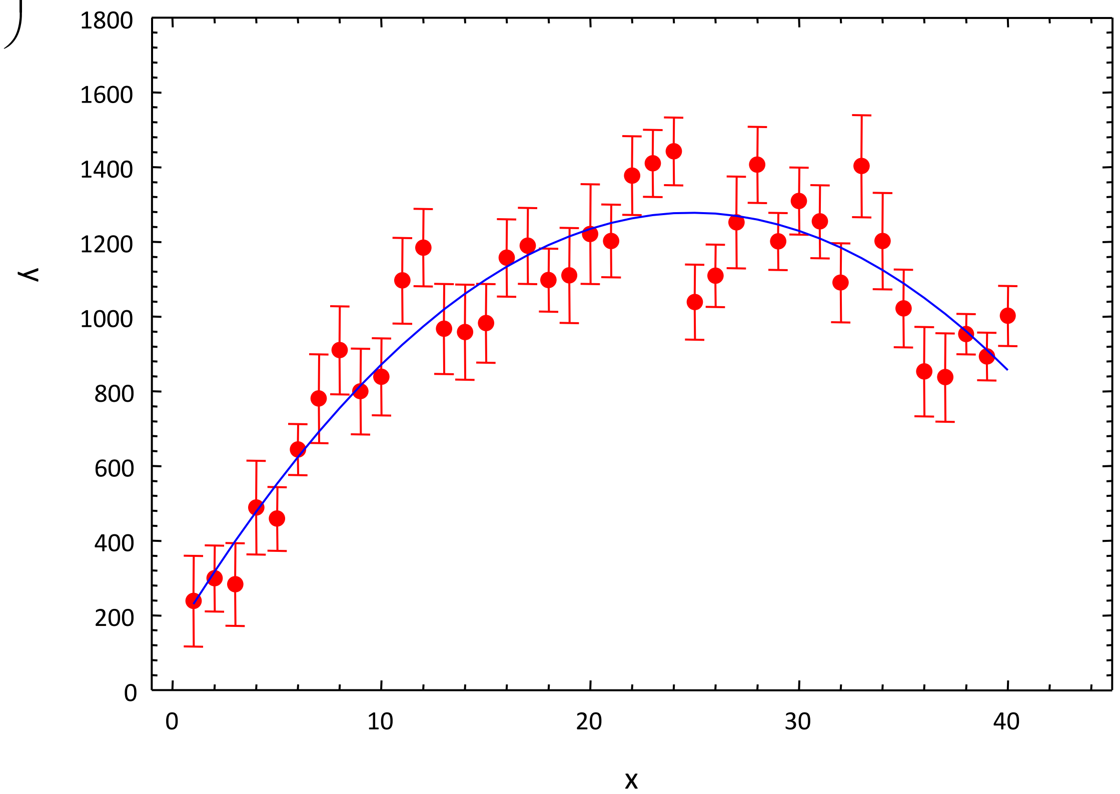
$$\mathbf{U} = \mathbf{B}\mathbf{V}^{-1}\mathbf{B}^T$$

