Úvod do praktické fyziky, cvičení 9
Metoda maximální věrohodnosti

Jan Matoušek

24. 11. 2020

UNIVERZITA KARLOVA
Matematicko-fyzikální fakulta

Jak odhadnout parametry rozdělení z experimentálních dat?

- Máme sadu (nezávislých) naměřených hodnot $\boldsymbol{x} = (x_1, x_2, ..., x_n))$.
- Rozdělení x_i je popsáno známou/předpokládanou hustotou pravděpodobnosti $f(x|\boldsymbol{\theta})$.
- Parametry $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_m)$ neznáme.
- Jaký je jejich nejlepší odhad (*estimátor*) $\hat{\boldsymbol{\theta}} = (\hat{\theta_1}, \hat{\theta_2}, ..., \hat{\theta_m})$?

Vlastnosti estimátorů

konzistence: pro $n \to \infty$ konverguje $\hat{\boldsymbol{\theta}} \to \boldsymbol{\theta}$,

předpojatost: anlicky *bias* $b = E[\hat{\theta}] - \theta$, $b = 0$ zamená *nepředpojatý* (*nevychýlený*) odhad.

efektivita: *mean square error (MSE)*: $E[(\hat{\theta} - \theta)^2]$

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

Příklad: estimátory parametrů normálního rozdělení $N(\mu, \sigma)$

- Odhad očekávané hodnoty: aritmetický průměr $\overline{x}$ (viz přednáška prof. Čížka),
- Odhad rozptylu: aritmetický průměr $(x_i - \mu)^2$,

$$\hat{\mu} = \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \qquad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 \approx \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2.$$

Jak odhadnout parametry rozdělení z experimentálních dat?

- Máme sadu (nezávislých) naměřených hodnot $\boldsymbol{x} = (x_1, x_2, ..., x_n))$.
- Rozdělení x_i je popsáno známou/předpokládanou hustotou pravděpodobnosti $f(x|\boldsymbol{\theta})$.
- Parametry $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_m)$ neznáme.
- Jaký je jejich nejlepší odhad (*estimátor*) $\hat{\boldsymbol{\theta}} = (\hat{\theta_1}, \hat{\theta_2}, ..., \hat{\theta_m})$?

Vlastnosti estimátorů

konzistence: pro $n \to \infty$ konverguje $\hat{\boldsymbol{\theta}} \to \boldsymbol{\theta}$,

předpojatost: anlicky *bias* $b = E[\hat{\theta}] - \theta$, $b = 0$ zamená *nepředpojatý* (*nevychýlený*) odhad.

efektivita: *mean square error (MSE)*: $E[(\hat{\theta} - \theta)^2]$

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

Příklad: estimátory parametrů normálního rozdělení $N(\mu, \sigma)$

- Odhad očekávané hodnoty: aritmetický průměr $\overline{x}$ (viz přednáška prof. Čížka),
- Odhad rozptylu: aritmetický průměr $(x_i - \mu)^2$,

$$\hat{\mu} = \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \qquad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 \approx \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2.$$

Jak odhadnout parametry rozdělení z experimentálních dat?

- Máme sadu (nezávislých) naměřených hodnot $\boldsymbol{x} = (x_1, x_2, ..., x_n))$.
- Rozdělení x_i je popsáno známou/předpokládanou hustotou pravděpodobnosti $f(x|\boldsymbol{\theta})$.
- Parametry $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_m)$ neznáme.
- Jaký je jejich nejlepší odhad (*estimátor*) $\hat{\boldsymbol{\theta}} = (\hat{\theta_1}, \hat{\theta_2}, ..., \hat{\theta_m})$?

Vlastnosti estimátorů

konzistence: pro $n \to \infty$ konverguje $\hat{\boldsymbol{\theta}} \to \boldsymbol{\theta}$,

předpojatost: anlicky *bias* $b = E[\hat{\theta}] - \theta$, $b = 0$ zamená *nepředpojatý* (*nevychýlený*) odhad.

efektivita: *mean square error (MSE)*: $E[(\hat{\theta} - \theta)^2]$

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

Příklad: estimátory parametrů normálního rozdělení $N(\mu, \sigma)$

- Odhad očekávané hodnoty: aritmetický průměr $\overline{x}$ (viz přednáška prof. Čížka),
- Odhad rozptylu: aritmetický průměr $(x_i - \mu)^2$,

$$\hat{\mu} = \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \qquad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 \approx \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2.$$

Jak odhadnout parametry rozdělení z experimentálních dat?

