Zobecněná metoda nejmenších čtverců

• pozorovatelné: $\boldsymbol{y} = (y_1, y_2, \ldots, y_J)^T$ $J \times 1$

• teoretický model: $\boldsymbol{y} = \boldsymbol{y}(\boldsymbol{\theta})$

• zpravidla $M < J$

• parametry: $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_M)^T$ $M \times 1$

• **apriorní informace:**

- odhad vektoru parametrů: $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots, \xi_M)^T$ $M \times 1$
- kovarianční matice: $A_{i,j} = \mathrm{cov}(\xi_i, \xi_j)$ $M \times M$

• apriorní hustota pravděpodobnosti: $f(\boldsymbol{\theta} | \boldsymbol{\xi}, \mathbf{A}) \propto \exp\left[-\frac{1}{2}(\boldsymbol{\xi} - \boldsymbol{\theta})^T \mathbf{A}^{-1} (\boldsymbol{\xi} - \boldsymbol{\theta})\right]$

• **posteriorní informace:**

- naměřená data: $\boldsymbol{\eta} = (\eta_1, \eta_2, \ldots, \eta_N)^T$ $N \times 1$
- kovarianční matice: $B_{i,j} = \mathrm{cov}(\eta_i, \eta_j)$ $N \times N$

• věrohodnost: $P(\boldsymbol{\eta} | \mathbf{y}(\boldsymbol{\theta}), \mathbf{B}) \propto \exp\left[-\frac{1}{2}(\boldsymbol{\eta} - \mathbf{y}(\boldsymbol{\theta}))^T \mathbf{B}^{-1} (\boldsymbol{\eta} - \mathbf{y}(\boldsymbol{\theta}))\right]$

Zobecněná metoda nejmenších čtverců

- **posteriorní hustota pravděpodobnosti:**

$$f(\theta|\eta, B, \xi, A) \propto P(\eta|y(\theta), B) f(\theta|\xi, A)$$

$$\boxed{f(\theta|\eta, B, \xi, A) \propto \exp\left[-\frac{1}{2}\left(\eta - y(\theta)\right)^T B^{-1}\left(\eta - y(\theta)\right) - \frac{1}{2}(\xi - \theta)^T A^{-1}(\xi - \theta)\right]}$$

$$\ln f = -\frac{1}{2}\left(\eta - y(\theta)\right)^T B^{-1}\left(\eta - y(\theta)\right) - \frac{1}{2}\left(\xi - \theta\right)^T A^{-1}\left(\xi - \theta\right) + K$$

- pokud jsou odhady parametrů ξ a naměřené hodnoty η nezávislé:

$$\ln f = -\frac{1}{2}\sum_{i=1}^{N}\frac{\left(\eta_i - y_i(\theta)\right)^2}{\sigma_{\eta_i}^2} - \frac{1}{2}\sum_{j=1}^{M}\frac{(\xi_j - \theta_j)^2}{\sigma_{\xi_j}^2} + K$$