$$\lambda\left(x|\theta\right) = \theta_0 + \theta_1 x + \theta_2 x^2$$

$$\theta_0 = 140 \pm 50$$

$$\theta_1 = 92 \pm 5$$

$$\theta_2 = -1.8 \pm 0.1$$

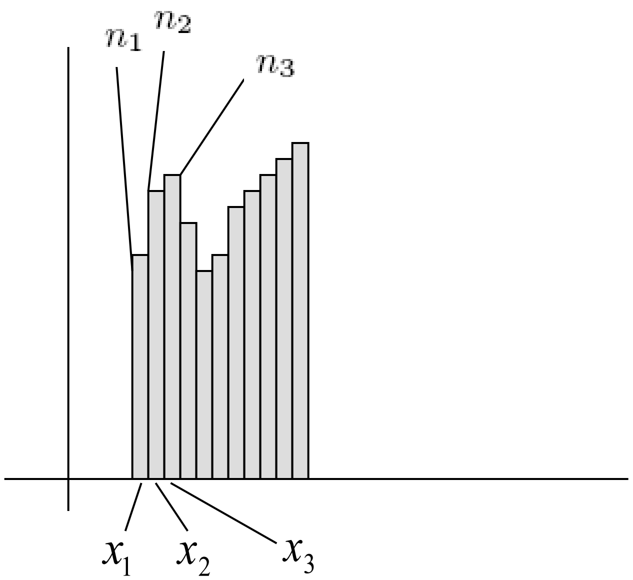


Fit histogramu

$x_1, x_2, \ldots x_N \qquad f(x|\boldsymbol{\theta}) \qquad \boldsymbol{\theta} = (\theta_1, \theta_2, \ldots \theta_M)$

- místo $x_1, x_2, \ldots x_N \longrightarrow$ histogram K binů $\mathbf{n} = (n_1, n_2, \ldots n_K)$
- pravděpodobnost, že výsledek padne do i-tého binu: $P_i = \int_{x_i - \frac{\Delta}{2}}^{x_i + \frac{\Delta}{2}} f(x|\boldsymbol{\theta}) \mathrm{d}x$
- očekávaná hodnota v i-tém binu: $\nu_i(\boldsymbol{\theta}) = NP_i = N \int_{x_i - \frac{\Delta}{2}}^{x_i + \frac{\Delta}{2}} f(x|\boldsymbol{\theta}) \mathrm{d}x$
- věrohodnost:

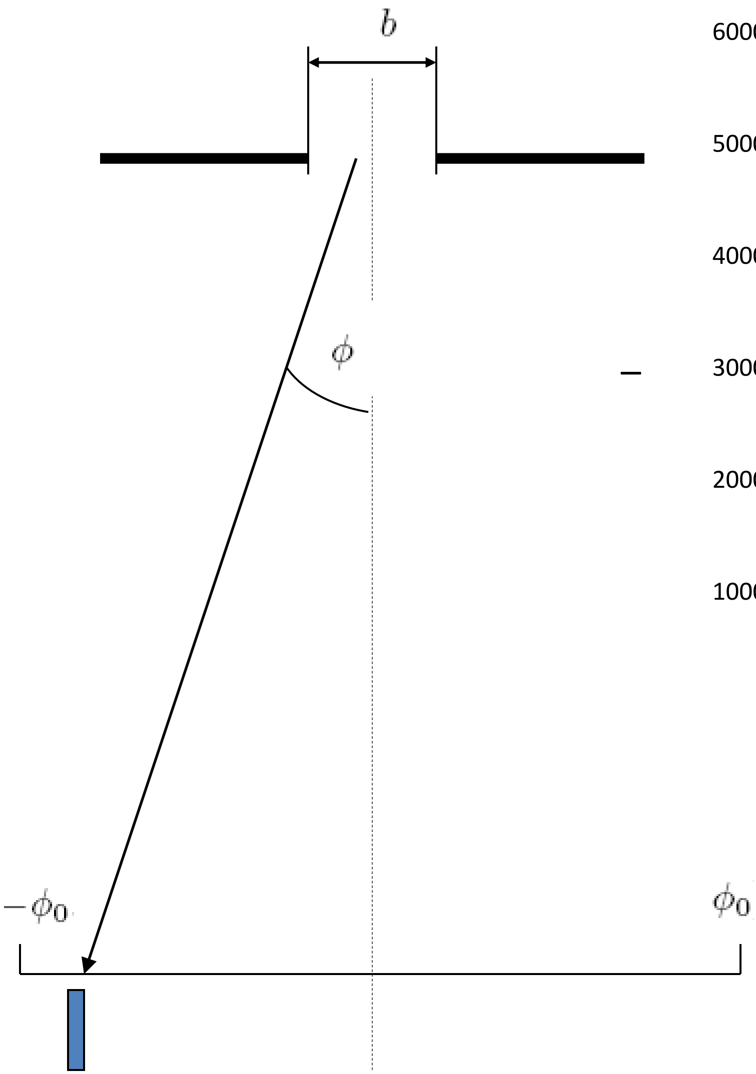
$$L(\boldsymbol{\theta}) = P(\mathbf{n}|\boldsymbol{\nu}) = \frac{N!}{n_1! n_2! \ldots n_K!} \left(\frac{\nu_1(\boldsymbol{\theta})}{N}\right)^{n_1} \ldots \left(\frac{\nu_K(\boldsymbol{\theta})}{N}\right)^{n_K}$$

