M5120 – cvičení

Odhady parametrů metodou maximální věrohodnosti a momentovou metodou

Ondřej Pokora (pokora@math.muni.cz)

Ústav matematiky a statistiky, Přírodovědecká fakulta, Masarykova univerzita, Brno

(podzim 2015)

Maximálně věrohodné odhady

Náhodný výběr $X_1, \ldots, X_n$ rozsahu n z rozdělení pravděpodobnosti P:

- $X \sim P$ $(i = 1, \ldots, n)$
- $X_1, \ldots, X_n$ jsou stochasticky nezávislé
- Co to znamená pro vztah mezi simultánní a marginální hustotou pravděpodobnosti $f(x)$ (pravděpodobnostní funkci $p(x)$) **?**

Rozdělení pravděpodobnosti závislé na parametru (parametrech) θ:

- $f(x), p(x)$ jako funkce proměnné $\theta \Rightarrow L(\theta)$
- **Věrohodnostní funkce** $L(\theta)$ a **logaritmická** věrohodnostní funkce $\ell(\theta)$:

$$L(\theta) = L(\theta; x_1, \ldots, x_n) = \prod_{i=1}^{n} f(x_i; \theta) = \prod_{i=1}^{n} p(x_i; \theta)$$

$$\ell(\theta) = \ell(\theta; x_1, \ldots, x_n) = \ln L(\theta; x_1, \ldots, x_n) = \sum_{i=1}^{n} \ln f(x_i; \theta) = \sum_{i=1}^{n} \ln p(x_i; \theta)$$

- Jak odhadnout θ ze znalosti $X_1, \ldots, X_n$?

Maximálně věrohodné odhady

Myšlenka: parametr θ odhadneme hodnotou, která je při daném náhodném výběru ze známého rozdělení pravděpodobnosti nejvíce pravděpodobná.

Maximálně věrohodný odhad (MLE = maximum likelihood estimator) $\widehat{\theta}_{\mathsf{ML}}$ parametru θ se získá maximalizací věrohodnostní funkce $L(\theta)$:

$$\widehat{\theta}_{\mathsf{ML}} : L(\theta; x_1, \ldots, x_n) \to \max_{\theta}, \quad \text{resp.} \quad \ell(\theta; x_1, \ldots, x_n) \to \max_{\theta}$$

To znamená najít stacionární bod funkce $\ell(\theta)$ vzhledem k θ,

$$\frac{\partial}{\partial\theta}\ell(\theta) = 0 \qquad \text{(věrohodnostní rovnice)},$$

a ověřit 2. diferenciál, resp. derivaci,

$$\left.\frac{\partial^2}{\partial\theta^2}\ell(\theta)\right|_{\theta=\widehat{\theta}_{\mathsf{ML}}} < 0 \ .$$

Poznámka: v případě věktoru parametrů $\boldsymbol{\theta}$ řešíme soustavu věrohodnostních rovnic pro $\boldsymbol{\theta}$ z 1. derivací a ověřujeme negativní definitnost matice 2. derivací.

Odhady parametrů metodou momentů

(Teoretické) momenty rozdělení pravděpodobnosti náhodné veličiny X

$$M_p(\theta) = \mathrm{E}(X^p), \qquad p = 1, 2, \dots$$

závisí na neznámých parametru (parametrech), které(ý) chceme odhadnout.

Výběrové momenty počítáme z realizace náhodného výběru:

$$m_p = \frac{1}{n} \sum_{i=1}^{n} X_i^p, \qquad p = 1, 2, \dots \ .$$

Položíme

$$M_p(\theta) = m_p, \qquad p = 1, 2, \dots \ ,$$

čímž obrdžíme soustavu rovnic pro parametr(y) θ.

V praxi volíme tolik rovnic (taková p), abychom jejich řešením (algebraickým, či numerickým) byli schopni spočítat odhady všech neznámých parametrů.

Příklad 1

X je diskrétní náhodná veličina s pravděpodobnostní funcí $p(x;\theta)$

x	0	1	2	3
$p(x;\theta)$	$\frac{2}{3}\theta$	$\frac{1}{3}\theta$	$\frac{2}{3}(1-\theta)$	$\frac{1}{3}(1-\theta)$

závislou na parametru θ.