Pokračování příkladu: Předpojatost estimátorů $\hat{\mu}$ a $\hat{\sigma}$ pro $N(\mu, \sigma)$

$$\hat{\mu} = \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \qquad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 \approx \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2.$$

- Aproximace – většinou neznáme skutečnou očekávanou hodnotu μ, použijeme estimátor.
- Předpojatost $b = E[\hat{\theta}] - \theta$:

$$E[\hat{\mu}] = \frac{1}{n}\sum_{i=1}^{n} E[x_i] = \mu \qquad \rightarrow b = 0.$$

$$E[\hat{\sigma}^2] = \frac{1}{n}\sum_{i=1}^{n} E[(x_i - \mu)^2] = \sigma^2 \qquad \rightarrow b = 0.$$

$$E[\hat{\sigma}^2] = E\left[\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2\right] = \frac{1}{n}\sum_{i=1}^{n} E[x_i^2] - E[\overline{x}^2] = E[x^2] - \mu^2 - E[\overline{x}^2] + \mu^2$$

$$= V[x] - V[\overline{x}] = V[x] - \frac{1}{n}V[x] = \frac{n-1}{n}\sigma^2 \qquad \rightarrow b = -\frac{\sigma^2}{n}.$$

$$\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \frac{2\overline{x}}{n}\sum_{i=1}^{n} x_i + \frac{\overline{x}^2}{n}\sum_{i=1}^{n} 1 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{x}^2$$

- Nepředpojatý odhad rozptylu (ale o něco větší EMS):

$$\hat{s}^2 = \frac{n}{n-1}\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2.$$

Jak odhadnout parametry rozdělení z experimentálních dat?

Pokračování příkladu: Předpojatost estimátorů $\hat{\mu}$ a $\hat{\sigma}$ pro $N(\mu, \sigma)$

$$\hat{\mu} = \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \qquad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 \approx \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2.$$

- Aproximace – většinou neznáme skutečnou očekávanou hodnotu μ, použijeme estimátor.
- Předpojatost $b = E[\hat{\theta}] - \theta$:

$$E[\hat{\mu}] = \frac{1}{n}\sum_{i=1}^{n} E[x_i] = \mu \qquad \rightarrow b = 0.$$

$$E[\hat{\sigma}^2] = \frac{1}{n}\sum_{i=1}^{n} E[(x_i - \mu)^2] = \sigma^2 \qquad \rightarrow b = 0.$$

$$E[\hat{\sigma}^2] = E\left[\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2\right] = \frac{1}{n}\sum_{i=1}^{n} E[x_i^2] - E[\overline{x}^2] = E[x^2] - \mu^2 - E[\overline{x}^2] + \mu^2$$

$$= V[x] - V[\overline{x}] = V[x] - \frac{1}{n}V[x] = \frac{n-1}{n}\sigma^2 \qquad \rightarrow b = -\frac{\sigma^2}{n}.$$

$$\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \frac{2\overline{x}}{n}\sum_{i=1}^{n} x_i + \frac{\overline{x}^2}{n}\sum_{i=1}^{n} 1 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{x}^2$$

- Nepředpojatý odhad rozptylu (ale o něco větší EMS):

$$\hat{s}^2 = \frac{n}{n-1}\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2.$$

Pokračování příkladu: Předpojatost estimátorů $\hat{\mu}$ a $\hat{\sigma}$ pro $N(\mu, \sigma)$

$$\hat{\mu} = \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \qquad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 \approx \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2.$$

- Aproximace – většinou neznáme skutečnou očekávanou hodnotu μ, použijeme estimátor.
- Předpojatost $b = E[\hat{\theta}] - \theta$:

$$E[\hat{\mu}] = \frac{1}{n}\sum_{i=1}^{n} E[x_i] = \mu \qquad \rightarrow b = 0.$$

$$E[\hat{\sigma}^2] = \frac{1}{n}\sum_{i=1}^{n} E[(x_i - \mu)^2] = \sigma^2 \qquad \rightarrow b = 0.$$

$$E[\hat{\sigma}^2] = E\left[\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2\right] = \frac{1}{n}\sum_{i=1}^{n} E[x_i^2] - E[\overline{x}^2] = E[x^2] - \mu^2 - E[\overline{x}^2] + \mu^2$$

$$= V[x] - V[\overline{x}] = V[x] - \frac{1}{n}V[x] = \frac{n-1}{n}\sigma^2 \qquad \rightarrow b = -\frac{\sigma^2}{n}.$$

$$\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \frac{2\overline{x}}{n}\sum_{i=1}^{n} x_i + \frac{\overline{x}^2}{n}\sum_{i=1}^{n} 1 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{x}^2$$