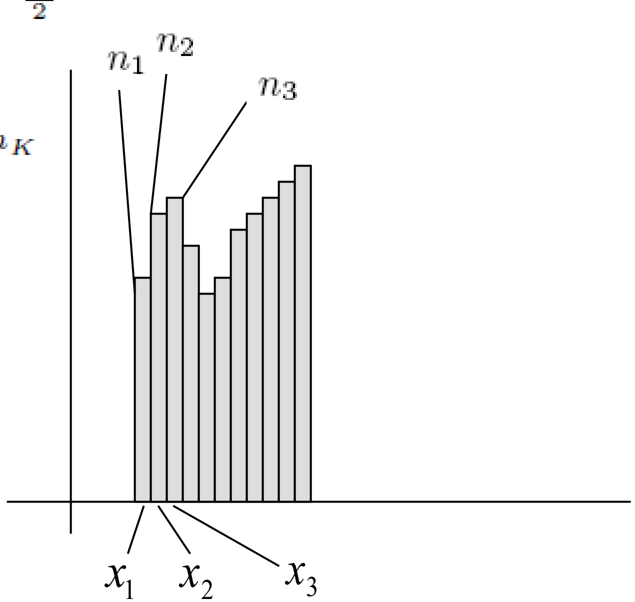
Fit histogramu

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

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

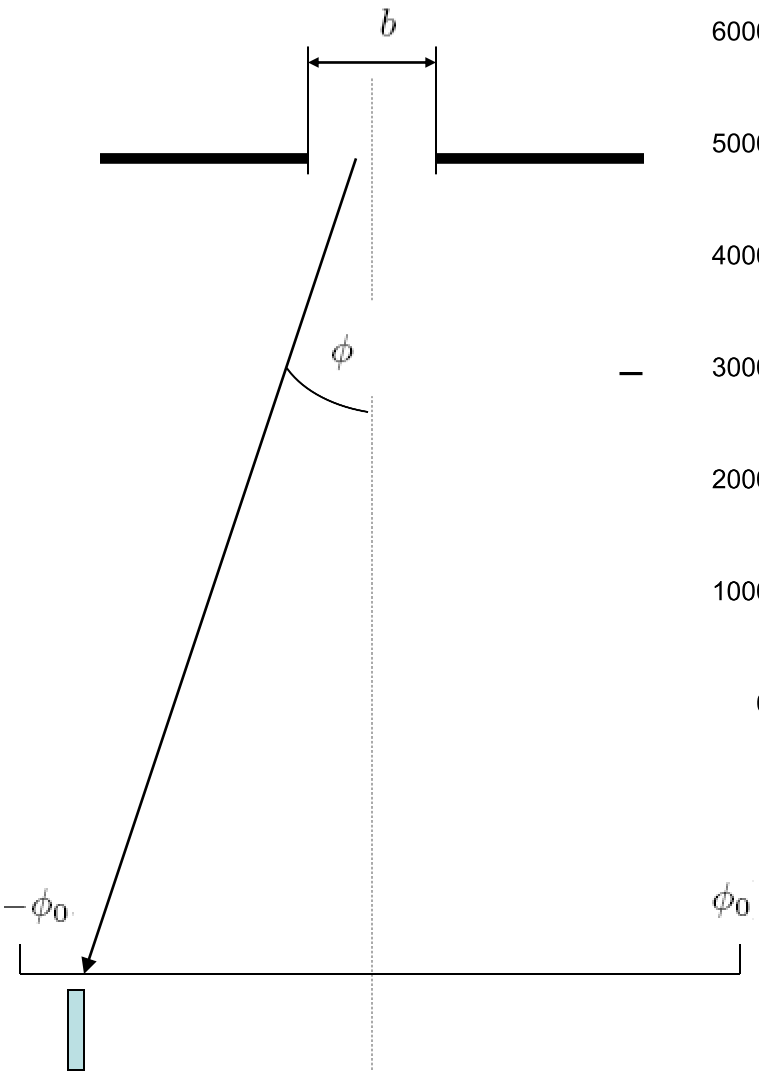
$$L(\boldsymbol{\theta}) = P(\boldsymbol{n}|\boldsymbol{\nu}) = \frac{N!}{n_1! n_2! \ldots n_K!} \left(\frac{\nu_1(\boldsymbol{\theta})}{N} \right)^{n_1} \ldots \left(\frac{\nu_K(\boldsymbol{\theta})}{N} \right)^{n_K}$$

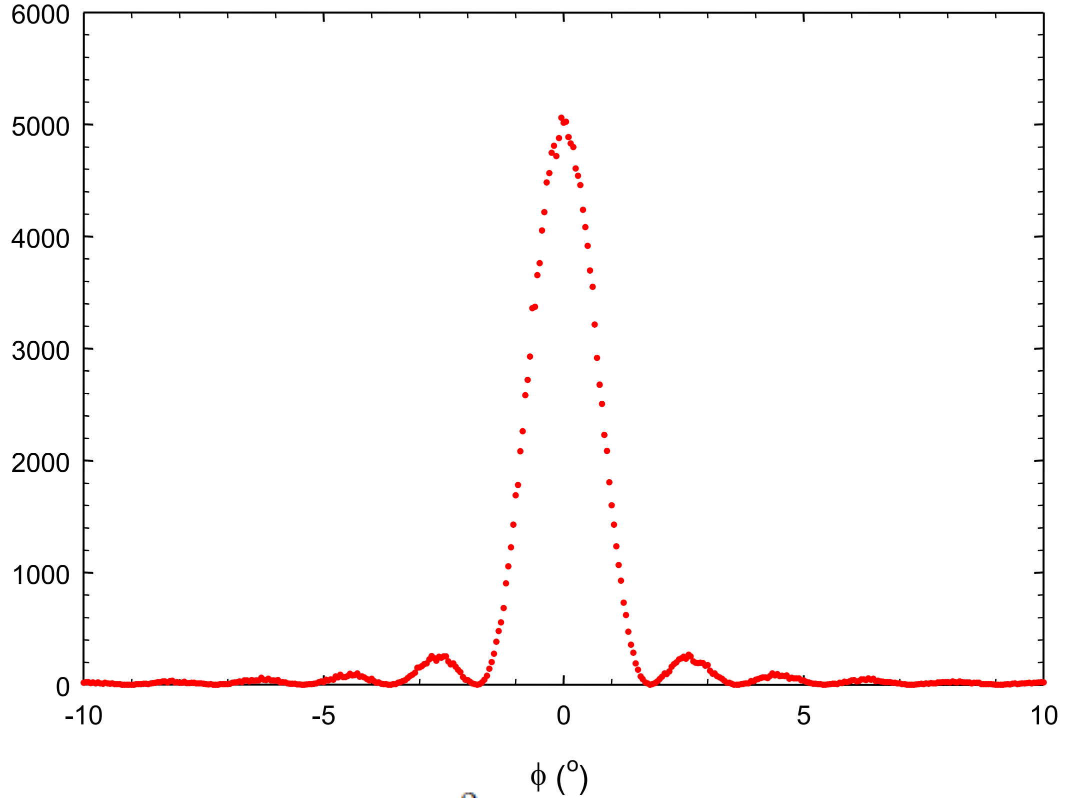
$$\ln L(\theta) = \sum_{i=1}^{K} n_i \ln [\nu_i(\theta] + \ln \left(\frac{N!}{n_1! n_2! \ldots n_K!} \right) - N \ln N$$



