§7.2 估计量的优良性准则

对于总体的一个参数, 可用各种不同的方法去估计它, 因此一个参数的估计量不唯一.

如，当 $X \sim U(0,\theta)$, θ的矩法估计量为 $2\overline{X}$，而极大似然估计量为 $\max\limits_{1\le i\le n}\{X_i\}$.

在众多的估计量中，选哪一个更好？选取的标准是什么？

三个常用的准则：

估计量的优良性准则(三个常用)

1. 无偏性(**样本容量n固定**)

定义：设 $\hat{\theta}(X_1, X_2,\dots, X_n)$ 是未知参数θ 的估计量，

若$E(\hat{\theta})=\theta$,则称$\hat{\theta}$为θ的无偏估计. 等价: $E(\hat{\theta}-\theta)=0$

2. 有效性: 设 $\hat{\theta}_1(X_1, X_2,\dots, X_n)$ 和 $\hat{\theta}_2(X_1, X_2,\dots, X_n)$

是未知参数θ 的两个无偏估计量, 如果 $D(\hat{\theta}_1)\le D(\hat{\theta}_2)$

则称 $\hat{\theta}_1$比$\hat{\theta}_2$有效 (**样本容量n固定**)

3. 相合性: 设$\hat{\theta}(X_1, X_2,\dots, X_n)$是未知参数$\theta$ 的估计量,

若对任意的ε>0, 有 $\lim\limits_{n\to\infty} P\{|\hat{\theta}-\theta|<\varepsilon\}=1$ 即$\hat{\theta}\xrightarrow{P}\theta$

则称 $\hat{\theta}$为θ 的相合估计或一致估计.

(**样本容量n充分大**,无无偏性的要求)

1. 无偏性的意义

$$E(\hat{\theta}-\theta)=0 \qquad \text{或} \qquad E(\hat{\theta})=\theta(\text{参数})$$

误差

- 随机误差：操作上和其他随机性原因，可消除(求平均);
- 系统误差：本身结构上问题(不能通过多次平均消除).

一个估计量如果不是无偏的，就称为有偏估计，称 $E(\hat{\theta}-\theta)$ 为估计量 $\hat{\theta}$ 的偏差，在科学技术中也称之为 $\hat{\theta}$ 的系统误差.

<u>无偏性的实际意义：</u>没有系统误差，消除随机误差.

S^2 是总体方差σ^2的无偏估计量(重要的结论)

例1: 设总体的方差$D(X)=\sigma^2>0$, 则样本方差

S^2 是σ^2的无偏估计.

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

$$=\sum_{i=1}^{n}X_i^2-2\bar{X}\sum_{i=1}^{n}X_i+n\bar{X}^2$$

$$=\sum_{i=1}^{n}X_i^2-n\bar{X}^2$$

S^2 是总体方差σ^2的无偏估计量(重要的结论)

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

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

样本独立同分布性质

$$=nE(X^2)-nE(\bar{X}^2)$$

$$=n\{D(X)+E(X)^2\}-n\{D(\bar{X})+E(\bar{X})^2\}$$

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

$$=(n-1)\sigma^2$$

$$\therefore E(S^2)=\sigma^2$$

重要结论1：S^2 是总体方差 σ^2 的无偏估计

注：$M_2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2=\frac{n-1}{n}S^2$ 不是 σ^2 的无偏估计

重要结论2：总体均值 μ 已知时，$\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2$ 是总体方差 σ^2 的无偏估计

$$E\left[\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2\right]\overset{\text{样本独立同分布的性质}}{=}\frac{1}{n}\sum_{i=1}^{n}E[X_i-E(X_i)]^2$$

$$=\frac{1}{n}\sum_{i=1}^{n}D(X_i)=\frac{1}{n}\times n\times\sigma^2=\sigma^2$$

思考：下列估计量是否总体均值 $E(X)$ 的无偏估计量？哪个更好？

1. $\bar{X}$ 2. X_1 3. $\frac{X_1+X_2}{2}$ 4. $0.1X_1+0.2X_2+0.7X_3$

注：一个未知参数可以有不同且很多的无偏估计量.

无偏估计性及意义 $E(\hat{\theta})=\theta$ 或 $E(\hat{\theta}-\theta)=0$

优点：没有系统误差, 相当于从平均意义上使随机误差刚好正、负抵消。

缺陷：但并不保证随机误差很小.

若随机误差很大，无偏估计给出的估计值并不令人放心的.

例：原料对产品质量偏差无法抵消.

一个优良的估计量，其方差应该较小，以控制随机误差.