Pro náhodný výběr $\boldsymbol{X} = (3,0,2,1,3,2,1,0,2,1)$ rozsahu 10 spočítejte

- *maximálně věrohodný odhad $\widehat{\theta}_{\mathrm{ML}}$ pomocí maximalizace $L(\theta)$,*
- *maximálně věrohodný odhad $\widehat{\theta}_{\mathrm{ML}}$ pomocí maximalizace $\ell(\theta)$,*
- *odhad $\widehat{\theta}_{\mathrm{M}}$ momentovou metodou.*

Analogické odhady pak spočítejte i pro náhodný výběr $\boldsymbol{X}^* = (3,0,2,1,3,2,1,0,2,0)$. Výsledky porovnejte.

řešení

$\mathrm{E}X = \frac{7}{3} - 2\,\theta$, $\boldsymbol{X}$: $\widehat{\theta}_{\mathrm{ML}} = \frac{1}{2}$, $\widehat{\theta}_{\mathrm{M}} = \frac{5}{12}$; $\boldsymbol{X}^*$: $\widehat{\theta}_{\mathrm{ML}} = \frac{1}{2}$, $\widehat{\theta}_{\mathrm{M}} = \frac{7}{15}$

Příklad 2

Pro náhodný výběr $\boldsymbol{X} = (X_1, \dots, X_n)$ *z binomického rozdělení pravděpodobnosti* $\mathrm{Bi}(m, \theta)$ *s daným* $m \in \mathbb{N}$ *odvoďte odhady neznámého parametru* $\theta \in (0; 1)$ *metodou maximální věrohodnosti a momentovou metodou.*
Odhady pak vyčíslete pro náhodný výběr $\boldsymbol{X} = (2, 3, 3, 2, 2, 3, 3, 4, 1, 3)$ *a* $m = 8$.

řešení

$$p(x; \theta) = \begin{cases} \binom{m}{x}\theta^x (1-\theta)^{m-x}, & x = 1, \dots, m, \\ 0, & \text{jinak.} \end{cases}$$

$$\ell(\theta) = \sum_{i=1}^{n} \ln \binom{m}{x_i} + \ln\theta \sum_{i=1}^{n} x_i + \ln(1-\theta) \sum_{i=1}^{n} (m - x_i)$$

$$\frac{\partial \ell}{\partial \theta} = \frac{1}{\theta} \sum_{i=1}^{n} x_i - \frac{1}{1-\theta} \sum_{i=1}^{n} (m - x_i)$$

věrohodnostní rovnice: $\theta \sum_{i=1}^{n} (m - x_i) = (1-\theta) \sum_{i=1}^{n} x_i \;\Rightarrow\; \widehat{\theta}_{\mathrm{ML}} = \frac{1}{m}\overline{X} = \frac{13}{40}$

$$\mathrm{E}X = m\,\theta \;\Rightarrow\; \widehat{\theta}_{\mathrm{M}} = \frac{1}{m}\overline{X}$$

Příklad 3

Pro náhodný výběr $\boldsymbol{X} = (X_1, \ldots, X_n)$ z rovnoměrného spojitého rozdělení pravděpodobnosti $\mathrm{Rs}\left([0;\theta]\right)$ odvoďte odhad neznámého parametru $\theta > 0$ metodou maximální věrohodnosti a momentovou metodou.

řešení

$$f(x;\theta) = \begin{cases} \frac{1}{\theta}, & 0 \leq x \leq \theta, \\ 0, & \text{jinak.} \end{cases}$$

$$L(\theta) = \begin{cases} \theta^{-n}, & 0 \leq x_i \leq \theta, \quad i = 1, \ldots, n, \\ 0, & \text{jinak.} \end{cases}$$

$$\widehat{\theta}_{\mathrm{ML}} = \max\left\{X_1, \ldots, X_n\right\}$$

Tento ML-odhad ale není vhodný, je totiž vychýlený, konkrétně podhodnocený. Proč? Odhad momentovou metodou je v tomto případě nestranný:

$$\mathrm{E}X = \frac{\theta}{2}, \qquad \widehat{\theta}_{\mathrm{M}} = 2\,\overline{X}$$

Příklad 4

Pro náhodný výběr $\boldsymbol{X} = (X_1, \ldots, X_n)$ z rovnoměrného spojitého rozdělení pravděpodobnosti $\mathrm{Rs}\left((0;\theta)\right)$ odvoďte odhad neznámého parametru $\theta > 0$ metodou maximální věrohodnosti a momentovou metodou.

řešení

$$f(x;\theta) = \begin{cases} \frac{1}{\theta}, & 0 < x < \theta, \\ 0, & \text{jinak.} \end{cases}$$

$$L(\theta) = \begin{cases} \theta^{-n}, & 0 < x_i < \theta, \quad i = 1, \ldots, n, \\ 0, & \text{jinak.} \end{cases}$$

ML-odhad zde neexistuje. Proč?
Odhad momentovou metodou vyjde stejně jako v předchozím případ.