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

Příklad – difrakce na štěrbině

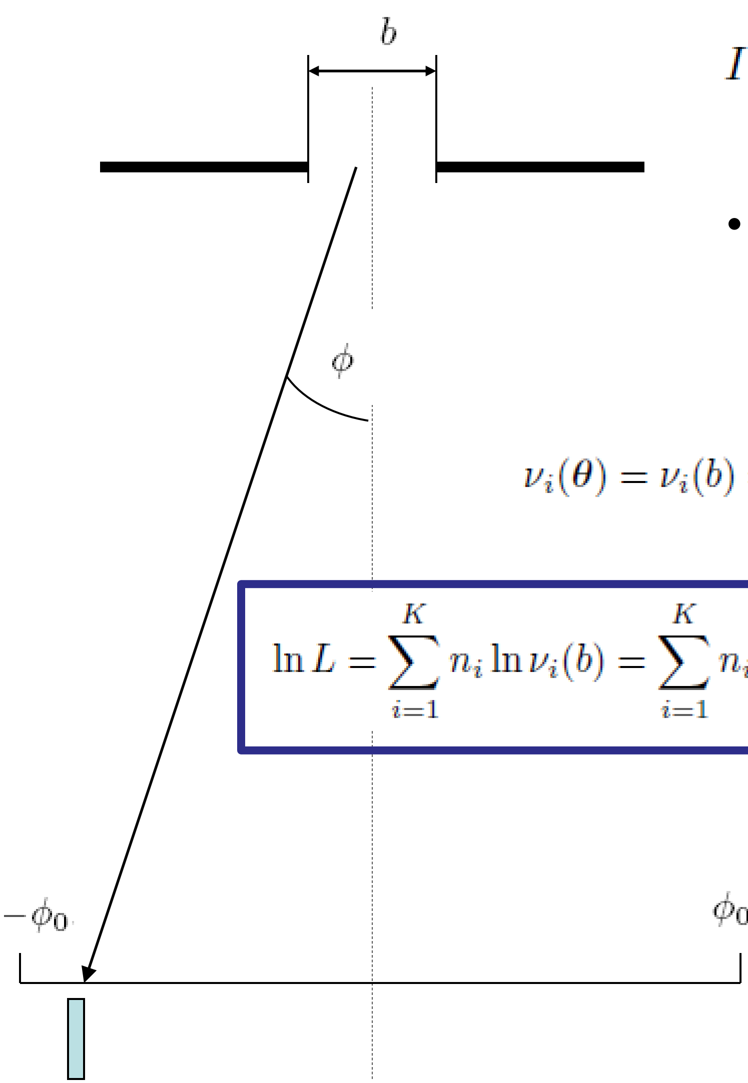




$$I = I_0 \left(\frac{\sin u}{u} \right)^2$$

$$u = \frac{1}{2} k b \phi = \frac{b\phi}{\lambda} \qquad \phi \in \langle -\phi_0, \phi_0 \rangle$$

Příklad – difrakce na štěrbině



$$I = I_0 \left(\frac{\sin u}{u} \right)^2 \qquad u = \frac{1}{2} k b \phi = \frac{b\phi}{\lambda} \qquad \phi \in \langle -\phi_0, \phi_0 \rangle$$