- Nepředpojatý odhad rozptylu (ale o něco větší EMS):

$$\hat{s}^2 = \frac{n}{n-1}\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2.$$

Jak odhadnout parametry rozdělení z experimentálních dat?

Pokračování příkladu: Předpojatost estimátorů $\hat{\mu}$ a $\hat{\sigma}$ pro $N(\mu, \sigma)$

$$\hat{\mu} = \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \qquad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 \approx \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2.$$

- Aproximace – většinou neznáme skutečnou očekávanou hodnotu μ, použijeme estimátor.
- Předpojatost $b = E[\hat{\theta}] - \theta$:

$$E[\hat{\mu}] = \frac{1}{n}\sum_{i=1}^{n} E[x_i] = \mu \qquad \rightarrow b = 0.$$

$$E[\hat{\sigma}^2] = \frac{1}{n}\sum_{i=1}^{n} E[(x_i - \mu)^2] = \sigma^2 \qquad \rightarrow b = 0.$$

$$E[\hat{\sigma}^2] = E\left[\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2\right] = \frac{1}{n}\sum_{i=1}^{n} E[x_i^2] - E[\overline{x}^2] = E[x^2] - \mu^2 - E[\overline{x}^2] + \mu^2$$

$$= V[x] - V[\overline{x}] = V[x] - \frac{1}{n}V[x] = \frac{n-1}{n}\sigma^2 \qquad \rightarrow b = -\frac{\sigma^2}{n}.$$

$$\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \frac{2\overline{x}}{n}\sum_{i=1}^{n} x_i + \frac{\overline{x}^2}{n}\sum_{i=1}^{n} 1 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{x}^2$$

- Nepředpojatý odhad rozptylu (ale o něco větší EMS):

$$\hat{s}^2 = \frac{n}{n-1}\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2.$$

Jak odhadnout parametry rozdělení z experimentálních dat?

Pokračování příkladu: Předpojatost estimátorů $\hat{\mu}$ a $\hat{\sigma}$ pro $N(\mu, \sigma)$

$$\hat{\mu} = \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \qquad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 \approx \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2.$$

- Aproximace – většinou neznáme skutečnou očekávanou hodnotu μ, použijeme estimátor.
- Předpojatost $b = E[\hat{\theta}] - \theta$:

$$E[\hat{\mu}] = \frac{1}{n}\sum_{i=1}^{n} E[x_i] = \mu \qquad \rightarrow b = 0.$$

$$E[\hat{\sigma}^2] = \frac{1}{n}\sum_{i=1}^{n} E[(x_i - \mu)^2] = \sigma^2 \qquad \rightarrow b = 0.$$

$$E[\hat{\sigma}^2] = E\left[\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2\right] = \frac{1}{n}\sum_{i=1}^{n} E[x_i^2] - E[\overline{x}^2] = E[x^2] - \mu^2 - E[\overline{x}^2] + \mu^2$$

$$= V[x] - V[\overline{x}] = V[x] - \frac{1}{n}V[x] = \frac{n-1}{n}\sigma^2 \qquad \rightarrow b = -\frac{\sigma^2}{n}.$$

$$\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \frac{2\overline{x}}{n}\sum_{i=1}^{n} x_i + \frac{\overline{x}^2}{n}\sum_{i=1}^{n} 1 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{x}^2$$

- Nepředpojatý odhad rozptylu (ale o něco větší EMS):

$$\hat{s}^2 = \frac{n}{n-1}\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2.$$

Metoda maximální věrohodnosti (maximum likelihood)

