Nestrannost a konzistence odhadů

$\mathbb{X} = (X_1, \dots, X_n)$ náh. výběr z $F(x, \Theta)$, $\Theta \in \tilde{\Theta} \subseteq \mathbb{R}^m$, konstrukce $T(\mathbb{X})$... odhad $g(\Theta)$

$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$, $S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$

$\mathbb{X} = (X_1, \dots, X_n)$ náh. výběr z rozdělení se stř. hodnotou μ.

V.1.6. $\bar{X}$ je nestranný odhad μ.

D.: $E\bar{X} = E\left(\frac{1}{n} \sum_{i=1}^{n} X_i\right) = \frac{1}{n} \sum_{i=1}^{n} \underbrace{EX_i}_{\mu} = \frac{1}{\cancel{n}} \cdot \cancel{n}\mu = \mu$.

V.1.7. $\mathbb{X} = (X_1, \dots, X_n)$ je náh. výběr z rozdělení s rozptylem σ^2.

Pak S^2 je nestranný odhad σ^2.

D.: Chceme $ES^2 = \sigma^2$.

$$ES^2 = E\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2 = \frac{1}{n-1}\sum_{i=1}^{n}E\left(\underbrace{X_i-\mu}+\underbrace{\mu-\bar{X}}\right)^2$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}E\left[\underline{(X_i-\mu)^2} + 2(X_i-\mu)(\mu-\bar{X}) + \underline{(\bar{X}-\mu)^2}\right]$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}(\underbrace{DX_i}_{\sigma^2} + D\bar{X}) + \frac{2}{n-1}E\sum_{i=1}^{n}(X_i-\mu)(\underbrace{\mu}_{\frac{1}{n}\sum\mu} - \frac{1}{n}\sum_{j=1}^{n}X_j)$$

$$= \frac{1}{n-1}\left(n\sigma^2 + \frac{\sigma^2}{n}\cdot n\right) - \frac{2}{n-1}E\sum_{i=1}^{n}\sum_{j=1}^{n}(X_i-\mu)\frac{1}{n}(X_j-\mu)$$

$$\boxed{D\bar{X} = D\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n}\underbrace{DX_i}_{\sigma^2} = \frac{\sigma^2}{n}}$$

$$= \frac{1}{n-1}\left(n\sigma^2+\sigma^2\right) - \frac{2}{n-1}\cdot\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}E(X_i-\mu)(X_j-\mu)$$

$$= \frac{1}{n-1}\sigma^2(n+1) - \frac{2}{(n-1)n}\sum_{i=1}^{n}\underbrace{E(X_i-\mu)^2}_{DX_i=\sigma^2}$$

$$= \frac{1}{n-1}\sigma^2(n+1) - \frac{2}{n-1}\sigma^2 = \sigma^2\frac{n+1-2}{n-1} = \underline{\underline{\sigma^2}}$$

V.1.8. $T_n = T(\mathbb{X})$ je nestranný nebo asymptoticky nestranný odhad $g(\Theta)$ a platí

$$\lim_{n\to\infty} DT_n = 0 .$$ Pak T_n je konzistentní odhad $g(\Theta)$.

D.: Čebyšev.: $\forall \varepsilon > 0$ $\quad P\left(|T_n - ET_n| \geq \frac{\varepsilon}{2}\right) \leq \frac{4 \cdot DT_n}{\varepsilon^2}$;

As. nestr.: $\exists n_0 : \forall n > n_0 :$

$$|ET_n - g(\Theta)| < \frac{\varepsilon}{2}$$

Nechť $n > n_0$:

$$P(|T_n - g(\Theta)| \geq \varepsilon) = 1 - P(|T_n - g(\Theta)| < \varepsilon)$$

$$= 1 - P(|\underline{T_n - ET_n} + \underline{ET_n - g(\Theta)}| < \varepsilon) \leq 1 - P\left(|T_n - ET_n| < \frac{\varepsilon}{2} \wedge |ET_n - g(\Theta)| < \frac{\varepsilon}{2}\right)$$

$$= 1 - P\left(|T_n - ET_n| < \frac{\varepsilon}{2}\right) = P\left(|T_n - ET_n| \geq \frac{\varepsilon}{2}\right) \overset{\check{c}.n.}{\leq} \frac{4 DT_n}{\varepsilon^2} \xrightarrow[n\to\infty]{} 0$$

Důsl. 1.9. $\overline{X}_n$ je konzistentní odhad μ.