2. 有效性

定义：设 $\hat{\theta}_1(X_1, X_2, \ldots, X_n)$ 和 $\hat{\theta}_2(X_1, X_2, \ldots, X_n)$ 是未知参数θ 的两个无偏估计量,

若对θ 的所有可能取值都有 $D(\hat{\theta}_1) \le D(\hat{\theta}_2)$

则称 $\hat{\theta}_1$比$\hat{\theta}_2$有效

设 $\hat{\theta}_0$ 是θ 的无偏估计，如果对θ 的任何一个无偏估计量 $\hat{\theta}$ 都有 $D(\hat{\theta}_0) \le D(\hat{\theta})$

则称 $\hat{\theta}_0$为θ 的最小方差无偏估计量。

证明无偏性并判断哪个有效例2和例3

例2:设$X_1, X_2 \ldots X_n$为总体X的一个样本, 证明:

$$\hat{\mu}=\sum_{i=1}^{n} c_i X_i, c_i \geq 0, \sum_{i=1}^{n} c_i = 1$$是总体均值无偏估计量,

其中 $\overline{X}$ 是其中最有效估计量.

证：$$E(\hat{\mu}) = E(\sum_{i=1}^{n} c_i X_i) = E(X)\sum_{i=1}^{n} c_i = E(X)$$

$$D(\hat{\mu}) = D(\sum_{i=1}^{n} c_i X_i) = \sigma^2 \sum_{i=1}^{n} c_i^2 \leq \sigma^2$$

函数$f(c_1, c_2, \cdots, c_n) = \sum_{i=1}^{n} c_i^2$ 条件极小值

利用拉格朗日乘数法求条件极值, 令

$$L(c_1, c_2, \cdots, c_n; \ \lambda) = \sum_{i=1}^{n} c_i^2 + \lambda(1 - \sum_{i=1}^{n} c_i)$$

例2 证明: $\hat{\mu}=\sum_{i=1}^{n}c_iX_i, c_i\ge 0, \sum_{i=1}^{n}c_i=1$是总体均值无偏估计量,

其中 $\overline{X}$ 是其中最有效估计量.

$$L(c_1,c_2\cdots c_n;\ \lambda)=\sum_{i=1}^{n}c_i^2+\lambda(1-\sum_{i=1}^{n}c_i)$$

$$\begin{cases}\dfrac{\partial L}{\partial c_i}=2c_i-\lambda=0;\ (i=1,2,\cdots,n)\\ \sum_{i=1}^{n}c_i=1.\end{cases}$$

解得: $\lambda=\dfrac{2}{n}\qquad c_i=\dfrac{1}{n},\qquad i=1,2,\cdots,n$

即函数

$$f\ (c_1,c_2,\cdots,c_n)=\sum_{i=1}^{n}c_i^2$$

的最小值点是$\left(\dfrac{1}{n},\dfrac{1}{n},\cdots,\dfrac{1}{n}\right)$

证明无偏性并判断哪个更有效

例3: 设总体$X\sim U(0,\theta),\theta>0$ 未知, (X_1, X_2, X_3)是取自X的一个样本:

1) 试证 $\hat{\theta}_1=\frac{4}{3}\max\limits_{1\le i\le 3}X_i, \hat{\theta}_2=4\min\limits_{1\le i\le 3}X_i, \hat{\theta}_3=2\overline{X}$ 都是θ的无偏估计;

2) 上述三个估计量中哪个方差最小?

回忆:当 $X\sim U(0,\theta)$, θ的矩法估计量为$2\overline{X}$，而极大似然估计量为 $\max\limits_{1\le i\le n}\{X_i\}$.

分析: 要判断估计量是否是无偏估计量,
需要计算统计量的数学期望.

令 $Y=\max\limits_{1\le i\le 3}X_i,\ Z=\min\limits_{1\le i\le 3}X_i$

证明: 1) 先求$Y=\max\limits_{1\le i\le 3} X_i$和$Z=\min\limits_{1\le i\le 3} X_i$的概率密度

已知 $X\sim U(0,\theta)$, $F_X(x)=\begin{cases}0, & x\le 0;\\ \dfrac{x}{\theta}, & 0<x<\theta;\\ 1, & x\ge\theta.\end{cases}$

$$F_Y(y)=P\{Y\le y\}=P\{\max_{1\le i\le 3} X_i\le y\}$$
$$=P\{X_1\le y, X_2\le y, X_3\le y\}$$
$$=P\{X_1\le y\}\cdot P\{X_2\le y\}\cdot P\{X_3\le y\}=[F_X(y)]^3$$

$$F_Y(y)=\begin{cases}0, & y\le 0\\ \left(\dfrac{y}{\theta}\right)^3, & 0<y<\theta\\ 1, & y\ge\theta\end{cases}\quad \therefore\ f_Y(y)=\begin{cases}\dfrac{3}{\theta}\cdot\left(\dfrac{y}{\theta}\right)^2, & 0<y<\theta\\ 0, & else\end{cases}$$

