VLASTNOSTI ODHADŮ

21.10.2019

1. $X_1, X_2, \ldots, X_n$ tvoří náhodný výběr z binomického rozdělení $\mathsf{Bi}(2, p)$, kde $p \in (0,1)$. Nechť $Y_i = \mathrm{I}[X_i = 0]$.
   (a) Rozhodněte, zda je $T_n = 1 - \sqrt{\overline{Y}_n}$ nestranný a konzistentní odhad parametru p.
   (b) Rozhodněte, zda je $U_n = \sqrt{\overline{X}_n + \overline{Y}_n - 1}$ nestranný a konzistentní odhad p.
   (c) Nalezněte asymptotické rozdělení U_n.
   (d) Uvažujte náhodné vektory tvaru $(\overline{X}_n/2, U_n)^T$. Vyšetřete jejich konvergenci v pravděpodobnosti a nalezněte asymptotické rozdělení.
   (e) Porovnejte asymptotické rozptyly $\overline{X}_n/2$ a U_n a rozhodněte, který z těchto dvou konzistentních odhadů p byste zvolili.

2. Nechť $X_1, \ldots, X_n$ je náhodný výběr z Poissonova rozdělení s parametrem $\lambda > 0$.
   (a) Uvažujte odhady λ tvaru $U_n = \overline{X}_n$ a $T_n = S_n^2$. Vyšetřete jejich nestrannost a konzistenci.
   (b) Uvažujte odhad $V_n = \frac{1}{2}(X_1 + X_n)$. Vyšetřete jeho nestrannost a konzistenci.
   (c) Uvažujte odhad $W_n = \frac{1}{n-1}\sum_{i=1}^n X_i$. Vyšetřete jeho nestrannost a konzistenci.
   (d) Spočtěte střední čtvercovou chybu odhadů U_n a W_n a porovnejte.
   (e) Který z odhadů U_n, T_n, V_n, W_n byste spíše doporučili a proč?
   Využijte toho, že pro Poissonovo rozdělení je $\mathsf{E}(X_1 - \lambda)^4 = \lambda + 3\lambda^2$.
   (f) Pro $n = 1$ najděte nestranný odhad parametru $\theta_X = \mathrm{e}^{-2\lambda}$. Zamyslete se nad jeho užitečností.

3. Nechť $X_1, \ldots, X_n$ je náhodný výběr z alternativního rozdělení s parametrem $p \in (0,1)$.
   (a) Rozhodněte, zda je $T_n = \overline{X}_n(1 - \overline{X}_n)$ nestranný a konzistentní odhad rozptylu $p(1-p)$. Pokud není, spočtěte jeho vychýlení.
   (b) Uvědomte si, jaký je vztah mezi T_n a výběrovým rozptylem S_n^2 pro alternativní rozdělení.
   (c) Ukažte, že neexistuje nestranný odhad parametru $\theta_X = \frac{1}{p}$.

4. Nechť $X_1, \ldots, X_n$ je náhodný výběr z $\mathsf{N}(\mu, \sigma^2)$, $\mu \in \mathbb{R}, \sigma^2 > 0$. Uvažujte odhady $S_n^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2$ a $\sigma_n^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \overline{X}_n)^2$.
   (a) Spočtěte střední čtvercovou chybu S_n^2.
   (b) Spočtěte střední čtvercovou chybu σ_n^2 a porovnejte s (a).
   (c) Rozhodněte, zda $S_n = \sqrt{S_n^2}$ je konzistentní a nestranný odhad parametru $\theta_X = \sigma$. Pokud není nestranný, spočítejte jeho vychýlení.
   Návod: Pro výpočet vychýlení využijte toho, že znáte rozdělení $(n-1)S_n^2/\sigma^2$ *a že hustota* χ_k^2 *rozdělení je* $f_k(x) = \dfrac{1}{2^{k/2}\Gamma(k/2)} x^{k/2-1} e^{-x/2} I[x \geq 0]$.

5. Nechť $X_1, \ldots, X_n$ je náhodný výběr z rovnoměrného rozdělení $R(0, \theta)$, kde $\theta > 0$.
   (a) Rozhodněte, zda $\widetilde{\theta}_n = 2\,\overline{X}_n$ je nestranný a konzistentní odhad parametru θ.
   (b) Spočtěte střední čtvercovou chybu odhadu $\widetilde{\theta}_n$.
   (c) Rozhodněte, zda $\widehat{\theta}_n = \max_{1 \leq i \leq n} X_i$ je nestranný a konzistentní odhad parametru θ.
   (d) Spočtěte střední čtvercovou chybu odhadu $\widehat{\theta}_n$.
   (e) Kterému z odhadů $\widehat{\theta}_n$, $\widetilde{\theta}_n$ byste dali přednost?