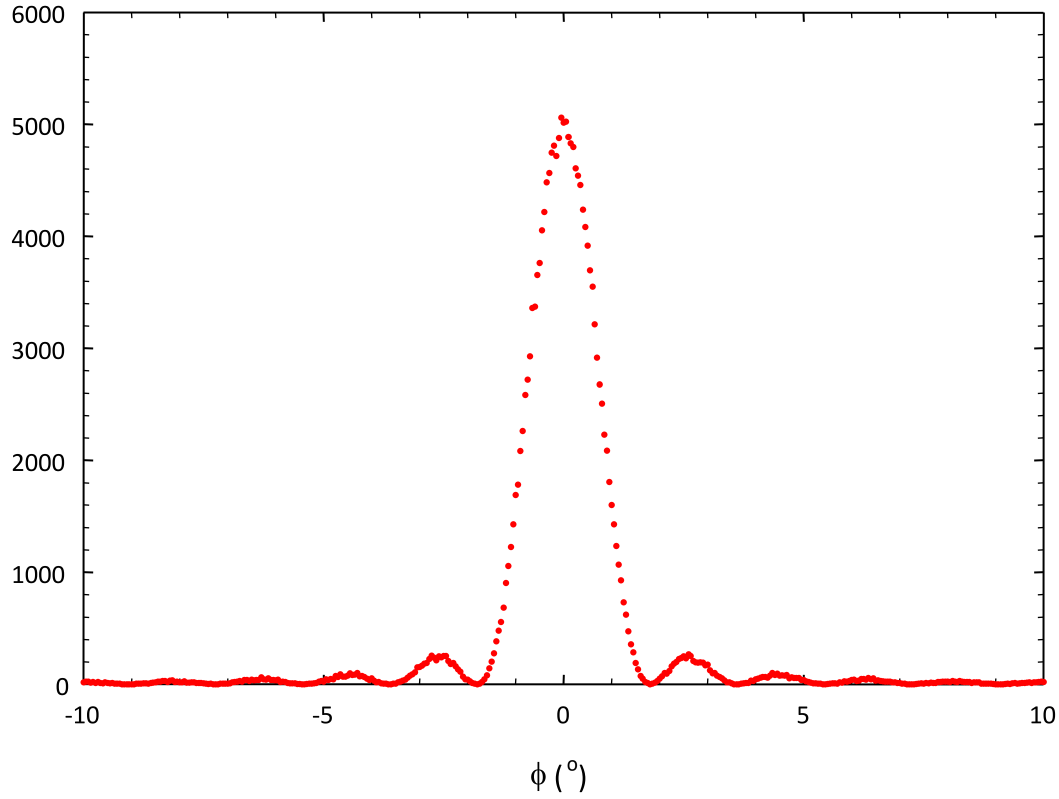


$$\ln L(\boldsymbol{\theta}) = \sum_{i=1}^{K} n_i \ln[\nu_i(\boldsymbol{\theta})] + \ln\left(\frac{N!}{n_1! n_2! \ldots n_K!}\right)$$

posteriorní hustota pravděpodobnosti: $f(\boldsymbol{\theta}|\mathbf{n}) \propto L(\boldsymbol{\theta}) f(\boldsymbol{\theta})$