先求$Y=\max\limits_{1\le i\le 3} X_i$和$Z=\min\limits_{1\le i\le 3} X_i$的概率密度函数:

$$\therefore f_Y(y)=\begin{cases}\dfrac{3}{\theta}\cdot\left(\dfrac{y}{\theta}\right)^2, & 0<y<\theta\\ 0, & else\end{cases}$$

$$\therefore E(Y)=\frac{3}{\theta^3}\int_0^\theta y^3dy=\frac{3}{4}\theta$$

同理可得, $f_Z(z)=\begin{cases}\dfrac{3}{\theta}\cdot\left(1-\dfrac{z}{\theta}\right)^2, & 0\le z\le\theta\\ 0, & else\end{cases}$

$$\therefore f_Y(y)=\begin{cases}\dfrac{3}{\theta}\cdot\left(\dfrac{y}{\theta}\right)^2, & 0\le y\le\theta\\ 0\quad, & else\end{cases}\qquad f_Z(z)=\begin{cases}\dfrac{3}{\theta}\cdot\left(1-\dfrac{z}{\theta}\right)^2, & 0\le z\le\theta\\ 0\quad, & else\end{cases}$$

$$\therefore E(Y)=\frac{3}{\theta^3}\int_0^\theta y^3dy=\frac{3}{4}\theta\qquad \hat{\theta}_1=\frac{4}{3}\max_{1\le i\le 3}X_i,\ \hat{\theta}_2=4\min_{1\le i\le 3}X_i,\ \hat{\theta}_3=2\overline{X}$$

$$\therefore E(Z)=\frac{3}{\theta^3}\int_0^\theta z\cdot(\theta-z)^2dz=\frac{1}{4}\theta\qquad Y=\max_{1\le i\le 3}X_i,\ Z=\min_{1\le i\le 3}X_i$$

$$\text{从而},E\left(\frac{4}{3}\max_{1\le i\le 3}X_i\right)=E\left(4\min_{1\le i\le 3}X_i\right)=\theta$$

$$E(\hat{\theta}_3)=E(2\overline{X})=2\cdot\frac{\theta}{2}=\theta$$

从而 $\hat{\theta}_1=\frac{4}{3}\max\limits_{1\le i\le 3}X_i,\ \hat{\theta}_2=4\min\limits_{1\le i\le 3}X_i,\ \hat{\theta}_3=2\overline{X}$ 都是 θ 的无偏估计.

$$\hat{\theta}_1 = \frac{4}{3}\max_{1\le i\le 3} X_i,\ \hat{\theta}_2 = 4\min_{1\le i\le 3} X_i,\ \hat{\theta}_3 = 2\overline{X} \qquad Y = \max_{1\le i\le 3} X_i,\ Z = \min_{1\le i\le 3} X_i$$

2) 上述三个估计量中哪个更有效?

$$\because D(Y) = E(Y^2) - E(Y)^2 = \frac{3}{80}\theta^2$$

$$D(Z) = E(Z^2) - E(Z)^2 = \frac{3}{80}\theta^2 \qquad D(\hat{\theta}_3) = D(2\overline{X}) = 4\cdot\frac{1}{3}\cdot\frac{\theta^2}{12}$$

$$\therefore D(\frac{4}{3}Y) = \frac{16}{9}D(Y) = \frac{16}{9}\cdot\frac{3}{80}\theta^2 \le D(4Z) = 16D(Z) = 16\cdot\frac{3}{80}\theta^2$$

$$D(\hat{\theta}_1) < D(\hat{\theta}_3)$$

即 $\frac{4}{3}\max\limits_{1\le i\le 3} X_i$ 最有效.

注: 这是最大似然估计优于矩估计的有名例子.

重要结论: $\overline{X}$ 和S^2 分别是μ和σ^2 的最小方差无偏估计.