- normalizace: $\displaystyle \frac{1}{\xi} = \int_{-\phi_0}^{\phi_0} I(\phi)\, d\phi$

$$\boxed{I = \xi(b) \left(\frac{\sin u(b)}{u(b)} \right)^2}$$

$$\nu_i(\theta) = \nu_i(b) = N \int_{\phi_i - \frac{\Delta}{2}}^{\phi_i + \frac{\Delta}{2}} \xi(b) \left(\frac{\sin\left(\frac{1}{2} k b \phi\right)}{\frac{1}{2} k b \phi} \right)^2 d\phi \approx N \xi(b) \left(\frac{\sin\left(\frac{1}{2} k b \phi_i\right)}{\frac{1}{2} k b \phi_i} \right)^2 \Delta$$

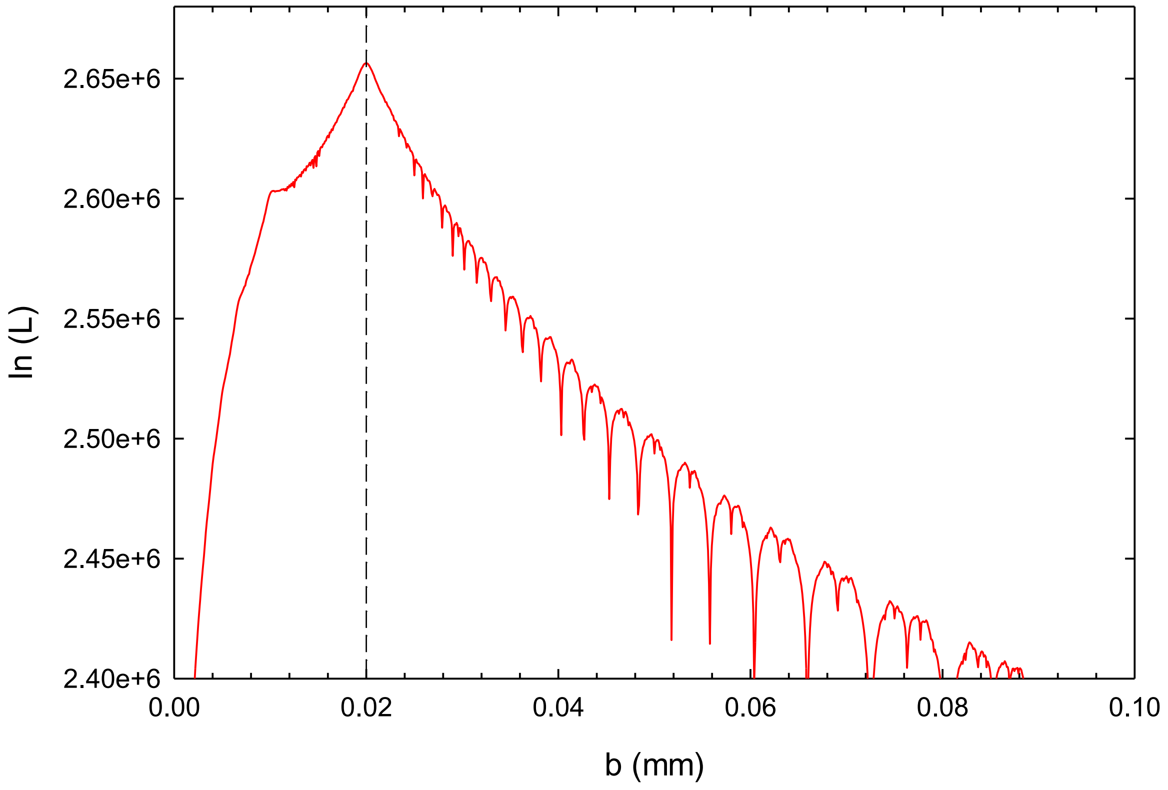
$$\boxed{\ln L = \sum_{i=1}^{K} n_i \ln \nu_i(b) = \sum_{i=1}^{K} n_i \ln\left(\xi(b)\right) + 2 \sum_{i=1}^{K} \ln\left(\sin\left(\frac{1}{2} k b \phi_i \right) \right) - 2 \sum_{i=1}^{K} n_i \ln\left(\frac{1}{2} k b \phi_i \right) + N \ln(N\Delta)}$$