- Máme sadu (nezávislých) naměřených hodnot $\boldsymbol{x} = (x_1, x_2, ..., x_n))$.
- Rozdělení x_i je popsáno známou/předpokládanou hustotou pravděpodobnosti $f(x|\boldsymbol{\theta})$.
- Parametry $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_m)$ neznáme.
- Pravděpodobnost pozorování hodnot $x_1, x_2, ...x_n$ závisí na parametrech $\boldsymbol{\theta}$,

$$P(\boldsymbol{x}, \mathrm{d}x|\boldsymbol{\theta}) = \prod_{i=1}^{n} P\big(x \in (x_i, x_i + \mathrm{d}x)|\boldsymbol{\theta}\big) = \prod_{i=1}^{n} f(x_i|\boldsymbol{\theta})\mathrm{d}x.$$

- Sada $\boldsymbol{x}$ kterou jsme pozorovali musí být pravděpodobná...

- Definujeme *věrohodnostní* (*likelihood*) funkci, která je funkcí $\boldsymbol{\theta}$ a má $\boldsymbol{x}$ jako parametry,

$$L(\boldsymbol{\theta}|\boldsymbol{x}) = \prod_{i=1}^{n} f(x_i|\boldsymbol{\theta}).$$

- Nejlepší odhad $\hat{\boldsymbol{\theta}}$ je ten, který maximalizuje L – řešíme soustavu m rovnic

$$\frac{\partial L}{\partial \theta_k} = 0, \qquad k = 1, 2, ..., m$$

- Často je výhodnější maximalizovat logaritmus L (součin $\rightarrow$ suma)

$$\ln L(\boldsymbol{\theta}|\boldsymbol{x}) = \sum_{i=1}^{n} \ln f(x_i|\boldsymbol{\theta}), \qquad \frac{\partial \ln L}{\partial \theta_k} = 0, \qquad k = 1, 2, ..., m$$

Metoda maximální věrohodnosti (maximum likelihood)

- Máme sadu (nezávislých) naměřených hodnot $\boldsymbol{x} = (x_1, x_2, ..., x_n))$.
- Rozdělení x_i je popsáno známou/předpokládanou hustotou pravděpodobnosti $f(x|\boldsymbol{\theta})$.
- Parametry $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_m)$ neznáme.
- Pravděpodobnost pozorování hodnot $x_1, x_2, ...x_n$ závisí na parametrech $\boldsymbol{\theta}$,

$$P(\boldsymbol{x}, \mathrm{d}x|\boldsymbol{\theta}) = \prod_{i=1}^{n} P\big(x \in (x_i, x_i + \mathrm{d}x)|\boldsymbol{\theta}\big) = \prod_{i=1}^{n} f(x_i|\boldsymbol{\theta})\mathrm{d}x.$$

- Sada $\boldsymbol{x}$ kterou jsme pozorovali musí být pravděpodobná...

- Definujeme *věrohodnostní (likelihood)* funkci, která je funkcí $\boldsymbol{\theta}$ a má $\boldsymbol{x}$ jako parametry,

$$L(\boldsymbol{\theta}|\boldsymbol{x}) = \prod_{i=1}^{n} f(x_i|\boldsymbol{\theta}).$$

- Nejlepší odhad $\hat{\boldsymbol{\theta}}$ je ten, který maximalizuje L – řešíme soustavu m rovnic

$$\frac{\partial L}{\partial \theta_k} = 0, \qquad k = 1, 2, ..., m$$

- Často je výhodnější maximalizovat logaritmus L (součin $\to$ suma)

$$\ln L(\boldsymbol{\theta}|\boldsymbol{x}) = \sum_{i=1}^{n} \ln f(x_i|\boldsymbol{\theta}), \qquad \frac{\partial \ln L}{\partial \theta_k} = 0, \qquad k = 1, 2, ..., m$$

Příklad: Odhad parametrů normálního rozdělení

- Maximální věrohodnost dostaneme pro $\hat{\mu} = \overline{x}$, $\hat{\sigma}^2 = \frac{1}{n} \sum_i^n (x_i - \overline{x})^2$.
- Odvození viz přednáška prof. Čížka.
- Víme: nevychýlený odhad $\hat{s}^2 = \frac{1}{n-1} \sum_i^n (x_i - \overline{x})^2$.
- Naměřené hodnoty: $\boldsymbol{x} = (51.99, 37.65, 49.82, 41.70, 50.24)$
- Spočítáme

$$\hat{\mu} = \frac{1}{5} \sum_{i=1}^{5} x_i = 46.28, \qquad \hat{\sigma}^2 = \frac{1}{5} \sum_{i=1}^{5} (x_i - 46.28)^2 = 31.25, \qquad \hat{s}^2 = \frac{1}{4} \sum_{i=1}^{5} (x_i - 46.28)^2 = 39.06.$$

