Momentový odhad a vlastnosti odhadu

28.11.2018

1. Lze předpokládat, že počet gólů vstřelených v jednom fotbalovém zápase v jedné konkrétní soutěži se řídí Poissonovým rozdělením s neznámým parametrem $\lambda > 0$ a že počty gólů vstřelené v různých zápasech jsou vzájemně nezávislé náhodné veličiny. Budeme zaznamenávat počty gólů v n zápasech a naměříme tak veličiny $X_1, \ldots, X_n$.

   (a) Nalezněte odhad parametru λ momentovou metodou a zjistěte, zda je tento odhad nestranný a konzistentní.

   (b) 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. Vyšetřete vlastnosti takového odhadu.

2. Na cvičení je nahlášeno 23 studentů, ale počet studentů, kteří opravdu dorazí je náhodná veličina. Lze předpokládat, že návštěvnosti v jednotlivých týdnech jsou na sobě nezávislé náhodné veličiny s binomickým rozdělením $\mathsf{Bi}(23, p)$ s neznámým parametrem $p \in (0, 1)$.

   (a) Navrhněte odhad $\widehat{p}_n$ parametru p momentovou metodou pro realizace $X_1, \ldots, X_n$. Rozhodněte, zda je tento odhad nestranný a konzistentní.

   (b) Spočítejte odhad z (a) pro skutečné návštěvnosti, které na 8 cvičeních byly:
   23, 23, 21, 22, 20, 21, 21, 21.

   (c) Navrhněte dva různé odhady pravděpodobnosti, že na cvičení dorazí více než 20 studentů. Diskutujte jejich vlastnosti a porovnejte jejich hodnoty pro naše data.

   (d) Odhadněte pravděpodobnost, že na cvičení nedorazí ani jeden student.

3. Uvažujte náhodný výběr $X_1, \ldots, X_n$ 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).

   (c) Jakým způsobem bychom mohli generovat veličiny z tohoto rozdělení?

   (d) Uvažujte odhad p ve tvaru $U_n = \dfrac{n}{\sum_{i=1}^n \log X_i}$. Vyšetřete jeho vlastnosti.

   (e) Navrhněte tři různé konzistentní odhady $P(X_i > 2)$.

Opakování

Bodový odhad:

- Nechť $X_1, \ldots, X_n$ je náhodný výběr z rozdělení, které závisí na neznámém parametru $\theta \in \Theta$ (např. $\mathsf{Alt}(\theta)$, $\mathsf{Po}(\theta)$, $\mathsf{Exp}(\theta)$). Odhadem θ je libovolná (měřitelná) funkce náhodného výběru $X_1, \ldots, X_n$, jejíž funkční předpis nezávisí na θ.
- **Konstrukce odhadu momentovou metodou:** Předpokládejme, že $\mathsf{E}X_i = g(\theta)$ pro nějakou známou spojitou funkci g. Pak momentový odhad $\widehat{\theta}_n$ parametru θ spočítáme tak, že

$$\overline{X}_n = g(\widehat{\theta}_n),$$

  kde $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ je výběrový průměr.
  Odhadujeme-li více parametrů, nebo pokud $\mathsf{E}X_i$ na θ nezávisí, zapojíme do výběrové momenty vyšších řádů, tj. $\frac{1}{n}\sum_{i=1}^n X_i^k$.

Vlastnosti odhadů:

- Řekneme, že T_n je **nestranný** odhad θ, jestliže

$$\mathsf{E}T_n = \mathsf{E}_\theta T_n = \theta, \qquad \text{pro všechna } \theta \in \Theta.$$

- 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 $\varepsilon > 0$

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

(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 $g(Y_n) \xrightarrow{\mathsf{P}} g(a)$.

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.