D.: V.1.6. $\overline{X}_n$ je nestranný odhad μ

$$D\overline{X}_n = D\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} \underbrace{DX_i}_{\sigma^2} = \frac{1}{n^2}\cdot n\sigma^2 = \frac{\sigma^2}{n} \xrightarrow[n\to\infty]{} 0$$

Důsl. 1.10. S^2 je konzistentní odhad σ^2.

D.: V.1.7. S^2 je nestranný odhad σ^2

$DS^2 = \dots$ viz skripta $\longrightarrow 0$ pro $n\to\infty$.

NNO ... T^* $\quad ET^* = \gamma(\theta)$

$DT^* \leq DT$ pro vš. T : $ET = \gamma(\theta)$

Př. $\mathbb{X} = (X_1, \dots, X_n)$ náh. výběr z rozdělení se stř. h. μ.

Uvaž. vš. lineární statistiky $T = \sum_{i=1}^{n} c_i X_i$. Jak volit $c_1, \dots, c_n$, aby T byl NNO par. μ?

Řeš.

$$ET = \mu \Rightarrow ET = E\sum_{i=1}^{n} c_i X_i = \sum_{i=1}^{n} c_i \underbrace{EX_i}_{\mu} = \mu \cdot \sum_{i=1}^{n} c_i = \mu$$

$$\Downarrow$$

$$\boxed{\sum_{i=1}^{n} c_i = 1}$$

$$DT \xrightarrow{?} \min_{c_1, \dots, c_n}$$

$$DT = D\left(\sum_{i=1}^{n} c_i X_i\right) = \sum_{i=1}^{n} c_i^2 \underbrace{DX_i}_{\sigma^2} = \sigma^2 \cdot \sum_{i=1}^{n} c_i^2 \longrightarrow \text{minimalizovat za podmínky}$$

$$L(c_1, \dots, c_n, \lambda) = \sum_{i=1}^{n} c_i^2 - \lambda\left(\sum_{i=1}^{n} c_i - 1\right)$$

$$\frac{\partial L}{\partial c_j} = 2c_j - \lambda = 0 \Rightarrow c_j = \frac{\lambda}{2} \qquad c_j = c$$

$$\frac{\partial L}{\partial \lambda} = -\sum_{i=1}^{n} c_i + 1 = 0 \Rightarrow \sum_{i=1}^{n} c_i = 1$$

$$\left.\begin{array}{l} \sum c = 1 \\ nc = 1 \end{array}\right\} \Rightarrow c = \frac{1}{n}$$

$$\Uparrow$$

$\bar{X}$ je NNO μ

Střední kvadratická chyba

Def. $\mathbb{X} = (X_1, \dots, X_n)$ je náh. výběr z rozdělení s $F(x, \theta)$; $\theta \in \Theta \subseteq \mathbb{R}$

$T_n = T(\mathbb{X})$ statistika .. odhad Θ.

Střední kvadratická chyba

$$MSE(T_n) = E(T_n - \theta)^2.$$

Pozn.

$$E(T_n - \theta)^2 = E(\underbrace{T_n - ET_n} + \underbrace{ET_n - \theta})^2 =$$

$$= E\left[(T_n - ET_n)^2 + \underbrace{2(T_n - ET_n)(ET_n - \theta)}_{0} + (ET_n - \theta)^2\right]$$

$$= \underbrace{E(T_n - ET_n)^2}_{DT_n} + \underbrace{(ET_n - \theta)^2}_{bias^2}$$