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

Příklad – difrakce na štěrbině

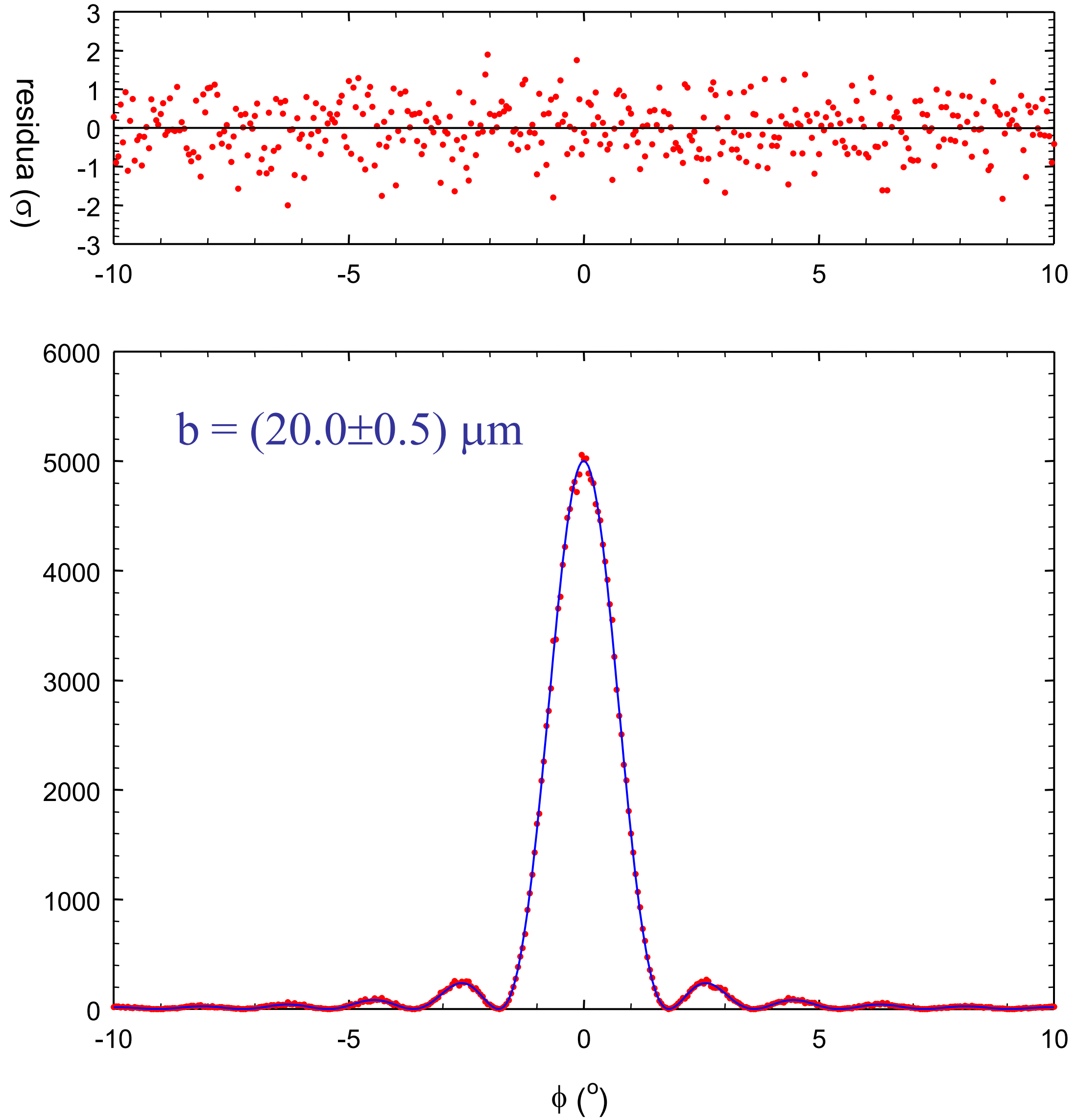
$$\ln L = \sum_{i=1}^{K} n_i \ln \nu_i(b) = \sum_{i=1}^{K} n_i \ln\left(\xi(b)\right) + 2\sum_{i=1}^{K} \ln\left(\sin\left(\frac{1}{2}kb\phi_i\right)\right) - 2\sum_{i=1}^{K} n_i \ln\left(\frac{1}{2}kb\phi_i\right) + N\ln(N\Delta)$$

