Bodové odhady

5.12.2017

1. Provádíme průzkum, jaký skutečný objem piva točí v nejmenované hospodě. Zakoupeno bylo $n = 10$ piv a jejich objem byl (v litrech):

$$0.510,\ 0.462,\ 0.491,\ 0.466,\ 0.451,\ 0.503,\ 0.475,\ 0.487, 0.512, 0.505.$$

Předpokládejte, že natočený objem piva se řídí normálním rozdělením a jednotlivá měření jsou nezávislá.

   (a) Odhadněte střední hodnotu objemu natočeného piva. Napište nejprve obecný vzorec a pak spočtěte konkrétní hodnotu pro naše data.
   (b) Je výše uvedený odhad nestranný a konzistetní?
   (c) Zkuste najít odhad střední hodnoty, který by byl konzistentní ale nebyl by nestranný. Podobně najděte odhad, který je nestranný ale není konzistentní.
   (d) Jak bychom odhadli rozptyl objemu piva?
   (e) Odhadněte pomocí empirické distribuční funkce pravděpodobnost, že dostaneme „podmírák“.
   (f) Mohli bychom pro konstrukci odhadu pravděpodobnosti z předchozího bodu využít i odhady z (a) a (d)?

2. Na cvičení je nahlášeno 27 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 $\mathsf{Bi}(27, p)$ s neznámým parametrem $p \in (0, 1)$.

   (a) Navrhněte odhad $\widehat{p}$ parametru p momentovou metodou pro realizace $X_1, \ldots, X_n$.
   (b) Spočítejte odhad z (a) pro skutečné návštěvnosti, které byly:
   25, 26, 26, 23, 25, 27, 26, 26,23.
   (c) Je (obecný) odhad $\widehat{p}$ z (a) nestranný?
   (d) Je (obecný) odhad $\widehat{p}$ z (a) konzistentní?
   (e) Pomocí odhadu z (a) odhadněte, s jakou pravděpodobností na cvičení nedorazí ani jeden student. Porovnejte s odhadem založeným na empirické distribuční funkci.
   (f) Jaké jsou vlastnosti odhadů z (e)?

3. Lze předpokládat, že délka telefonního hovoru je náhodná veličina s exponenciálním rozdělením s hustotou $f(x) = \lambda e^{-\lambda x} I[x \geq 0]$ pro nějaký neznámý parametr $\lambda > 0$ a že délky jednotlivých hovorů jsou na sobě nezávislé náhodné veličiny $X_1, \ldots, X_n$.

   (a) Zkonstruujte momentový odhad parametru λ.
   (b) Je tento odhad konzistentní?
   (c) Navrhněte dva různé způsoby, jak z dat $X_1, \ldots, X_n$ odhadnout pravděpodobnost, že hovor bude delší než 5 minut.

Opakování

Teorie odhadu:

- Nechť $X_1, \dots, X_n$ je náhodný výběr z rozdělení, které závisí na **neznámém** parametru $\theta \in \Theta$ (např. $\mathsf{Alt}(\theta)$, $\mathsf{Po}(\theta)$, $\mathsf{Exp}(\theta)$ aj.). **Odhadem** θ je libovolná (měřitelná) funkce náhodného výběru $X_1, \dots, X_n$, jejíž funkční předpis nezávisí na θ.
  Odhad $T_n = T_n(X_1, \dots, X_n)$ je tedy **náhodná veličina**.

Vlastnosti odhadů.

- V praxi pracujeme pouze s odhady, které mají „pěkné" vlastnosti.
- Řekneme, že odhad $T_n = T_n(X_1, \dots, X_n)$ je **nestranný** odhad θ, jestliže

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

- Řekneme, že odhad $T_n = T_n(X_1, \dots, X_n)$ je **konzistentní** odhad θ, jestliže $T_n \to \theta$ v pravděpodobnosti pro $n \to \infty$ pro všechna $\theta \in \Theta$.

Konstrukce bodových odhadů 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. $\frac{1}{n}\sum_{i=1}^n X_i^k$.

(Slabý) zákon velkých čísel. Nechť $X_1, \dots,$ jsou nezávislé a stejně rozdělené náhodné veličiny s konečnou střední hodnotou $\mathsf{E}X_1 = \mu$ a konečným rozptylem. Pak $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i \to \mu$ v pravděpodobnosti pro $n \to \infty$.