Properties of estimators: MSE and Bias trade-off

MSE of an estimator $\hat{\theta}$ for θ is

$$MSE(\hat{\theta}) = E(\hat{\theta} - \theta)^2$$

Bias of an estimator $\hat{\theta}$ for θ is

$$B(\hat{\theta}) = E\hat{\theta} - \theta$$

Proposition 1:

$$MSE(\hat{\theta}) = E\left[(\theta - \theta)^2\right] = Var(\hat{\theta}) + B^2$$

Proof

Remarks(1)

Methods for estimating parameters

Definition:

Let $x_1, \dots, x_n$ be a from a distribution $F(x;\theta)$ which can be discrete or continuous The function

$$L(\boldsymbol{x}, \theta) \quad f(x_1;\theta) \times f(x_2;\theta) \times \dots \times f$$

The method of Maximum Likelihood

Def: the value $\hat{\theta}$ from sample space A for which $L(\boldsymbol{x}, \theta)$ obtains its largest value withing A is called maximum likelihood estimate (MLE) of θ:

$$\hat{\theta}_{MLE} = \underset{\theta \in A}{\arg\max}\, L(\boldsymbol{x}, \theta)$$

Def
the logarithmic likelihood function

$$l(\boldsymbol{x}, \theta) = \ln(L(\boldsymbol{x}, \theta)) = \begin{cases} \sum_{i=1}^{n} \ln f(x_i, \theta) \\ \sum_{i=1}^{n} \ln p(x_i, \theta) \end{cases}$$

* Important for MLE we make an assumption about type of distribution

The Method of Least Squares

Let $x_1, x_2, \ldots, x_n$ be a sample from distribution with mean $E(X) = \mu(\theta)$ where $\mu(\theta)$ is a known function and θ an unknown parameter with parameter space A. Let

$$Q(\theta) = \sum_{i=1}^{n} (x_i - \mu(\theta))^2$$

be the sum of squares of the deviations of the observations from $\mu(\theta)$

Def: the value $\hat{\theta}$, for which $Q(\theta)$ obtains its minimal value within A, is called the LS estimate of θ

$$\hat{\theta}_{LS} = \underset{\theta \in A}{\arg\min}\ Q(\theta)$$

* We assume the form of mean of random variable

Лекция 03 Баженов

Метод максимального правдоподобия (Фишера)

Пусть $\vec{X} = (X_1, \dots, X_n)$ - распред-е изв. типа, задаваемая параметрами $\theta = (\theta_1, \dots, \theta_k)$

Идея метода: подбираем параметры выборки таким образом что вероятность получения выборки максимальна

Например если распр-е дискретное

$P_\theta(X_1, \dots, X_n) = P_\theta(X = X_1) \cdot P_\theta(X = X_2) \dots P(X = X_n)$

независимые случ-е величины

Функция правдоподобия

$L(\vec{X}, \theta)$ наз-ся функция

$$L(\vec{X}, \theta) = \prod_{i=1}^{n} P(X = X_i)$$ - дискретное распред-е

Логарифмической ф-ей правдоподобия наз-ся

$\ln L(\vec{X}, \theta)$

Замечание

$y = \ln x$ возрастающая ф-я

экстремумы совпадают, а искать их проще во втором случае

Оценкой максимального правдоподобия $\hat{\theta}$ называется значение θ при котором ф-я L достигает наибольшего значения при ф-х значениях $X_1, \dots, X_n$. Идея приводит к эф-м оценкам

Пример 1

$\exists\ \vec{X} = (X_1, \dots, X_n)$ из распр-я Пуассона с неизвестным параметром $\lambda > 0$

$$\underline{X} \in \Pi_\lambda : \quad P(X = x_i) = \frac{\lambda^{x_i}}{x_i!} e^{-\lambda} \qquad x_i \in \mathbb{N} \cup \{0\}$$

$$L(\vec{X}, \lambda) = \prod_{i=1}^{n} \frac{\lambda^{x_i}}{x_i!} e^{-\lambda} = \frac{\lambda^{n \cdot \bar{X}}}{\prod_{i=1}^{n} x_i!} \cdot e^{-n\lambda}$$

$$\ln L(\vec{X}, \lambda) = n\bar{X} \ln\lambda - \sum_{i=1}^{n} \ln x_i! - n\lambda$$

$$\frac{d}{d\lambda} \ln L(\vec{X}, \lambda) = \frac{n\overline{X}}{\lambda} - n = 0 \Rightarrow$$

$$\hat{\lambda} = \overline{X}$$

Покажем, что это точка максимума

$$\frac{d^2}{d\lambda^2} \ln L(\vec{X}, \lambda) = -\frac{n\overline{X}}{\lambda^2} < 0 \Rightarrow$$

$$\hat{\lambda} = \overline{X} \quad \text{— т. максимума}$$

Пример 2

Нормальное распр-е

$\exists\ \vec{X} = (X_1, \dots, X_n)$ — выборка из $N(a, \sigma^2)$

$a \in \mathbb{R} \quad \sigma > 0$

$$f_{a,\sigma^2}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}$$

$$L(\vec{X}, a, \sigma) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_i-a)^2}{2\sigma^2}} =$$

$$= \frac{1}{\sigma^n (2\pi)^{n/2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-a)^2}$$

$$\ln L(\vec{X}, a, \sigma) = -n\ln\sigma - \frac{n}{2}\ln 2\pi - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i - a)^2$$