1. 无偏性

定义：设 $\hat{\theta}(X_1, X_2, ..., X_n)$ 是未知参数θ 的估计量，

若$E(\hat{\theta})=\theta$,则称$\hat{\theta}$为θ的无偏估计. 等价: $E(\hat{\theta}-\theta)=0$

2. 有效性: 设 $\hat{\theta}_1(X_1, X_2, ..., X_n)$ 和 $\hat{\theta}_2(X_1, X_2, ..., X_n)$

是未知参数θ 的两个无偏估计量, 如果 $D(\hat{\theta}_1)\le D(\hat{\theta}_2)$

则称 $\hat{\theta}_1$比$\hat{\theta}_2$有效

估计量$\hat{\theta}=\hat{\theta}(X_1, X_2, \cdots X_n)$与样本容量$n$也有关系,

随着n增加, 估计量的值越来越趋向于被估参数的真值, 提出了

3. 相合性: 设$\hat{\theta}(X_1, X_2, ..., X_n)$是未知参数$\theta$ 的估计量,

若对任意的ε>0, 有 $\lim\limits_{n\to\infty} P\{|\hat{\theta}-\theta|<\varepsilon\}=1$ 即 $\hat{\theta}\xrightarrow{P}\theta$

则称 $\hat{\theta}$为θ 的相合估计或一致估计.

相合估计量的证明

例4 设 $X\sim N(0,\sigma^2)$, 证明: $\frac{1}{n}\sum_{i=1}^{n}X_i^2$是$\sigma^2$的相合估计量.

分析: 1) 证明相合性往往用到切比雪夫不等式, 其中涉及期望与方差;

$$P\{|X-E(X)|\ge\varepsilon\}\le\frac{D(X)}{\varepsilon^2}$$

2) 这里计算方差较难, 可以先化为χ^2分布, 再利用卡方分布的性质计算.

令$Y=\sum_{i=1}^{n}\left(\frac{X_i}{\sigma}\right)^2\sim\chi^2(n)$　则$\frac{1}{n}\sum_{i=1}^{n}X_i^2=\frac{\sigma^2}{n}Y$

例4. 设 $X\sim N(0,\sigma^2)$，证明：$\frac{1}{n}\sum_{i=1}^{n}X_i^2$ 是σ^2的相合估计量。

证明：

$$E\left(\frac{1}{n}\sum_{i=1}^{n}X_i^2\right)=\frac{1}{n}\sum_{i=1}^{n}E\left(X_i^2\right)=\frac{1}{n}nE\left(X^2\right)=\sigma^2$$

令 $Y=\sum_{i=1}^{n}\left(\frac{X_i}{\sigma}\right)^2\sim\chi^2(n)$，则 $\frac{1}{n}\sum_{i=1}^{n}X_i^2=\frac{\sigma^2}{n}Y$

$$\therefore D\left(\frac{1}{n}\sum_{i=1}^{n}X_i^2\right)=D\left(\frac{\sigma^2}{n}Y\right)$$

$$=\frac{\sigma^4}{n^2}\cdot D(Y)=\frac{\sigma^4}{n^2}\cdot 2n=\frac{2\sigma^4}{n}$$

例4.设 $X\sim N(0,\sigma^2)$，证明：$\frac{1}{n}\sum_{i=1}^{n}X_i^2$ 是σ^2的相合估计量。

由切比雪夫不等式, 有

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i^2 - E\left(\frac{1}{n}\sum_{i=1}^{n}X_i^2\right)\right| \geq \varepsilon\right\} \leq \frac{D(\frac{1}{n}\sum_{i=1}^{n}X_i^2)}{\varepsilon^2}$$

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i^2 - \sigma^2\right| \geq \varepsilon\right\} \leq \frac{D\left(\frac{1}{n}\sum_{i=1}^{n}X_i^2\right)}{\varepsilon^2}$$

$$= \frac{2\sigma^4}{n\varepsilon^2} \xrightarrow[\varepsilon\text{固定}]{n\to\infty} 0$$

故 $\frac{1}{n}\sum_{i=1}^{n}X_i^2$ 是σ^2 的相合估计量.

小 结

重要结论: $\overline{X}$ 和S^2 分别是μ和σ^2 的最小方差无偏估计;

$\overline{X}$是μ(总体期望)的相合估计量; (例7.2.4)

S^2 和M_2 都是σ^2(总体方差)的相合估计量。

样本的k 阶原点矩A_k为总体矩γ_k的相合估计估计量

注: 由于相合性是在极限意义下定义的.

因此, 只有当样本容量充分大时, 才显示出优越性,

而在实际生活中往往难以增大样本容量,

而且证明估计量的相合性并非容易, 因此, 在实际生活

常常使用无偏性和有效性这两个标准(固定样本容量n).