OPAKOVÁNÍ Z PŘEDNÁŠKY

Nechť $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_n)$ je náhodný výběr z rozdělení $F_X \in \mathcal{F}$, kde $\theta_X = t(F_X)$ je nějaký parametr tohoto rozdělení. Nechť $\widehat{\theta}_n = T_n(\mathbf{X})$ je odhad θ_X. Pak $\widehat{\theta}_n$ může mít následující **vlastnosti**:

– $\widehat{\theta}_n$ je nestranný, pokud $\mathsf{E}\widehat{\theta}_n = \theta_X$ pro všechna $F_X \in \mathcal{F}$,

– $\widehat{\theta}_n$ je konzistentní, pokud $\widehat{\theta}_n \xrightarrow{P} \theta_X$ pro $n \to \infty$ pro všechna $F_X \in \mathcal{F}$.

Číselně nás může zajímat

– vychýlení $\mathsf{E}(\widehat{\theta}_n - \theta_X)$,

– střední čtvercová chyba

$$\mathrm{MSE}(\widehat{\theta}_n) = E(\widehat{\theta}_n - \theta_X)^2 = \mathsf{Var}\,\widehat{\theta}_n + \Big(\mathsf{E}(\widehat{\theta}_n - \theta_X)\Big)^2,$$

– rozptyl asymptotického rozdělení, tj. σ^2 takové, že platí $\sqrt{n}(\widehat{\theta}_n - \theta_X) \xrightarrow{D} \mathsf{N}(0, \sigma^2)$ pro $n \to \infty$.

JENSENOVA NEROVNOST. Nechť X je náhodná veličina s hodnotami v intervalu I, tj. $\mathsf{P}(X \in I) = 1$. Nechť g je konvexní funkce taková, že existuje $\mathsf{E}g(X)$. Pak

$$\mathsf{E}g(X) \geq g(\mathsf{E}X)$$

a rovnost nastává právě tehdy, když je $g(x) = ax + b$ pro $x \in I$, nebo když je X rovno konstantě skoro jistě.

DELTA METODA
– *Obecná verze*: Nechť $\{\mathbf{T}_n\}$ splňuje $\sqrt{n}(\mathbf{T}_n - \boldsymbol{\mu}) \xrightarrow{D} \mathsf{N}_k(\mathbf{0}, \boldsymbol{\Sigma})$ pro nějaké $\boldsymbol{\mu} \in \mathbb{R}^k$ a $\boldsymbol{\Sigma} \in \mathbb{R}^{k\times k}$ pozitivně semi-definitní, a nechť $\mathbf{g} = (g_1, \dots, g_m)^\mathsf{T} : \mathbb{R}^k \to \mathbb{R}^m$ má spojitý Jakobián

$$\mathbb{D}\,\mathbf{g}(\mathbf{x}) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_1}{\partial x_k} \\ \frac{\partial g_2}{\partial x_1} & \dots & \frac{\partial g_2}{\partial x_k} \\ \dots & \dots & \dots \\ \frac{\partial g_m}{\partial x_1} & \dots & \frac{\partial g_m}{\partial x_k} \end{pmatrix}$$

na nějakém okolí bodu $\mathbf{x} = \boldsymbol{\mu}$. Pak

$$\sqrt{n}\,(\mathbf{g}(\mathbf{T}_n) - \mathbf{g}(\boldsymbol{\mu})) \xrightarrow{D} \mathsf{N}_m(\mathbf{0}, (\mathbb{D}\,\mathbf{g}(\mu))\boldsymbol{\Sigma}(\mathbb{D}\,\mathbf{g}(\mu))^\mathsf{T}), \quad n \to \infty.$$

CENTRÁLNÍ LIMITNÍ VĚTA (CLV): Nechť $\{\mathbf{X}_n\}$ je náhodný výběr z rozdělení se střední hodnotou $\boldsymbol{\mu} \in \mathbb{R}^k$ a konečnou rozptylovou maticí $\boldsymbol{\Sigma} \in \mathbb{R}^{k\times k}$. Pak platí

$$\sqrt{n}(\overline{\mathbf{X}}_n - \boldsymbol{\mu}) \xrightarrow{D} \mathsf{N}_k(\mathbf{0}, \boldsymbol{\Sigma}), \quad n \to \infty.$$