b_0



Příklad – difrakce na štěrbině

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



Testování hypotéz

- hypotéza H
- naměřená data D

věrohodnost, tj. pravděpodobnost,
že naměříme data D
pokud hypotéza H platí

apriorní víra v pravdivost hypotézy H

posteriorní víra
v pravdivost
hypotézy H

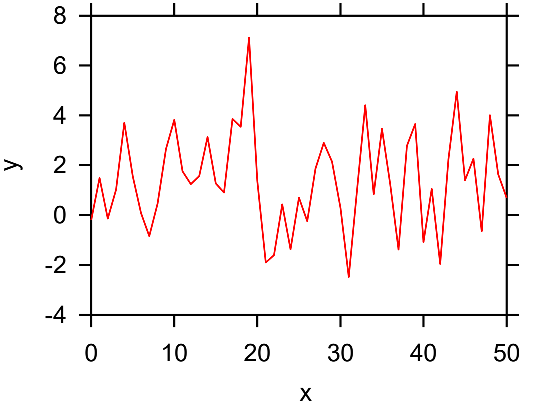
$$P(H|D,I) = \frac{P(D|H,I)P(H|I)}{P(D|I)}$$

normovací člen
pravděpodobnost, že naměříme data D

- použití zákona celkové pravděpodobnosti $P(D|I) = P(D|H,I)P(H|I) + P(D|\overline{H},I)P(\overline{H},I)$

$$P(H|D,I) = \frac{P(D|H,I)P(H|I)}{P(D|H,I)P(H|I) + P(D|\overline{H},I)P(\overline{H},I)}$$

Dvě alternativní hypotézy



- hypotéza A: veličina $y = 0$
- hypotéza B: veličina y je nějaká konstanta λ

poměr věrohodností

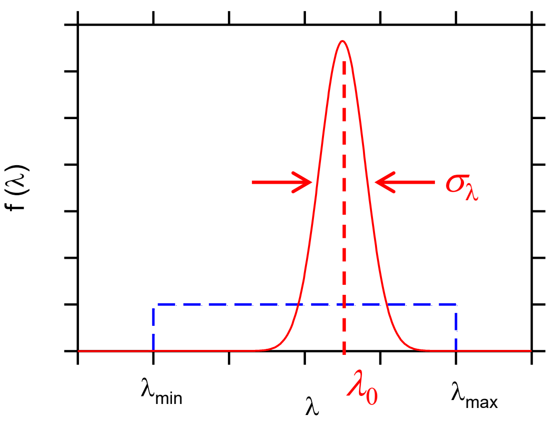
poměr apriorních pravděpodobností

poměr posteriorních pravděpodobností

$$\frac{P(A|D,I)}{P(B|D,I)} = \frac{P(D|A,I)}{P(D|B,I)}\frac{P(A,I)}{P(B,I)}$$

- věrohodnost pro hypotézu B: $P(D|B,I) = \int P(D,\lambda|B,I)\,d\lambda = \int P(D|\lambda,B,I) f(\lambda|B,I)\,d\lambda$

- apriorní hustota pravděpodobnosti parametru λ: $f(\lambda|B,I) = \begin{cases} \frac{1}{\lambda_{max}-\lambda_{min}} & \lambda \in \langle\lambda_{min},\lambda_{max}\rangle \\ 0 & \text{jinak} \end{cases}$



$$P(D|\lambda,B,I) = P(D|\lambda_0,B,I)\exp\left[-\frac{(\lambda-\lambda_0)^2}{2\sigma_\lambda^2}\right]$$

Dvě alternativní hypotézy

- hypotéza A: veličina $y = 0$

- hypotéza B: veličina y je nějaká konstanta λ

poměr věrohodností — poměr apriorních pravděpodobností

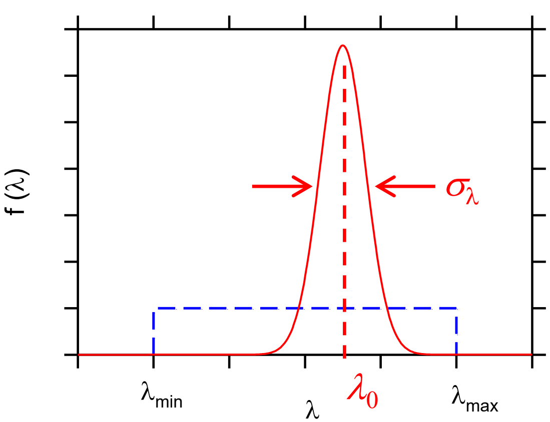
poměr posteriorních pravděpodobností

$$\frac{P(A|D,I)}{P(B|D,I)} = \frac{P(D|A,I)}{P(D|B,I)} \frac{P(A,I)}{P(B,I)}$$

- věrohodnost pro hypotézu B: $$P(D|B,I) = \frac{1}{\lambda_{max} - \lambda_{min}} \int_{\lambda_{min}}^{\lambda_{max}} P(D|\lambda, B, I)\, d\lambda$$