- Odhad šířky rozdělení „celé populace“ a současně odhad chyby jednoho měření $\hat{\sigma} = 5.60$, nevychýlený $\hat{s} = 6.25$.
- Chyba aritmetického průměru (viz cvičení 8):

$$\sigma_{\overline{x}} = \frac{\hat{s}}{\sqrt{n}} = 2.80.$$

- Zápis výsledku měření:

$$x = \overline{x} \pm \sigma_{\overline{x}} = 46 \pm 3.$$

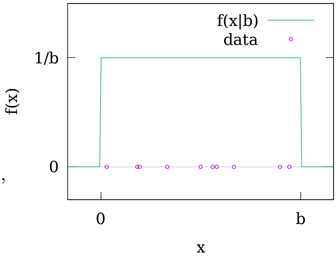
Příklad: Odhad parametru rovnoměrného rozdělení

- Máme nezávislá měření $x_1, x_2, ..., x_n \in U(0, b)$.
- Hustota pravděpodobnosti

$$f(x|b) = \begin{cases} \frac{1}{b} & \text{pro}\, 0 < x < b, \\ 0 & \text{jinde.} \end{cases}$$

- Věrohodnost:

$$L(b|\boldsymbol{x}) = \begin{cases} \prod_{i=1}^{n} \frac{1}{b} = b^{-n} & \text{pro}\, x_i < b \qquad \forall i = 1, 2, ..., n, \\ 0 & \text{jinak.} \end{cases}$$



- Pozn.: pro diskrétní x: „German tank problem“.
- Věrohodnost je maximální pokud b je minimální a současně všechna $x_i < b$.
- Tedy $\hat{b} = \max x_i$.
- Ovšem je to odhad vychýlený (vždy bude $\hat{b} < b$). Předpojatost:

$$E[\max x_i] = \int_0^b x f_{\max}(x) \mathrm{d}x = \int_0^b \frac{n}{b^n} x^n \mathrm{d}x = \left[\frac{n}{n+1} \frac{x^{n+1}}{b^n} \right]_0^b = \frac{n}{n+1} b$$

Rozdělení pravděpodobnosti pro maximum:

$$F_{\max}(x) = P(x_{\max} < x) = P(x_i < x, \forall i) = \prod_{i=1}^{n} P(x_i < x) = \prod_{i=1}^{n} \frac{x}{b} = \left(\frac{x}{b} \right)^n \qquad \text{pro}\, 0 < x < b.$$

$$f_{\max}(x) = \frac{\mathrm{d}F_{\max}}{\mathrm{d}x} = \frac{n}{b^n} x^{n-1} \qquad \text{pro}\, 0 < x < b.$$

- Nepředpojatý odhad (*unbiased estimator*): $\hat{b}_{\mathrm{UE}} = \frac{n+1}{n} \max x_i$

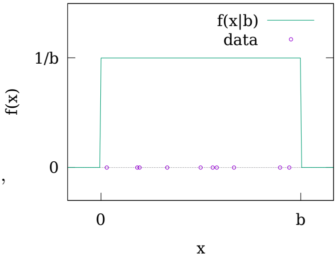
Příklad: Odhad parametru rovnoměrného rozdělení

- Máme nezávislá měření $x_1, x_2, ..., x_n \in U(0, b)$.
- Hustota pravděpodobnosti

$$f(x|b) = \begin{cases} \frac{1}{b} & \text{pro}\, 0 < x < b, \\ 0 & \text{jinde.} \end{cases}$$

- Věrohodnost:

$$L(b|\boldsymbol{x}) = \begin{cases} \prod_{i=1}^{n} \frac{1}{b} = b^{-n} & \text{pro}\, x_i < b \qquad \forall i = 1, 2, ..., n, \\ 0 & \text{jinak.} \end{cases}$$



