VLASTNOSTI ODHADŮ

4.CVIČENÍ

1. Nechť $X_1, \dots, X_n$ je náhodný výběr z Poissonova rozdělení s parametrem $\lambda > 0$.
   (a) Uvažujte odhady λ tvaru $U_n = \overline{X}_n$ (výběrový průměr) a $T_n = S_n^2$ (výběrový rozptyl). 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?
   Návod: Pro výpočet asymptotického rozptyly T_n 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 $\theta = \mathrm{e}^{-2\lambda}$. Zamyslete se nad jeho užitečností.

2. Nechť $X_1, \dots, X_n$ 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 $\frac{1}{p}$.

3. $X_1, X_2, \dots, X_n$ tvoří náhodný výběr z binomického rozdělení $\mathsf{Bi}(2, p)$, kde $p \in (0, 1)$. Nechť $Y_i = \mathrm{I}[X_i = 0]$.
   (a) Zkonstruujte momentový odhad parametru p založený na $Y_1, \dots, Y_n$.
   (b) Rozhodněte, zda je odhad z (a) nestranný a konzistentní.
   (c) Rozhodněte, zda je $U_n = \sqrt{\overline{X}_n + \overline{Y}_n - 1}$ nestranný a konzistentní odhad p.
   (d) Porovnejte asymptotické rozptyly $\overline{X}_n/2$ a U_n a rozhodněte, který z těchto dvou odhadů p byste zvolili.

4. Nechť $X_1, \dots, X_n$ 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 $\widehat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \overline{X}_n)^2$.
   (a) Spočtěte střední čtvercovou chybu S_n^2.
   (b) Spočtěte střední čtvercovou chybu $\widehat{\sigma}_n^2$ a porovnejte s (a).
   (c) Rozhodněte, zda $S_n = \sqrt{S_n^2}$ je konzistentní a nestranný odhad parametru σ. Pokud není nestranný, spočítejte jeho vychýlení.
   Návod: Využijte toho, že znáte rozdělení $(n-1)S_n^2/\sigma^2$. V (c) pak využijte znalosti hustoty χ_k^2 rozdělení, ktrerá je $f_k(x) = \dfrac{1}{2^{k/2}\Gamma(k/2)} x^{k/2-1} e^{-x/2}\, \mathbb{I}\{x \geq 0\}$.

Opakování z přednášky

Vlastnosti odhadů. 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 $\widehat{\theta}_n \xrightarrow[n\to\infty]{\mathsf{P}} \theta_X$ pro všechna $F_X \in \mathcal{F}$.

Číselné charakteristiky odhadů. Při porovnávání různých odhadů θ_X nás číselně mohou zajímat následující charakteristiky:

- vychýlení $\mathsf{E}\,(\widehat{\theta}_n - \theta_X)$,
- střední čtvercová chyba

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

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

Jensenova nerovnost. Nechť X je náhodná veličina s hodnotami v intervalu $I \subset \mathbb{R}$, tj. platí $\mathsf{P}(X \in I) = 1$. Nechť g je konvexní funkce na I 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.

Momentová metoda Nechť $X = (X_1, \dots, X_n)$ je náhodný výběr z rozdělení $F_X \in \mathcal{F}$.

- Nechť $\theta_X = t(F_X) \in \mathbb{R}$ je nějaký parametr tohoto rozdělení a nechť $\mathsf{E}\,X_1 = g(\theta_X)$ pro nějakou ryze monotónní funkci $g\colon \mathbb{R} \to \mathbb{R}$. Momentový odhad $\widehat{\theta}_n$ se získá jako řešení $\overline{X}_n = g(\widehat{\theta}_n)$, neboli

$$\widehat{\theta}_n = g^{-1}\left(\overline{X}_n\right).$$

- V případě, že $\mathsf{E}\,X_1$ nezávisí na θ_X, nebo pokud $\theta_X \in \mathbb{R}^d$, $d > 1$, pak do odhadovacího procesu analogickým způsobem zapojíme momenty vyšších řádů, tj. $\mathsf{E}\,X^k$ a $\frac{1}{n}\sum_{i=1}^n X_i^k$ pro $k > 1$.