Příklad 5

Pro náhodný výběr $\boldsymbol{X} = (X_1, \ldots, X_n)$ z rovnoměrného spojitého rozdělení pravděpodobnosti $\mathrm{Rs}\left([\theta; \theta+1]\right)$ odvoďte odhad neznámého parametru $\theta \in \mathbb{R}$ metodou maximální věrohodnosti a momentovou metodou.

řešení

$$f(x;\theta) = \begin{cases} 1, & \theta \le x \le \theta + 1, \\ 0, & \text{jinak.} \end{cases}$$

$$L(\theta) = \begin{cases} 1, & \theta \le x_i \le \theta + 1, \quad i = 1, \ldots, n, \\ 0, & \text{jinak.} \end{cases}$$

Odtud lze pro hledaný odhad nalézt podmínku

$$\max\{X_1 \ldots . X_n\} - 1 \le \widehat{\theta}_{\mathrm{ML}} \le \min\{X_1 \ldots . X_n\},$$

ML-odhad tedy v tomto případě není určen jednoznačně. Proč?
Odhad momentovou metodou je v tomto případě nestranný:

$$\mathrm{E}X = \theta + \frac{1}{2}, \qquad \widehat{\theta}_{\mathrm{M}} = \overline{X} - \frac{1}{2}$$

Příklad 6

Pro náhodný výběr $\boldsymbol{X} = (X_1, \ldots, X_n)$ *z geometrického rozdělení* $\mathrm{Ge}(\theta)$, *kde* X *značí počet neúspěchů před prvním úspěchem, odvoďte odhady neznámého parametru* $\theta \in (0;1)$ *metodou maximální věrohodnosti a momentovou metodou. Odhady pak vyčíslete pro náhodný výběr* $\boldsymbol{X} = (0,1,0,2,3,1,1,0,3,0)$.

řešení

$$p(x;\theta) = \begin{cases} \theta(1-\theta)^x, & x = 0,1,2,\ldots, \\ 0, & \text{jinak.} \end{cases}$$

$\ell(\theta) = n \ln \theta + \ln(1-\theta) \sum_{i=1}^n x_i$,
věrohodnostní rovnice: $\theta \sum_{i=1}^n x_i = (1-\theta)n$,
$\widehat{\theta}_{\mathrm{ML}} = \frac{1}{1+\overline{\mathrm{X}}} = \frac{10}{21}$,
Jakým způsobem spočítáte střední hodnotu? $\mathrm{E}X = \frac{1-\theta}{\theta}$
$\widehat{\theta}_{\mathrm{M}} = \frac{1}{1+\overline{\mathrm{X}}}$

Příklad 7

Předchozí úlohu vyřešte pro náhodný výběr z geometrického rozdělení, kde X značí počet všech pokusů až do prvního úspěchu.
Odhady pak vyčíslete pro náhodný výběr $\boldsymbol{X} = (1, 2, 1, 3, 4, 2, 2, 1, 4, 1)$.

řešení

$$p(x;\theta) = \begin{cases} \theta(1-\theta)^{x-1}, & x = 0, 1, 2, \dots, \\ 0, & \text{jinak.} \end{cases}$$

$\widehat{\theta}_{\mathrm{ML}} = \frac{1}{\overline{\mathrm{X}}} = \frac{10}{21}$,
$\mathrm{EX} = \frac{1}{\theta}$, $\qquad \widehat{\theta}_{\mathrm{M}} = \frac{1}{\overline{\mathrm{X}}}$

Příklad 8

Pro náhodný výběr $\boldsymbol{X} = (X_1, \ldots, X_n)$ z normálního (Gaussova) rozdělení $\mathrm{N}(\mu, \sigma^2)$, odvoďte odhady neznámých parametrů μ, σ^2.

řešení

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

$$L(\mu, \sigma^2) = \left(2\,\pi\,\sigma^2\right)^{-n/2} \prod_{i=1}^{n} \exp\left[-\frac{(x_i-\mu)^2}{2\,\sigma^2}\right]$$

$$\ell(\mu, \sigma^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

Hledaáme stacionární bod μ, σ^2, derivujeme dle μ a σ^2:

$$\frac{\partial \ell}{\partial \mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu), \qquad \frac{\partial \ell}{\partial(\sigma^2)} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2$$

Uvědomte si geometrický význam parametru μ a první věrohodnostní rovnice. Při odvozování druhé rovnice je výhodné se vyhnout derivování podle σ.