- Pozn.: pro diskrétní x: „German tank problem“.
- Věrohodnost je maximální pokud b je minimální a současně všechna $x_i < b$.
- Tedy $\hat{b} = \max x_i$.
- Ovšem je to odhad vychýlený (vždy bude $\hat{b} < b$). Předpojatost:

$$E[\max x_i] = \int_0^b x f_{\max}(x) \mathrm{d}x = \int_0^b \frac{n}{b^n} x^n \mathrm{d}x = \left[\frac{n}{n+1} \frac{x^{n+1}}{b^n} \right]_0^b = \frac{n}{n+1} b$$

Rozdělení pravděpodobnosti pro maximum:

$$F_{\max}(x) = P(x_{\max} < x) = P(x_i < x, \forall i) = \prod_{i=1}^{n} P(x_i < x) = \prod_{i=1}^{n} \frac{x}{b} = \left(\frac{x}{b} \right)^n \qquad \text{pro}\, 0 < x < b.$$

$$f_{\max}(x) = \frac{\mathrm{d}F_{\max}}{\mathrm{d}x} = \frac{n}{b^n} x^{n-1} \qquad \text{pro}\, 0 < x < b.$$

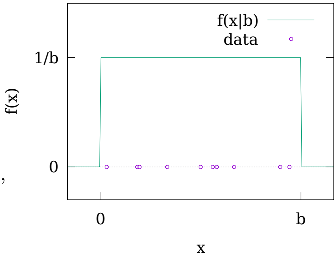
- Nepředpojatý odhad (*unbiased estimator*): $\hat{b}_{\mathrm{UE}} = \frac{n+1}{n} \max x_i$

- Máme nezávislá měření $x_1, x_2, ..., x_n \in U(0, b)$.
- Hustota pravděpodobnosti

$$f(x|b) = \begin{cases} \frac{1}{b} & \text{pro}\, 0 < x < b, \\ 0 & \text{jinde.} \end{cases}$$

- Věrohodnost:

$$L(b|\boldsymbol{x}) = \begin{cases} \prod_{i=1}^{n} \frac{1}{b} = b^{-n} & \text{pro}\, x_i < b \qquad \forall i = 1, 2, ..., n, \\ 0 & \text{jinak.} \end{cases}$$



- Pozn.: pro diskrétní x: „German tank problem“.
- Věrohodnost je maximální pokud b je minimální a současně všechna $x_i < b$.
- Tedy $\hat{b} = \max x_i$.
- Ovšem je to odhad vychýlený (vždy bude $\hat{b} < b$). Předpojatost:

$$E[\max x_i] = \int_0^b x f_{\max}(x) \mathrm{d}x = \int_0^b \frac{n}{b^n} x^n \mathrm{d}x = \left[\frac{n}{n+1} \frac{x^{n+1}}{b^n} \right]_0^b = \frac{n}{n+1} b$$

Rozdělení pravděpodobnosti pro maximum:

$$F_{\max}(x) = P(x_{\max} < x) = P(x_i < x, \forall i) = \prod_{i=1}^{n} P(x_i < x) = \prod_{i=1}^{n} \frac{x}{b} = \left(\frac{x}{b} \right)^n \qquad \text{pro}\, 0 < x < b.$$

$$f_{\max}(x) = \frac{\mathrm{d}F_{\max}}{\mathrm{d}x} = \frac{n}{b^n} x^{n-1} \qquad \text{pro}\, 0 < x < b.$$

- Nepředpojatý odhad (*unbiased estimator*): $\hat{b}_{\mathrm{UE}} = \frac{n+1}{n} \max x_i$

Příklad: Odhad parametru rovnoměrného rozdělení

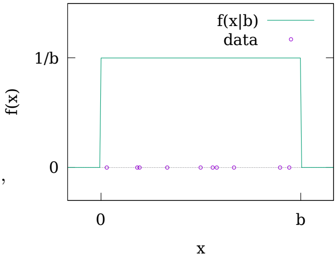
- Máme nezávislá měření $x_1, x_2, ..., x_n \in U(0, b)$.
- Hustota pravděpodobnosti

$$f(x|b) = \begin{cases} \frac{1}{b} & \text{pro}\, 0 < x < b, \\ 0 & \text{jinde.} \end{cases}$$

- Věrohodnost:

$$L(b|\boldsymbol{x}) = \begin{cases} \prod_{i=1}^{n} \frac{1}{b} = b^{-n} & \text{pro}\, x_i < b \qquad \forall i = 1, 2, ..., n, \\ 0 & \text{jinak.} \end{cases}$$

- Pozn.: pro diskrétní x: „German tank problem“.



