无偏性、有效性
和 正态分布最短区间估计证明过程

一、无偏性

若 $X_1, X_2, \cdots, X_n$ 为总体 X 的一个样本，

$\theta \in \Theta$ 是包含在总体 X 的分布中的待估参数，

(Θ 是 θ 的取值范围)

若估计量 $\hat{\theta} = \theta(X_1, X_2, \cdots, X_n)$ 的数学期望 $E(\hat{\theta})$ 存在, 且对于任意 $\theta \in \Theta$ 有 $E(\hat{\theta}) = \theta$, 则称 $\hat{\theta}$ 是 θ 的无偏估计量.

无偏估计的实际意义: 无系统误差.

例3 设总体 X 服从参数为 θ 的指数分布，其概率密度为

$$f(x)=\begin{cases}\dfrac{1}{\theta}e^{-x/\theta}, & x>0,\\ 0, & \text{其它},\end{cases}$$

其中 $\theta>0$ 为未知，$X_1,X_2,\ldots X_n$ 是取自总体的一个样本，

试证 $\bar{X}$ 和 $nZ=n\min(X_1,\ldots,X_n)$ 都是参数 θ 的无偏估计量。

$$L(\theta)=\prod_{i=1}^{n} f(x_i,\theta)=\prod_{i=1}^{n}\frac{1}{\theta}e^{-\frac{x_i}{\theta}}=\frac{1}{\theta^n}e^{-\frac{\sum_{i=1}^{n}x_i}{\theta}}$$

$$\ln L(\theta)=-n\ln\theta-\frac{\sum_{i=1}^{n}x_i}{\theta}$$

$$\frac{\partial \ln L(\theta)}{\partial\theta}=-\frac{n}{\theta}+\frac{\sum_{i=1}^{n}x_i}{\theta^2}=0$$

$\Rightarrow$ 极大似然估计为 $\widehat{\theta}_{\mathrm{MLE}}=\frac{1}{n}\sum_{i=1}^{n}x_i=\bar{X}$

无偏性：

$$E\widehat{\theta}_{\mathrm{MLE}} = E\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \frac{1}{n}E\left(X_1 + \ldots + X_n\right) = \frac{1}{n}\left(\theta + \ldots + \theta\right) = \theta$$

方差为：

$$D\left(\widehat{\theta}_{\mathrm{MLE}}\right) = D\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \frac{1}{n^2}D\left(X_1 + \ldots + X_n\right)$$

$$= \frac{1}{n^2}\left[D\left(X_1\right) + \ldots + D\left(X_n\right)\right]$$

$$= \frac{1}{n^2}\left[\theta^2 + \ldots + \theta^2\right] = \frac{1}{n^2}n\theta^2 = \frac{\theta}{n}$$

最小值$Z=\min\{X_1,\ldots,X_n\}$的分布函数F(z)

$$\mathrm{F}(\mathrm{z})=P(Z\le z)=P(\min\{X_1,\ldots,X_n\}\le z)$$

当 $z<0$时， $\mathrm{F}(\mathrm{z})=0$

$z\ge 0$时，

$$\mathrm{F}(\mathrm{z})=1-P(\min\{X_1,\ldots,X_n\}>z)$$

$$=1-P(X_1>z,\ldots,X_n>z)$$

$$=1-\prod_{i=1}^{n}P(X_i>z)=1-\prod_{i=1}^{n}\left(\int_z^{\infty}\frac{1}{\theta}e^{-\frac{x}{\theta}}dx\right)$$

$$=1-\prod_{i=1}^{n}e^{-\frac{z}{\theta}}=1-e^{-\frac{nz}{\theta}}=1-\exp\left(-\frac{z}{\theta/n}\right)$$

最小值Z服从指数分布，参数为$\dfrac{1}{\theta/n}$

$$Z=\min\{X_1,\ldots,X_n\}$$

无偏性：

$$E(nZ)=nE(Z)=n\frac{\theta}{n}=\theta$$

方差为：

$$D(nZ)=n^2D(Z)=n^2\left(\frac{\theta}{n}\right)^2=\theta^2$$

$\widehat{\theta}_{\mathrm{MLE}}=\overline{X}$ 和 nZ的有效性：

$$E(\overline{X})=E(nZ)=\theta$$

当$n=1$时候，

$$X_1=\min\{X_1\}=\frac{1}{1}\sum_{i=1}^{1}X_i=\overline{X},D(\overline{X})=D(nZ)=\theta^2$$

当$n>1$时， $D(\overline{X})=\dfrac{\theta^2}{n}<D(nZ)=\theta^2,\overline{X}$更有效

极大似然估计的无偏性讨论

设总体 X 在 $[a,b]$ 上服从均匀分布，其中 a, b 未知，$x_1, x_2, \cdots, x_n$ 是来自总体 X 的一个样本值，求 a,b 的最大似然估计量.