- po dosazení:

$$P(D|B,I) = \frac{P(D|\lambda_0, B, I)\sqrt{2\pi}\sigma_\lambda}{\lambda_{max} - \lambda_{min}}$$



$$P(D|\lambda, B, I) = P(D|\lambda_0, B, I) \exp\left[-\frac{(\lambda - \lambda_0)^2}{2\sigma_\lambda^2}\right]$$

Dvě alternativní hypotézy

- hypotéza A: veličina $y = 0$
- hypotéza B: veličina y je nějaká konstanta λ

poměr apriorních pravděpodobností

poměr věrohodností

poměr posteriorních pravděpodobností

$$\boxed{\frac{P(A|D,I)}{P(B|D,I)} = \frac{P(A,I)}{P(B,I)}\,\frac{P(D|A,I)}{P(D|\lambda_0,B,I)}\,\frac{\lambda_{max}-\lambda_{min}}{\sqrt{2\pi}\sigma_\lambda}}$$

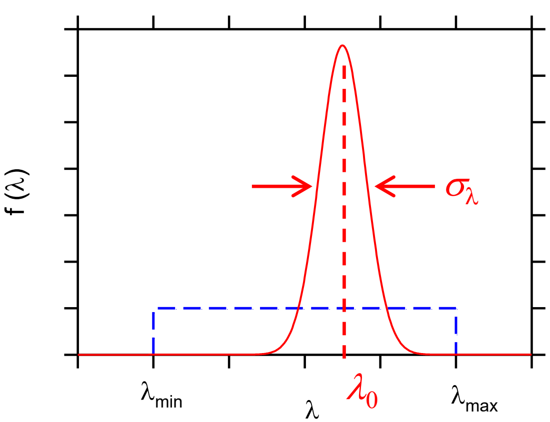
Ockhamův člen (penalizační faktor pro hypotézu B)

- pokud je $P(A,I) = P(B,I)$

$P(D|A,I) \approx P(D|\lambda_0,B,I)$

$\sigma_\lambda << \lambda_{\mathrm{max}} - \lambda_{\mathrm{min}}$

preferujeme hypotézu A



$$P(D|\lambda,B,I) = P(D|\lambda_0,B,I)\exp\left[-\frac{(\lambda-\lambda_0)^2}{2\sigma_\lambda^2}\right]$$

Dvě alternativní hypotézy

- hypotéza A: veličina $y = 0$
- hypotéza B: veličina y je nějaká konstanta λ

poměr apriorních pravděpodobností

poměr věrohodností

poměr posteriorních pravděpodobností

$$\boxed{\frac{P(A|D,I)}{P(B|D,I)} = \frac{P(A,I)}{P(B,I)} \frac{P(D|A,I)}{P(D|\lambda_0,B,I)} \frac{\lambda_{max} - \lambda_{min}}{\sqrt{2\pi}\sigma_\lambda}}$$

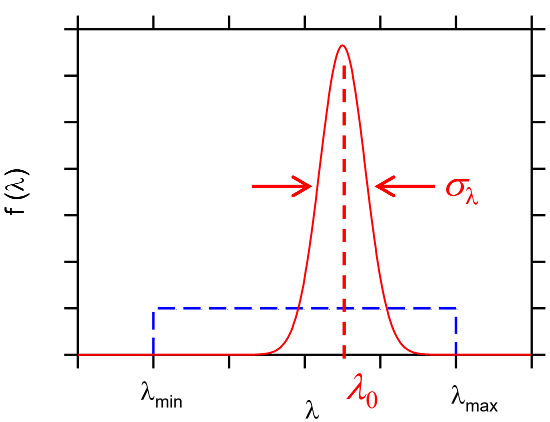
Ockhamův člen (penalizační faktor pro hypotézu B)

- pokud je $P(A,I) = P(B,I)$

$$\left.\begin{array}{l} P(D|A,I) << P(D|\lambda_0,B,I) \\ \sigma_\lambda << \lambda_{\max} - \lambda_{\min} \end{array}\right\}$$

preferujeme hypotézu B

tj. hypotéza s více parametry musí vést k výrazně lepšímu souhlasu s experimentem



$$P(D|\lambda,B,I) = P(D|\lambda_0,B,I) \exp\left[-\frac{(\lambda-\lambda_0)^2}{2\sigma_\lambda^2}\right]$$