řešení (pokračování)

Dostáváme věrohodnostní rovnice

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 = 0, \qquad n\,\sigma^2 = \sum_{i=1}^{n}(x_i-\mu)^2.$$

Z první rovnice určíme ML-odhad μ, do druhé jej pak dosadíme a určíme ML-odhad σ^2:

$$\widehat{\mu}_{\mathrm{ML}} = \overline{\mathrm{X}}, \qquad \widehat{\sigma^2}_{\mathrm{ML}} = \frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{\mathrm{X}})^2$$

Tyto výsledky už pro nás přitom nejsou novinkou. Jak se ML-odhad σ^2 liší od výběrového rozptylu?

řešení (pokračování)

Pro kontrolu maxima potřebujeme spočítat Hessovu matici

$$H = \begin{pmatrix} \frac{\partial^2 \ell}{\partial \mu^2} & \frac{\partial^2 \ell}{\partial \mu (\sigma^2)} \\ \frac{\partial^2 \ell}{\partial (\sigma^2) \mu} & \frac{\partial^2 \ell}{\partial (\sigma^2)^2} \end{pmatrix} = \begin{pmatrix} -\frac{n}{\sigma^2} & 0 \\ 0 & -\frac{n}{2\sigma^4} \end{pmatrix}$$

a ověřit její negativní definitnost (ve stacionárním bodě)

$$|H_1| = -\frac{n}{\sigma^2} < 0, \qquad |H_2| = -\frac{n^2}{2\sigma^6} > 0$$

Odhady momentovou metodou:

$$\widehat{\mu}_{\mathrm{M}} = \overline{X}, \qquad \widehat{\sigma^2}_{\mathrm{M}} = m_2 - (\overline{X})^2$$

Příklad 9

Pro náhodný výběr $\boldsymbol{X} = (X_1, \ldots, X_n)$ *z logaritmického normálního rozdělení pravděpodobnosti* $\mathrm{LN}(\mu, \sigma^2)$ *s hustotou*

$$f(x) = \frac{1}{\sqrt{2\,\pi\,\sigma^2}\,x} \exp\left[-\frac{(\ln x - \mu)^2}{2\,\sigma^2}\right] \quad \textit{pro} \;\; x > 0; \qquad f(x) = 0 \;\; \textit{jinak.}$$

odvoďte odhady neznámých parametrů μ, σ^2.

řešení

$$\ell(\mu, \sigma^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 - \sum_{i=1}^{n}\ln x_i - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(\ln x_i - \mu)^2$$

$$\widehat{\mu}_{\mathrm{ML}} = \overline{\ln \boldsymbol{X}}, \qquad \widehat{\sigma^2}_{\mathrm{ML}} = \frac{1}{n}\sum_{i=1}^{n}\left(\ln x_i - \overline{\ln \boldsymbol{X}}\right)^2$$

$$\mathrm{E}X = \exp\left(\mu + \frac{1}{2}\sigma^2\right), \quad \mathrm{E}X^2 = \exp\left(2\mu + 2\sigma^2\right)$$

$$\widehat{\mu}_{\mathrm{M}} = 2\ln\overline{X} - \frac{1}{2}m2, \qquad \widehat{\sigma^2}_{\mathrm{M}} = \ln m_2 - 2\ln\overline{X}$$

Příklad 10

Pro náhodný výběr $\boldsymbol{X} = (X_1, \ldots, X_n)$ *z exponenciálního rozdělení pravděpodobnosti* $\mathrm{Ex}(\mu)$ *s hustotou*

$$f(x) = \frac{1}{\mu} \exp\left[-\frac{x}{\mu}\right] \quad \textit{pro} \quad x \geq 0; \qquad f(x) = 0 \quad \textit{jinak.}$$

odvoďte odhady neznámého parametru $\mu > 0$.

řešení

$$\ell(\mu) = -n \ln \mu - \frac{1}{\mu} \sum_{i=1}^{n} x_i, \qquad \widehat{\mu}_{\mathrm{ML}} = \overline{\mathrm{X}}$$

$$\mathrm{EX} = \mu, \qquad \mathrm{EX}^2 = 2\mu^2, \qquad \widehat{\mu}_{\mathrm{M}} = \overline{\mathrm{X}}$$

Příklad 11

Pro náhodný výběr $\boldsymbol{X} = (X_1, \dots, X_n)$ *z exponenciálního rozdělení pravděpodobnosti* $\mathrm{Ex}(\lambda)$ *s hustotou*

$$f(x) = \lambda \exp\left(-\lambda x\right) \quad \textit{pro} \quad x \geq 0; \qquad f(x) = 0 \quad \textit{jinak.}$$

odvoďte odhady neznámého parametru $\lambda > 0$.