$$\frac{d}{da}\ln L(\vec{X}, a, \sigma) = -\frac{1}{2\sigma^2}\sum_{i=1}^{n} 2(X_i - a)(-1) =$$

$$= \frac{1}{\sigma^2}(n\overline{X} - na)$$

$$\frac{d}{d\sigma}\ln L(\overline{X}, a, \sigma) = -\frac{n}{\sigma} - \frac{1}{2}(-2)\cdot\sigma^{-3}\sum^{n}(X_i - a)^2 =$$

$$= \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i - a)^2 - \frac{n}{\sigma}$$

$$\begin{cases} \dfrac{n\overline{X} - na}{\sigma^2} = 0 \\ \dfrac{1}{\sigma^3}\sum_{i=1}^{n}(x_i - a)^2 - \dfrac{n}{\sigma} = 0 \end{cases} \Leftrightarrow \begin{cases} \hat{a} = \overline{X} \\ \hat{\sigma}^2 = \dfrac{1}{n}\sum^{n}(x_i - a)^2 \end{cases}$$

Var

смещённая оценка

Пример 3

равномерного распр-е

$\exists\ \vec{X} = (X_1, \dots, X_n)$ из $U(0; \theta)$, $\theta > 0$

начало/конец интервала

Найти оценки θ методом моментов и максимального правдоподобия

а) методом моментов

$$EX = \frac{a+b}{2} = \frac{\theta}{2} \qquad \overline{X} = \frac{\theta^*}{2} \Rightarrow$$

$D\theta^* =$

$$\theta^* = 2\overline{X}$$ оценка по методу моментов

б) методом макс. правд-я

$$f_\theta = \begin{cases} 0, & X < 0 \\ \frac{1}{\theta}, & 0 \le X \le \theta \\ 0, & X > 0 \end{cases}$$

Обозн-м $X_{(n)} = \max(X_1, \dots, X_n)$ — послед. порядковая стат-ка

θ не меньше чем макс элемент выборки

$$L(\vec{X}, \theta) = \prod_{i=1}^{n} f_\theta(X_i) = \begin{cases} 0, & \theta < X_{(n)} \\ \frac{1}{\theta^n}, & \theta \geq X_{(n)} \end{cases}$$

Ясно, что максимальное значение при $X_{(n)}$

$$\hat{\theta} = X_{(n)}$$

в) сравним эти оценки

1. $\theta^* = 2\overline{X}$ несмещённые т.к.

теор. мат. ожидание

$$E\theta^* = E2\overline{X} = 2E\overline{X} = 2EX = \theta$$

$$E(\theta^* - \theta)^2 = D\theta^* = D2\overline{X} = 4D\overline{X} =$$

$$= 4\frac{DX}{n} = \frac{4}{n}\frac{\theta^2}{12} = \frac{\theta^2}{3n}$$

2. $X_{(n)} = \max(X_1, \dots, X_n)$

$$F_{X_{(n)}} = P(X_{(n)} \leq X) = P(X_1 \leq X) \cdot P(X_2 \leq X) \cdot \ldots$$

$$\cdot P(X_n \leq X) = (F(X))^n$$

при $\overline{X} \in U(0,\theta)$,

$$F(x) = \begin{cases} 0, & x < 0 \\ \dfrac{x}{\theta}, & 0 \le x \le \theta \\ 1, & x > \theta \end{cases}$$

Функция распр-я

$$F_{\overline{X}_{(n)}}(x) = \begin{cases} 0, & x < 0 \\ \dfrac{x^n}{\theta^n}, & 0 \le x \le \theta \\ 1, & x > \theta \end{cases}$$

$$f_{X_{(n)}} = F'_{X_{(n)}} = \begin{cases} 0 & x < 0 \\ \dfrac{n x^{n-1}}{\theta^n} & 0 \le x \le \theta \\ 0, & x > \theta \end{cases}$$

$$E X_{(n)} = \int_0^\theta x \frac{n x^{n-1}}{\theta^n} dx = \frac{n}{\theta^n} \int_0^\theta x^n dx =$$

$$= \frac{n}{\theta^n} \frac{x^{n+1}}{n+1} \Big|_0^\theta = \frac{n}{n+1} \theta$$ – смещённая вниз оценка

$$\tilde{\theta} = \frac{n+1}{n}\hat{\theta} = \frac{n+1}{n} X_{(n)}$$ – несм-ая оценка

$$E\tilde{\theta}^2 = E\left[\left(\frac{n+1}{n} X_{(n)}\right)^2\right] = \frac{(n+1)^2}{n^2}\int_0^{\theta} x^2 \frac{n x^{n-1}}{\theta^n}\,dx =$$

$$= \frac{(n+1)^2}{n\,\theta^n}\,\frac{x^{n+2}}{n+2}\Big|_0^{\theta} = \frac{(n+1)^2}{n(n+2)}\cdot\theta^2$$

$$E(\tilde{\theta}-\theta)^2 = Var\,\tilde{\theta} = D\tilde{\theta} = E\tilde{\theta}^2 - (E\tilde{\theta})^2 =$$

$$= \frac{(n+1)^2\cdot\theta^2}{n(n+2)} - \theta^2 = \theta^2\left(\frac{\cancel{n^2}+\cancel{2n}+1-\cancel{n^2}-\cancel{2n}}{n(n+2)}\right) =$$

$$= \frac{\theta^2}{n(n+2)}$$

$$D\theta^* = \theta$$