- Věrohodnost je maximální pokud b je minimální a současně všechna $x_i < b$.
- Tedy $\hat{b} = \max x_i$.
- Ovšem je to odhad vychýlený (vždy bude $\hat{b} < b$). Předpojatost:

$$E[\max x_i] = \int_0^b x f_{\max}(x) \mathrm{d}x = \int_0^b \frac{n}{b^n} x^n \mathrm{d}x = \left[\frac{n}{n+1} \frac{x^{n+1}}{b^n} \right]_0^b = \frac{n}{n+1} b$$

Rozdělení pravděpodobnosti pro maximum:

$$F_{\max}(x) = P(x_{\max} < x) = P(x_i < x, \forall i) = \prod_{i=1}^{n} P(x_i < x) = \prod_{i=1}^{n} \frac{x}{b} = \left(\frac{x}{b}\right)^n \qquad \text{pro}\, 0 < x < b.$$

$$f_{\max}(x) = \frac{\mathrm{d}F_{\max}}{\mathrm{d}x} = \frac{n}{b^n} x^{n-1} \qquad \text{pro}\, 0 < x < b.$$

- Nepředpojatý odhad (*unbiased estimator*): $\hat{b}_{\mathrm{UE}} = \frac{n+1}{n} \max x_i$

Příklad: Odhad parametru rovnoměrného rozdělení

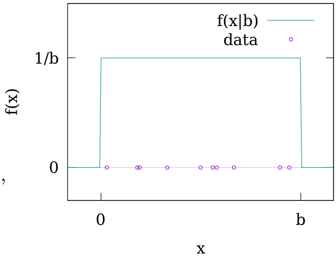
- Máme nezávislá měření $x_1, x_2, ..., x_n \in U(0, b)$.
- Hustota pravděpodobnosti

$$f(x|b) = \begin{cases} \frac{1}{b} & \text{pro}\, 0 < x < b, \\ 0 & \text{jinde.} \end{cases}$$

- Věrohodnost:

$$L(b|\boldsymbol{x}) = \begin{cases} \prod_{i=1}^{n} \frac{1}{b} = b^{-n} & \text{pro}\, x_i < b \qquad \forall i = 1, 2, ..., n, \\ 0 & \text{jinak.} \end{cases}$$

- Pozn.: pro diskrétní x: „German tank problem“.



- Věrohodnost je maximální pokud b je minimální a současně všechna $x_i < b$.
- Tedy $\hat{b} = \max x_i$.
- Ovšem je to odhad vychýlený (vždy bude $\hat{b} < b$). Předpojatost:

$$E[\max x_i] = \int_0^b x f_{\max}(x)\mathrm{d}x = \int_0^b \frac{n}{b^n} x^n \mathrm{d}x = \left[\frac{n}{n+1}\frac{x^{n+1}}{b^n}\right]_0^b = \frac{n}{n+1} b$$

Rozdělení pravděpodobnosti pro maximum:

$$F_{\max}(x) = P(x_{\max} < x) = P(x_i < x, \forall i) = \prod_{i=1}^{n} P(x_i < x) = \prod_{i=1}^{n} \frac{x}{b} = \left(\frac{x}{b}\right)^n \qquad \text{pro}\, 0 < x < b.$$

$$f_{\max}(x) = \frac{\mathrm{d}F_{\max}}{\mathrm{d}x} = \frac{n}{b^n} x^{n-1} \qquad \text{pro}\, 0 < x < b.$$

- Nepředpojatý odhad (*unbiased estimator*): $\hat{b}_{\mathrm{UE}} = \frac{n+1}{n} \max x_i$

Příklad: Numerický odhad parametrů Breitova–Wignerova rozdělení

- Máme nezávislá měření $\boldsymbol{x} = (x_1, x_2, ..., x_n)$, x_i mají Breitovo–Wignerovo rozdělění

$$f(x|\lambda, \xi) = \frac{1}{\pi} \frac{\lambda}{\lambda^2 + (x - \xi)^2}.$$

- Chceme odhadnout λ a ξ.
- Logaritmus věrohodnosti:

$$\ln L(\lambda, \xi|\boldsymbol{x}) = \sum_{i=1}^{n} \ln \left(\frac{1}{\pi} \frac{\lambda}{\lambda^2 + (x_i - \xi)^2} \right) = -n \ln \pi + n \ln \lambda - \sum_{i=1}^{n} \ln \left[\lambda^2 + (x_i - \xi)^2 \right]$$