řešení

$$\ell(\lambda) = n \ln \lambda - \lambda \sum_{i=1}^{n} x_i, \qquad \widehat{\lambda}_{\mathrm{ML}} = \frac{1}{\overline{\overline{X}}}$$

$$\mathrm{E}X = \frac{1}{\lambda}, \qquad \mathrm{E}X^2 = \frac{2}{\lambda^2}, \qquad \widehat{\lambda}_{\mathrm{M}} = \frac{1}{\overline{\overline{X}}}$$

Příklad 12

Pro náhodný výběr $\boldsymbol{X} = (X_1, \ldots, X_n)$ *z Weibullova rozdělení pravděpodobnosti* $\mathrm{Wb}(\lambda, k)$ *s hustotou*

$$f(x) = k\,\lambda\,x^{k-1} \exp\left[-\lambda\,x^k\right] \quad \textit{pro} \quad x \geq 0; \qquad f(x) = 0 \quad \textit{jinak.}$$

odvoďte odhady neznámého parametru $\lambda > 0$ *při známé hodnotě parametru* $k > 0$.

řešení

$$\ell(\lambda) = n \ln k + n \ln \lambda + (k-1) \sum_{i=1}^{n} \ln x_i - \lambda \sum_{i=1}^{n} x_i^k, \qquad \widehat{\lambda}_{\mathrm{ML}} = \frac{1}{\overline{\boldsymbol{X}^k}}$$

$$\mathrm{E}X = \frac{1}{k}\,\lambda^{-\frac{1}{k}}\,\Gamma\left(\frac{1}{k}\right), \qquad \mathrm{E}X^2 = \lambda^{-\frac{2}{k}}\,\Gamma\left(1 + \frac{2}{k}\right)$$

Příklad 13

Pro náhodný výběr $\boldsymbol{X} = (X_1, \ldots, X_n)$ *z Rayleighova rozdělení pravděpodobnosti* $\mathrm{Ra}(s)$ *s hustotou*

$$f(x) = \frac{x}{s} \exp\left[-\frac{x^2}{2\,s}\right] \quad \textit{pro} \quad x \geq 0; \qquad f(x) = 0 \quad \textit{jinak.}$$

odvoďte odhady neznámého parametru $s > 0$.

řešení

$$\ell(\lambda) = n \ln \lambda - \lambda \sum_{i=1}^{n} x_i, \qquad \widehat{s}_{\mathsf{ML}} = \frac{1}{2\,n} \sum_{i=1}^{n} x_i^2$$

$$\mathrm{EX} = \sqrt{\frac{\pi\, s}{2}}, \qquad \mathrm{EX}^2 = 2\,s$$

$$\widehat{s}_{\mathsf{M}} = \frac{2\,m_1^2}{\pi}\ , \text{ anebo } \ \widehat{s}_{\mathsf{M}} = \frac{m_2}{2}$$

Příklad 14

Pro náhodný výběr $\boldsymbol{X} = (X_1, \ldots, X_n)$ *z Gamma rozdělení pravděpodobnosti* $\Gamma(\lambda, k)$ *s hustotou*

$$f(x) = \frac{\lambda^k}{\Gamma(k)} x^{k-1} \exp\left[-\lambda\, x\right] \quad \textit{pro} \quad x \geq 0; \qquad f(x) = 0 \quad \textit{jinak.}$$

odvoďte (rovnice pro) odhady neznámých parametrů $\lambda > 0$ *a* $k > 0$.

Pomůcka: $\frac{\partial}{\partial k} \ln \Gamma(k) = \Psi(k) =$ *digamma funkce.*

řešení

$$\ell(\lambda, k) = n\, k\, \ln \lambda - n \ln \Gamma(k) + (k-1) \sum_{i=1}^{n} \ln x_i - \lambda \sum_{i=1}^{n} x_i$$

věrohodnostní rovnice: $\widehat{\lambda}_{\mathrm{ML}} = \dfrac{\widehat{k}_{\mathrm{ML}}}{\overline{X}}, \qquad \Psi(\widehat{k}_{\mathrm{ML}}) = \ln \widehat{k}_{\mathrm{ML}} - \ln \overline{X} + \overline{X}$

$$\mathrm{E}X = \frac{k}{\lambda}, \qquad \mathrm{E}X^2 = \frac{k(k+1)}{\lambda^2}, \qquad \widehat{\lambda}_{\mathrm{M}} = \frac{m_1}{m_2 - m_1^2}, \qquad \widehat{k}_{\mathrm{M}} = \frac{m_1^2}{m_2 - m_1^2}$$