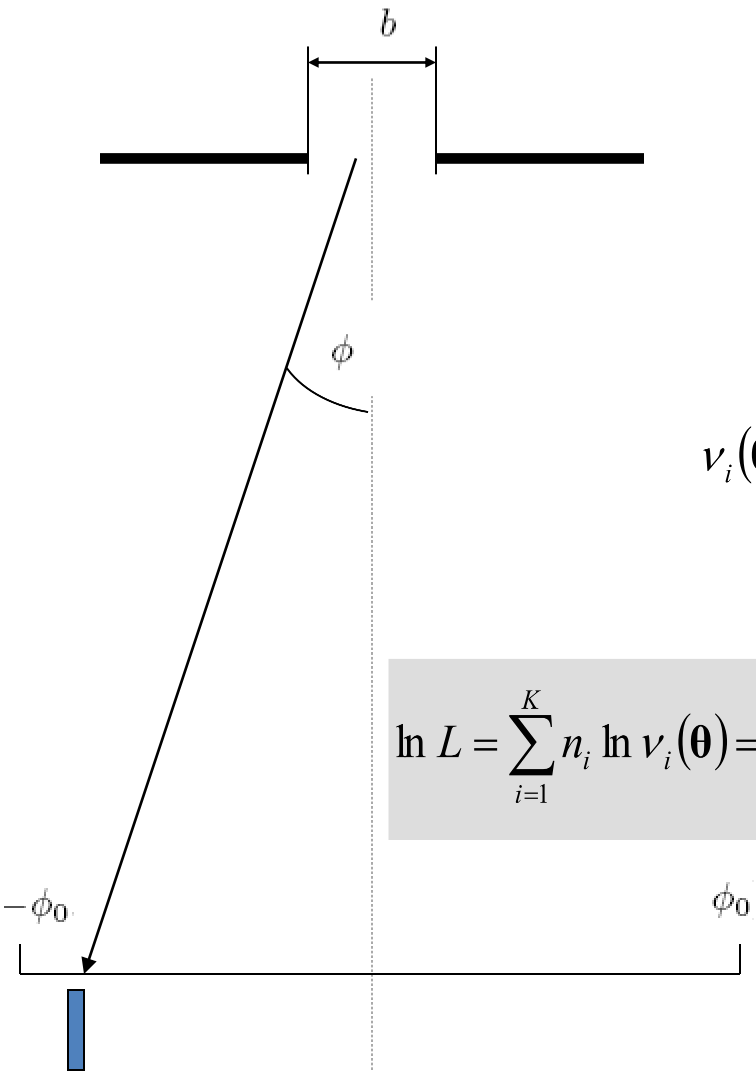
Příklad – difrakce na štěrbině





$$I = I_0\left(\frac{\sin u}{u}\right)^2 \qquad u = \frac{1}{2}kb\varphi \qquad (-\varphi_0, \varphi_0)$$

Příklad – difrakce na štěrbině



$$I = I_0\left(\frac{\sin u}{u}\right)^2 \qquad u = \frac{1}{2}kb\varphi \qquad (-\varphi_0, \varphi_0)$$

- normalizace: $\frac{1}{\xi} = \int_{-\varphi_0}^{\varphi_0} I(\varphi)\,\mathrm{d}\varphi \qquad I = \xi(b)\left(\frac{\sin u(b)}{u(b)}\right)^2$