- Analytická maximalizace pomocí derivací dává rovnice

$$0 = \frac{\partial \ln L}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} \frac{2\lambda}{\lambda^2 + (x_i - \xi)^2}, \qquad 0 = \frac{\partial \ln L}{\partial \xi} = -\sum_{i=1}^{n} \frac{2(x_i - \xi)}{\lambda^2 + (x_i - \xi)^2},$$

  které není snadné vyřešit.
- Můžeme hledat maximum numericky: `upf_cv9_BW.cc`

Příklad: Numerický odhad parametrů Breitova–Wignerova rozdělení

- Máme nezávislá měření $\boldsymbol{x} = (x_1, x_2, ..., x_n)$, x_i mají Breitovo–Wignerovo rozdělení

$$f(x|\lambda, \xi) = \frac{1}{\pi} \frac{\lambda}{\lambda^2 + (x - \xi)^2}.$$

- Chceme odhadnout λ a ξ.
- Logaritmus věrohodnosti:

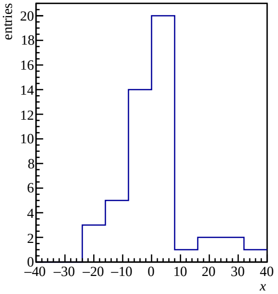
$$\ln L(\lambda, \xi|\boldsymbol{x}) = \sum_{i=1}^{n} \ln \left(\frac{1}{\pi} \frac{\lambda}{\lambda^2 + (x_i - \xi)^2} \right) = -n \ln \pi + n \ln \lambda - \sum_{i=1}^{n} \ln \left[\lambda^2 + (x_i - \xi)^2 \right]$$

- Analytická maximalizace pomocí derivací dává rovnice

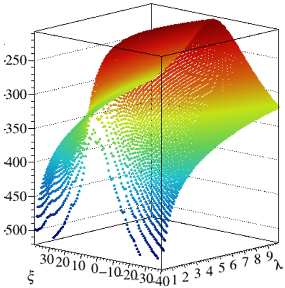
$$0 = \frac{\partial \ln L}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} \frac{2\lambda}{\lambda^2 + (x_i - \xi)^2}, \qquad 0 = \frac{\partial \ln L}{\partial \xi} = -\sum_{i=1}^{n} \frac{2(x_i - \xi)}{\lambda^2 + (x_i - \xi)^2},$$

  které není snadné vyřešit.
- Můžeme hledat maximum numericky: `upf_cv9_BW.cc`

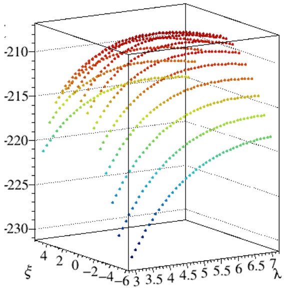
Příklad: Numerický odhad parametrů Breitova–Wignerova rozdělení



Histogram naměřených hodnot.



Logaritmus věrohodnosti
$\ln L(\lambda, \xi|\boldsymbol{x})$.



Přiblížení kolem maxima.

- 50 čísel s Breitovým–Wignerovým rozdělením $\boldsymbol{x}$ (pro představu: histogram vlevo).
- Jednoduchá metoda minimalizace:
  - $\ln L(\lambda, \xi|\boldsymbol{x})$ spočítáno ve 100×100 bodech $0.1 < \lambda < 10$, $-40 < \xi < 40$.
  - $\hat{\lambda} = \lambda_j, \hat{\xi} = \xi_k$, že $\ln L(\lambda_j, \xi_k|\boldsymbol{x})$ je maximální.
- Existují lepší metody maximalizace, např.:
  - Newtonova metoda (aproximace 2. řádem Taylorovy řady – paraboloidem).
  - Výstup po nejprudším gradientu.
  - Levenbergův–Marquardtův algoritmus (kombinace prvních dvou).
- Statistickou chybu odhadu lze zjistit z šířky maxima.