Dvě alternativní hypotézy

- hypotéza A: veličina y je nějaká konstanta μ
- hypotéza B: veličina y je nějaká konstanta λ

poměr apriorních pravděpodobností

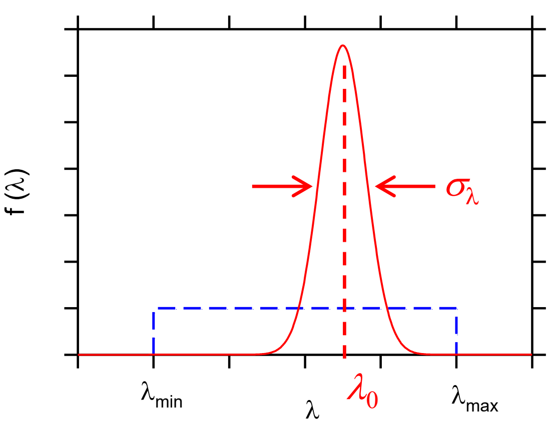
poměr věrohodností

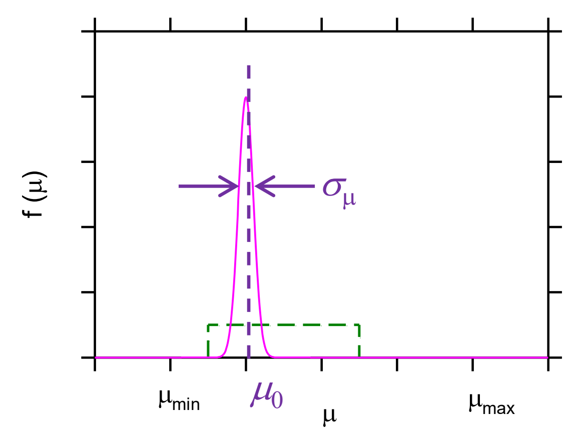
poměr posteriorních pravděpodobností

$$\frac{P(A|D,I)}{P(B|D,I)} = \frac{P(A,I)}{P(B,I)}\,\frac{P(D|\mu_0,A,I)}{P(D|\lambda_0,B,I)}\,\frac{\lambda_{max}-\lambda_{min}}{\mu_{max}-\mu_{min}}\,\frac{\sigma_\mu}{\sigma_\lambda}$$

$$P(D|\lambda,B,I) = P(D|\lambda_0,B,I)\exp\left[-\frac{(\lambda-\lambda_0)^2}{2\sigma_\lambda^2}\right]$$

$$P(D|\mu,B,I) = P(D|\mu_0,B,I)\exp\left[-\frac{(\mu-\mu_0)^2}{2\sigma_\mu^2}\right]$$





Dvě alternativní hypotézy

- hypotéza A: veličina y je nějaká konstanta μ
- hypotéza B: veličina y je nějaká konstanta λ

poměr apriorních pravděpodobností

poměr věrohodností

poměr posteriorních pravděpodobností

$$\boxed{\frac{P(A|D,I)}{P(B|D,I)} = \frac{P(A,I)}{P(B,I)} \frac{P(D|\mu_0,A,I)}{P(D|\lambda_0,B,I)} \frac{\lambda_{max}-\lambda_{min}}{\mu_{max}-\mu_{min}} \frac{\sigma_\mu}{\sigma_\lambda}}$$

- pokud je $P(A,I) = P(B,I)$

$$\left.\begin{array}{l} P(D|\mu_0,A,I) = P(D|\lambda_0,B,I) \\ \mu_{\max} - \mu_{\min} = \lambda_{\max} - \lambda_{\min} \end{array}\right\} \quad \frac{P(A|D,I)}{P(B|D,I)} = \frac{P(A,I)}{P(B,I)} \frac{P(D|\mu_0,A,I)}{P(D|\lambda_0,B,I)} \frac{\sigma_\mu}{\sigma_\lambda}$$

- pokud je $P(A,I) = P(B,I)$

$$\left.\begin{array}{l} P(D|\mu_0,A,I) = P(D|\lambda_0,B,I) \\ \sigma_\mu = \sigma_\lambda \end{array}\right\} \quad \frac{P(A|D,I)}{P(B|D,I)} = \frac{P(A,I)}{P(B,I)} \frac{P(D|\mu_0,A,I)}{P(D|\lambda_0,B,I)} \frac{\lambda_{max}-\lambda_{min}}{\mu_{max}-\mu_{min}}$$