Vlastnosti bodových odhadů

Model: $X_1,\dots,X_n$ je náhodný výběr (tj. X_i jsou vzájemně nezávislé se stejným rozdělením) z rozdělení s distribuční funkcí $F(x,\theta)$, kde $\theta\in\Theta$ je neznámý parametr.

Cíl: Na základě náh. výběru $X_1,\dots,X_n$ získat odhad parametru θ.

Bodový odhad (parametru θ) $T=T_n=T(X_1,\dots,X_n)$ je libovolná (měřitelná) funkce náh. výběru $X_1,\dots,X_n$. Tedy je to opět náh. veličina.

Vlastnosti:

Mějme T odhad parametru θ.

T je nestranný odhad θ, jestliže $ET=\theta$, $\forall\theta\in\Theta$.

T je asymptoticky nestranný odhad θ, jestliže $\lim\limits_{n\to\infty} ET=\theta$, $\forall\theta\in\Theta$.

T je konzistentní odhad θ, jestliže $T\xrightarrow[n\to\infty]{P}\theta$, $\forall\theta\in\Theta$, tj. $\forall\varepsilon>0$ $P(|T-\theta|>\varepsilon)\xrightarrow{n\to\infty}0$.

Věta

Nechť T je asymptoticky nestranný odhad parametru θ a platí $\lim\limits_{n\to\infty} D(T)=0$ ($\forall\theta\in\Theta$), pak T je konzistentní odhad θ.

1.) $X_1,\dots,X_n$ je náhodný výběr z $Rs(0,\theta)$, kde $\theta>0$ je neznámý parametr. Uvažujme jeho 2 odhady:

$$T_1=\max\{X_1,\dots,X_n\}$$

$$T_2=2\cdot\overline{X}=2\cdot\frac{1}{n}\sum_{i=1}^{n}X_i$$

a) Jsou T_1 a T_2 nestranné odhady θ?

Začneme s T_2

$$ET_2=E\left(2\cdot\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{2}{n}\sum_{i=1}^{n}\overset{\frac{\theta}{2}}{EX_i}=\frac{2}{n}\sum_{i=1}^{n}\frac{\theta}{2}=\frac{2}{n}\cdot n\cdot\frac{\theta}{2}=\underline{\theta},\ \forall\theta>0.$$

$\Rightarrow$ T_2 je nestranný odhad θ.

$ET_1 = \underbrace{E\max\{X_1,\dots,X_n\}}_{\text{potřebujeme znát rozdělení}}$

označme přísl. distrib. fci $F_{T_1}(t) = P(\max\{X_1,\dots,X_n\} \le t) \overset{\text{NEZ.}}{=} P(X_1 \le t, \dots, X_n \le t) = \underbrace{P(X_1 \le t)}_{F(t)} \cdot \ldots \cdot \underbrace{P(X_n \le t)}_{F(t)}$

$= [F(t)]^n$, kde F je distrib. fce X_i, tj. $Rs(0,\theta) \Rightarrow F(t) = \frac{t}{\theta}$, pro $0 \le t \le \theta$.

$$= \frac{t^n}{\theta^n}, \quad 0 \le t \le \theta$$

Odtud $f_{T_1}(t) = F'_{T_1}(t) = \frac{n}{\theta^n} \cdot t^{n-1}$, $0 \le t \le \theta$

$= 0$, jinak

Počítejme

$$ET_1 = E\max\{X_1,\dots,X_n\} = \int t \cdot f_{T_1}(t)\,dt = \int_0^\theta t \frac{n}{\theta^n} \cdot t^{n-1}\,dt = \frac{n}{\theta^n} \cdot \int_0^\theta t^n\,dt =$$

$$\frac{n}{\theta^n} \cdot \left[\frac{t^{n+1}}{n+1}\right]_0^\theta = \frac{n}{\theta^n} \cdot \frac{\theta^{n+1}}{n+1} = \frac{n}{n+1} \cdot \theta \neq \theta.$$ T_1 není nestranný odhad θ.

Můžeme T_1 „opravit", aby nestranný byl?

