Momentový odhad a vlastnosti odhadu

4.12.2018

1. Lze předpokládat, že počet gólů vstřelených v jednom fotbalovém zápase v jedné konkrétní soutěži se řídí Poissonovým rozdělením, tj.

$$\mathsf{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \dots,$$

kde $\lambda > 0$ je neznámý parametr. Dále lze uvažovat, že počty gólů vstřelené v různých zápasech jsou vzájemně nezávislé a stejně rozdělené náhodné veličiny. Budeme zaznamenávat počty gólů v n zápasech a naměříme tak veličiny $X_1, \dots, X_n$.

(a) Nalezněte odhad $\widehat{\lambda}_n$ parametru λ momentovou metodou.

(b) Rozhodněte, zda je odhad $\widehat{\lambda}_n$ nestranný a konzistentní.

(c) Uvažujte odhady $\check{\lambda}_n = (X_1 + X_n)/2$ a $\widetilde{\lambda}_n = \frac{1}{n-1} \sum_{i=1}^n X_i$. Zjistěte, zda jsou tyto odhady nestranné a konzistentní odhady λ. Který z nich je rozumnější?

Dále nás bude zajímat odhad pravděpodobnosti, že v daném zápase nepadne ani jeden gól, tj. odhad $p_0 = \mathrm{P}(X_1 = 0) = e^{-\lambda}$. Uvažujme následující dva odhady p_0:

$$W_n = \frac{1}{n} \sum_{i=1}^n \mathrm{I}[X_i = 0], \quad V_n = e^{-\overline{X}_n}.$$

(d) Vyšetřete vlastnosti odhadů W_n a V_n

(e) Obdrželi jsme následující hodnoty: 0, 3, 1, 4, 1, 3, 2, 2, 2, 3, 3, 1, 3, 2, 3, 5, 3, 4, 5, 3. Spočtěte pro ně odhad λ i odhady p_0.

(f) Navrhněte odhad pravděpodobnosti, že v zápase padne přesně 6 gólů. Spočtěte pro výše uvedená data.

2. Uvažujte náhodný výběr $X_1, \dots, X_n$ z rozdělení s hustotou

$$f(x) = \begin{cases} \dfrac{p}{x^{p+1}} & x \geq 1, \\ 0 & \text{jinak,} \end{cases}$$

kde $p > 2$ je neznámý parametr.

(a) Odhadněte parametr p momentovou metodou.

(b) Vyšetřete vlastnosti odhadu z (a).

3. Náhodná veličina X má normální rozdělení $\mathsf{N}(0, 1)$. Nechť $\mu \in \mathbb{R}$ a $\sigma^2 > 0$ a $Y = \sigma X + \mu$.

(a) Vyjádřete distribuční funkci Y pomocí distribuční funkce Φ.

(b) Ukažte, že Y má $\mathsf{N}(\mu, \sigma^2)$ rozdělení. Diskutujte, jak se hustota tohoto rozdělení mění, budeme-li měnit μ a σ^2.

(c) Určete střední hodnotu X a následně pomocí (b) střední hodnotu Y.

(d) Dokažte, že platí $\Phi(x) + \Phi(-x) = 1$

(e) Určete pravděpodobnost $\mathsf{P}(|Y - \mu| \leq 2\sigma)$.
Můžete využít, že $\Phi(2) = 0.977$.

Opakování

Vlastnosti odhadů:

- Řekneme, že T_n je **nestranný** odhad θ, jestliže $\mathsf{E}T_n = \mathsf{E}_\theta T_n = \theta$ pro všechna $\theta \in \Theta$.
- Řekneme, že T_n je **konzistentní** odhad θ, jestliže $T_n \xrightarrow{\mathsf{P}} \theta$ pro $n \to \infty$ pro všechna $\theta \in \Theta$, tj. platí, že pro všechna $\varepsilon > 0$

$$\mathsf{P}_\theta(|T_n - \theta| > \varepsilon) \to 0 \quad \text{pro } n \to \infty \quad \text{pro všechna } \theta \in \Theta.$$

Jensenova nerovnost. Nechť Y je náhodná veličina a nechť g je konvexní funkce taková, že existuje $\mathsf{E}g(Y)$. Pak platí

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

a rovnost nastává jen v případě, kdy je Y konstanta (skoro jistě) nebo g je lineární funkce. Je-li g konkávní, dostáváme opačnou nerovnost.

Markovova nerovnost. Nechť $Y \geq 0$ je náhodná veličina, $\varepsilon > 0$ je libovolné a $k \in \mathbb{N}$. Pak

$$\mathsf{P}(Y > \varepsilon) \leq \frac{\mathsf{E}Y^k}{\varepsilon^k}.$$

(Slabý) zákon velkých čísel. Nechť $X_1, X_2, \ldots,$ jsou nezávislé a stejně rozdělené náhodné veličiny s konečnou střední hodnotou $\mathsf{E}X_1 = \mu$ a konečným rozptylem. Pak

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{\mathsf{P}} \mu \quad \text{pro} \quad n \to \infty.$$

Věta o spojité transformaci. Jestliže pro posloupnost náhodných veličin $\{Y_n\}$ a $a \in \mathbb{R}$ platí $Y_n \xrightarrow{\mathsf{P}} a$ a g je spojitá funkce, pak $g(Y_n) \xrightarrow{\mathsf{P}} g(a)$.

Normální rozdělení. Normální rozdělení $\mathsf{N}(\mu, \sigma^2)$ má hustotu

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}, \quad x \in \mathbb{R},$$

kde $\mu \in \mathbb{R}$ a $\sigma^2 > 0$ jsou parametry.

- Je-li $\mu = 0$ a $\sigma^2 = 1$, tj. $f(x) = \frac{1}{\sqrt{2\pi}} \exp\{-x^2/2\}$, pak se toto rozdělení nazývá **normované** normální rozdělení a značí se $\mathsf{N}(0,1)$.
- **Distribuční funkce** rozdělení $\mathsf{N}(0,1)$ se značí jako Φ, tj. $\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\{-t^2/2\} dt$. Tento určitý integrál je možné spočítat jen **numericky**, a proto hodnoty funkce Φ nalezneme **v tabulkách** (nebo získáme pomocí vhodného softwaru).