解　记　$x_{(1)} = \min(x_1, x_2, \cdots, x_n)$,

$x_{(n)} = \max(x_1, x_2, \cdots, x_n)$,

X 的概率密度为

$$f(x; a, b) = \begin{cases} \dfrac{1}{b-a}, & a \le x \le b \\ 0, & \text{其它} \end{cases}$$

$\boldsymbol{a, b}$ **的最大似然估计值**

$$\hat{a} = x_{(1)} = \min_{1\le i\le n} x_i, \quad \hat{b} = x_{(n)} = \max_{1\le i\le n} x_i,$$

下面证明 b的极大似然估计不是无偏估计。

最大值$\mathrm{Y} = \max\{X_1, \ldots, X_n\}$的分布函数$\mathrm{F}(\mathrm{y})$

$$\mathrm{F}(\mathrm{y}) = P(Y \le y) = P(\max\{X_1, \ldots, X_n\} \le y)$$

当 $\mathrm{y} < a$时， $\mathrm{F}(\mathrm{z}) = 0$; 当$y > b$时， $\mathrm{F}(\mathrm{z}) = 1$

$a \le y \le b$时，

$$\mathrm{F}(\mathrm{y}) = P(\max\{X_1, \ldots, X_n\} \le y) = P(X_1 \le \mathrm{z}, \ldots, X_n \le y)$$

$$= \prod_{i=1}^{n} P(X_i \le y) = \prod_{i=1}^{n} \left(\int_a^y \frac{1}{b-a} dx \right) = \left(\frac{y-a}{b-a} \right)^n$$

$$Y = \max\{X_1, \ldots, X_n\}\text{的密度函数为}f(y) = n\frac{(y-a)^{n-1}}{(b-a)^n}$$

b的极大似然估计为 $\hat{b}=\max\{X_1,\ldots,X_n\}=Y$

$$E\left(\hat{b}\right)=E(\max\{X_1,\ldots,X_n\})=E(Y)$$

$$=\int_{-\infty}^{\infty} yf(y)dy=\int_a^b yn\frac{(y-a)^{n-1}}{(b-a)^n}dy$$

$$=\int_a^b (y-a+a)n\frac{(y-a)^{n-1}}{(b-a)^n}dy$$

$$=\int_a^b n\frac{(y-a)^n}{(b-a)^n}dy+a\int_a^b n\frac{(y-a)^{n-1}}{(b-a)^n}dy$$

$$=\frac{n}{(n+1)(b-a)^n}(y-a)^{n+1}\Big|_a^b+a$$

$$=\frac{n(b-a)}{n+1}+a=\frac{nb+a}{n+1}\neq b,$$ 所以 $\hat{b}$ 不是 b 的无偏估计

正态分布最短区间估计证明过程

单个正态总体参数的置信区间

一、σ已知时μ的置信区间

$\bar{x} \sim N(\mu, \sigma^2/n)$

在这种情况下，枢轴量可选为 $G=\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$，***c***和***d***应满足***P*(*c*<*G*<*d*)=*Φ*(*d*)−*Φ*(*c*)= 1−*α***，经过不等式变形可得

$$P_{\mu}\left(\bar{x}-d\sigma/\sqrt{n}<\mu<\bar{x}-c\sigma/\sqrt{n}\right)=1-\alpha$$

该区间长度为$(d-c)\sigma/\sqrt{n}$。当***d*=−*c*=z**$_{\alpha/2}$时，***d*−*c***达到最小，由此给出了的同等置信区间为

$$\left(\bar{x}-z_{\alpha/2}\sigma/\sqrt{n},\ \bar{x}+z_{\alpha/2}\sigma/\sqrt{n}\right)。$$

这是一个以$\bar{x}$ 为中心，半径为 $z_{\alpha/2}\sigma/\sqrt{n}$ 的对称区间，常将之表示为 $\bar{x}\pm z_{\alpha/2}\sigma/\sqrt{n}$

拉格朗日分析法

$$P\left(c<\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}<d\right)=P\left(c<N(0,1)<d\right)=\Phi(d)-\Phi(c)=1-\alpha$$

约束条件$\int_c^d \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx=1-\alpha$。$P\left(\bar{X}-\frac{d\sigma}{\sqrt{n}}<\mu<\bar{X}-\frac{c\sigma}{\sqrt{n}}\right)=1-\alpha,$

平均区间长度为 $\frac{(d-c)\sigma}{\sqrt{n}},$

构造拉格朗日算子$f(c,d)=(d-c)+\lambda\left(\int_c^d \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx-1+\alpha\right)$

$$\frac{\partial f}{\partial c}=-1-\lambda\frac{1}{\sqrt{2\pi}}e^{-\frac{c^2}{2}}=0, \qquad \frac{\partial f}{\partial d}=1+\lambda\frac{1}{\sqrt{2\pi}}e^{-\frac{d^2}{2}}=0$$

$$\frac{1}{\sqrt{2\pi}}e^{-\frac{c^2}{2}}=-\frac{1}{\lambda}=\frac{1}{\sqrt{2\pi}}e^{-\frac{d^2}{2}} \qquad \Rightarrow c^2=d^2,(d+c)(d-c)=0$$

又因为$d>c,\Rightarrow d+c=0 \qquad \Rightarrow d=-c,$

$$\Phi(d)-\Phi(c)=\Phi(d)-\Phi(-d)=2\Phi(d)-1=1-\alpha,$$

$$\Phi(d)=1-\frac{\alpha}{2}, \quad d=z_{\frac{\alpha}{2}},\ c=-z_{\frac{\alpha}{2}}$$