$\tilde{T}_1 = \frac{n+1}{n} \cdot T_1 = \frac{n+1}{n} \cdot \max\{X_1,\dots,X_n\}$. Pak zřejmě $E\tilde{T}_1 = \theta$, $\forall \theta > 0$. $\tilde{T}_1$ je nestranný odhad θ.

b) Jsou T_1, $\tilde{T}_1$ a T_2 asymptoticky nestranné odhady θ?

Je-li T nestranný odhad θ, pak T je zřejmě i asymptoticky nestranný odhad $\theta \Rightarrow \tilde{T}_1$ a T_2 jsou as. nestranné

$$\lim_{n\to\infty} ET_1 = \lim_{n\to\infty} \overbrace{\frac{n}{n+1}}^{\to 1} \cdot \theta = \theta, \ \forall \theta.$$ T_1 je asymptoticky nestranný.

c) Jsou $T_1, \tilde{T}_1$ a T_2 konzistentní odhady parametru θ?

- $T_1, \tilde{T}_1$ a T_2 jsou as. nestranné odhady θ
- zbývá ukázat, že $D(T_1), D(\tilde{T}_1)$ a $D(T_2)$ konvergují pro $n \to \infty$ k 0.

Sami odvodíte, že $D(T_1) = \theta^2 \cdot \frac{n}{(n+2)(n+1)^2} \xrightarrow{n\to\infty} 0, \forall \theta > 0.$ T_1 je konzistentní odhad θ

$$D(\tilde{T}_1) = D\left(\frac{n+1}{n} \cdot T_1\right) = \frac{(n+1)^2}{n^2} \cdot D(T_1) = \frac{\theta^2}{n(n+2)} \xrightarrow{n\to\infty} 0, \forall \theta > 0$$

$\tilde{T}_1$ je konzistentní odhad θ

$$D(T_2) = D\left(\frac{2}{n} \sum_{i=1}^{n} X_i\right) = \frac{4}{n^2} D\left(\sum_{i=1}^{n} X_i\right) \overset{\text{NEZ}}{=} \frac{4}{n^2} \cdot \sum_{i=1}^{n} \overset{\frac{\theta^2}{12}}{D(X_i)} = \frac{4}{n^2} \cdot n \cdot \frac{\theta^2}{12} = \theta^2 \cdot \frac{1}{3n} \xrightarrow{n\to\infty} 0. \forall \theta > 0$$

T_2 je konzistentní odhad θ

d) Který z odhadů $T_1, \tilde{T}_1$ a T_2 je „nejlepší“?

Střední čtvercová (kvadratická) chyba odhadu T je $E(T-\theta)^2$

$$E(T-\theta)^2 = \ldots = DT + (\underbrace{ET-\theta}_{\text{bias } T = \text{vychýlení } T})^2$$

Je-li T nestranný odhad θ, pak $E(T-\theta)^2 = DT$.

- $E(T_1-\theta)^2 = DT_1 + (ET_1-\theta)^2 = \theta^2 \cdot \frac{n}{(n+2)(n+1)^2} + \left(\theta \cdot \frac{n}{n+1} - \theta\right)^2 = \ldots = \theta^2 \cdot \frac{2}{(n+2)(n+1)}$

- $E(\tilde{T}_1-\theta)^2 = D\tilde{T}_1 = \theta^2 \cdot \frac{1}{n(n+2)}$

- $E(T_2-\theta)^2 = DT_2 = \theta^2 \cdot \frac{1}{3n}$

Zřejmě platí:

$$E(\tilde{T}_1-\theta)^2 \leq E(T_1-\theta)^2 < E(T_2-\theta)^2 \quad , \forall \theta > 0$$

tedy $\tilde{T}_1$ je „nejlepší".

2.) Nechť T_1 a T_2 jsou dva nezávislé odhady parametru θ. Nechť rozptyl odhadu T_1 je dvojnásobkem rozptylu T_2. Nechť T_1 a T_2 jsou navíc nestranné odhady θ.

Definujme odhad $T = k_1 T_1 + k_2 T_2$ pro nějaké $k_1, k_2 \in \mathbb{R}$.

Určete konstanty k_1, k_2 tak, aby T byl nestranný odhad θ a měl nejmenší možný rozptyl.