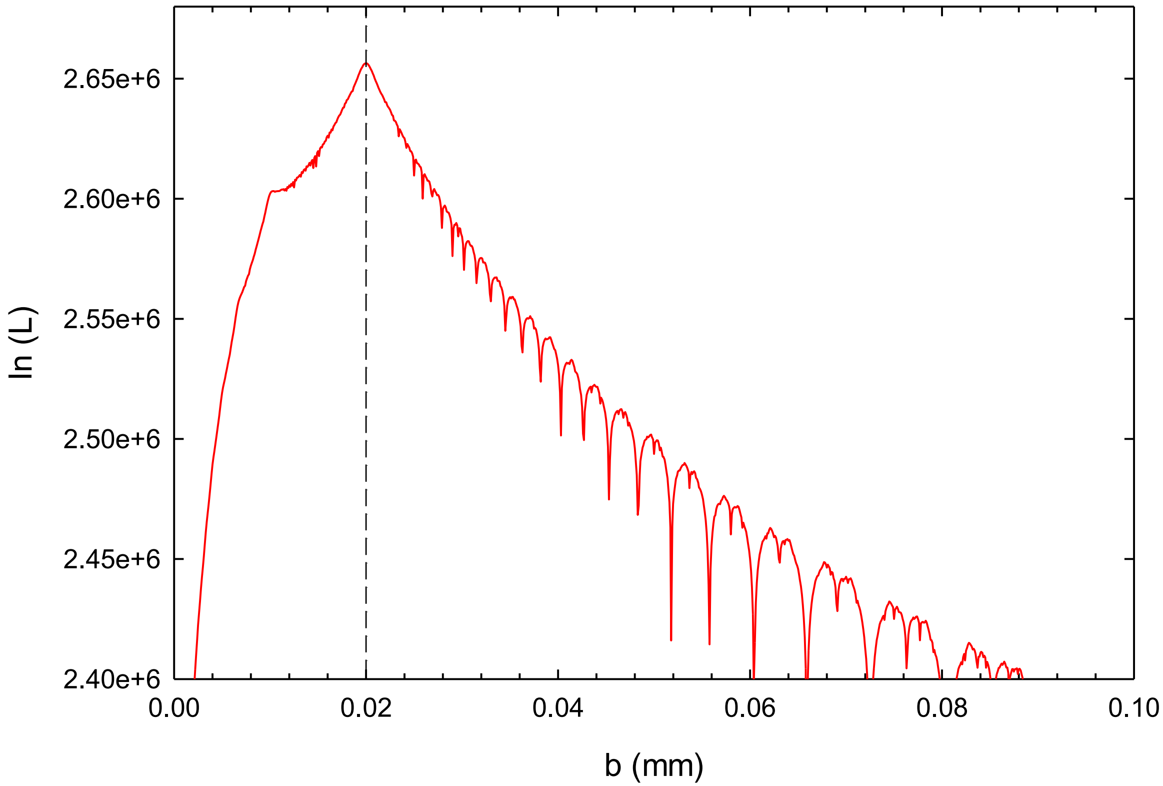
$$\nu_i(\boldsymbol{\theta}) = N\int_{\varphi_i-\frac{\Delta}{2}}^{\varphi_i+\frac{\Delta}{2}} \xi(b)\left(\frac{\sin\left(\frac{1}{2}kb\varphi\right)}{\frac{1}{2}kb\varphi}\right)^2 \mathrm{d}\varphi \approx N\xi(b)\left(\frac{\sin\left(\frac{1}{2}kb\varphi\right)}{\frac{1}{2}kb\varphi}\right)^2 \Delta$$

$$\ln L = \sum_{i=1}^{K} n_i \ln \nu_i(\boldsymbol{\theta}) = \sum_{i=1}^{K} n_i \ln(\xi(b)) + 2\sum_{i=1}^{K} n_i \ln\left(\sin\left(\frac{1}{2}kb\varphi_i\right)\right) - 2\sum_{i=1}^{K} n_i \ln\left(\frac{1}{2}kb\varphi_i\right) + \ln\Delta$$

$$\sigma_b \approx \left(-\left.\frac{\mathrm{d}^2 \ln L}{\mathrm{d}b^2}\right|_{b=b_0}\right)^{-\frac{1}{2}}$$

Příklad – difrakce na štěrbině

$$\ln L = \sum_{i=1}^{K} n_i \ln \nu_i(\boldsymbol{\theta}) = \sum_{i=1}^{K} n_i \ln\left(\xi(b)\right) + 2\sum_{i=1}^{K} n_i \ln\left(\sin\left(\frac{1}{2}kb\varphi_i\right)\right) - 2\sum_{i=1}^{K} n_i \ln\left(\frac{1}{2}kb\varphi_i\right) + \ln\Delta$$



Příklad – difrakce na štěrbině

$$\frac{n_i - I(\phi_i)}{\sqrt{n_i}}$$