§7.2 点估计的评价标准

对于同一个未知参数，不同的方法得到的估计量可能不同，于是提出问题

- 应该选用哪一种估计量？
- 用何标准来评价一个估计量的好坏？

常用标准

(1) 无偏性

(2) 有效性

(3) 一致性

无偏性

定义 若 $E(\hat{\theta}) = \theta$

则称 $\hat{\theta}$ 是 θ 的无偏估计量.

定义的合理性

我们不可能要求每一次由样本得到的估计值与真值都相等，但可以要求这些估计值的期望与真值相等.

例1 设总体X的k阶矩 $\mu_k = E(X^k)$存在

$(X_1, X_2, \cdots, X_n)$ 是总体X的样本,

证明: 不论X服从什么分布(但期望存在),

则 $A_k = \dfrac{1}{n}\sum_{i=1}^{n} X_i^k$ 是μ_k 的无偏估计量.

证 由于$E(X_i^k) = \mu_k \quad i = 1,2,\cdots,n$ 因而

$$E(A_k) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i^k\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i^k)$$

$$= \frac{1}{n}\cdot n \cdot \mu_k = \mu_k$$

特别地

样本均值 $\overline{X}$ 是总体期望 $E(X)$ 的无偏估计量

样本二阶原点矩 $A_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$ 是总体二阶原点矩 $\mu_2 = E(X^2)$ 的无偏估计量

例2 设总体 X 的期望 与方差存在, X 的样本为 $(X_1, X_2, \cdots, X_n)$ $(n>1)$. 证明

(1) $S_n^2 = \dfrac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$ 不是 $D(X)$ 的无偏估量;

(2) $S^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ 是 $D(X)$ 的无偏估计量.

证 前已证 $\dfrac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \dfrac{1}{n}\sum_{i=1}^{n}X_i^2 - \bar{X}^2$

$$E(X_i) = E(X) = \mu,\ D(X_i) = D(X) = \sigma^2$$

$$E(\bar{X}) = E(X) = \mu,\ D(\bar{X}) = \frac{\sigma^2}{n}$$

因而

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

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

$$=\frac{n-1}{n}\sigma^2\neq\sigma^2$$

故 $E\left(\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2\right)=\sigma^2$ 证毕.

例3 设$(X_1, X_2, \cdots, X_m)$是总体X的一个样本，$X \sim B(n, p)$ $n > 1$，求p^2的无偏估计量.

解 由于样本矩是总体矩的无偏估计量以及数学期望的线性性质，只要将未知参数表示成总体矩的线性函数，然后用样本矩作为总体矩的估计量，这样得到的未知参数的估计量即为无偏估计量.

令 $\overline{X} = E(X) = np$

$$\frac{1}{m}\sum_{i=1}^{m} X_i^2 = E(X^2) = (np)^2 + np(1-p)$$

故 $$(n^2-n)p^2=\frac{1}{m}\sum_{i=1}^{m}X_i^2-\overline{X}$$

因此, p^2 的无偏估计量为

$$\hat{p}^2=\frac{1}{n^2-n}\left(\frac{1}{m}\sum_{i=1}^{m}X_i^2-\overline{X}\right)$$

$$=\frac{\frac{1}{m}\sum_{i=1}^{m}X_i(X_i-1)}{n(n-1)}$$

例4 设总体 X 的密度函数为

$$f(x;\theta)=\begin{cases}\dfrac{1}{\theta}e^{-\frac{x}{\theta}} & x>0,\\ 0 & x\le 0\end{cases}\qquad \theta>0 \text{ 为常数}$$

$(X_1,X_2,\cdots,X_n)$ 为 X 的一个样本

证明 $\overline{X}$ 与 $n\min\{X_1,X_2,\cdots,X_n\}$ 都是θ 的无偏估计量

证 $X\sim E\left(\dfrac{1}{\theta}\right)$ $\qquad E(X)=\theta$

故 $E(\overline{X})=E(X)=\theta$

$\overline{X}$ 是θ 的无偏估计量.

令 $Z=\min\{X_1,X_2,\cdots,X_n\}$

$$F_Z(z)=1-P(X_1>z,X_2>z,\cdots,X_n>z)$$

$$=1-P(X_1>z)P(X_2>z)\cdots P(X_n>z)$$

$$=1-\prod_{i=1}^{n}(1-P(X_i\le z))=\begin{cases}0 & z<0\\ 1-e^{-\frac{nz}{\theta}} & z\ge 0\end{cases}$$

$$f_Z(z)=\begin{cases}0 & z<0\\ \dfrac{n}{\theta}e^{-\frac{nz}{\theta}} & z\ge 0\end{cases}$$

即 $Z\sim E\left(\dfrac{n}{\theta}\right)$ $\quad E(Z)=\dfrac{\theta}{n}$ $\quad E(nZ)=\theta$

故 nZ 是 θ 的无偏估计量.

有效性

定义 设 $\hat{\theta}_1=\theta_1(X_1,X_2,\cdots,X_n)$

$$\hat{\theta}_2=\theta_2(X_1,X_2,\cdots,X_n)$$

都是总体参数θ的无偏估计量, 且

$$D(\hat{\theta}_1)<D(\hat{\theta}_2)$$

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

例5 设总体 X 的密度函数为

$$f(x;\theta)=\begin{cases}\dfrac{1}{\theta}e^{-\frac{x}{\theta}} & x>0,\\ 0 & x\le 0\end{cases}\qquad \theta>0 \text{ 为常数}$$

由例4可知，$\bar{X}$ 与 $n\min\{X_1,X_2,\cdots,X_n\}$ 都是θ的无偏估计量,问哪个估计量更有效?

解 $D(\bar{X})=\dfrac{\theta^2}{n}$，$D(n\min\{X_1,X_2,\cdots,X_n\})=\theta^2$

所以,$\bar{X}$ 比$n\min\{X_1,X_2,\cdots,X_n\}$ 更有效.

例6 设总体 X，且 $E(X)=\mu$，$D(X)=\sigma^2$

$(X_1, X_2, \cdots, X_n)$ 为总体 X 的一个样本

(1) 设常数 $c_i \neq \dfrac{1}{n}$ $i=1,2,\cdots,n.$ $\sum_{i=1}^{n} c_i = 1.$

证明 $\hat{\mu}_1 = \sum_{i=1}^{n} c_i X_i$ 是 μ 的无偏估计量

(2) 证明 $\hat{\mu} = \overline{X}$ 比 $\hat{\mu}_1 = \sum_{i=1}^{n} c_i X_i$ 更有效

证 (1) $E(\hat{\mu}_1) = \sum_{i=1}^{n} c_i E(X_i) = \sum_{i=1}^{n} c_i \mu = \mu$

(2) $$D(\hat{\mu}_1)=\sum_{i=1}^{n}c_i^2 D(X_i)=\sigma^2\sum_{i=1}^{n}c_i^2$$

而 $$1=\left(\sum_{i=1}^{n}c_i\right)^2=\sum_{i=1}^{n}c_i^2+2\sum_{1\le i<j\le n}c_i c_j$$

$$<\sum_{i=1}^{n}c_i^2+\sum_{1\le i<j\le n}(c_i^2+c_j^2)=n\sum_{i=1}^{n}c_i^2$$

$$\Rightarrow \sum_{i=1}^{n}c_i^2>\frac{1}{n} \quad \Rightarrow D(\hat{\mu})=\frac{1}{n}\sigma^2<D(\hat{\mu}_1)$$

结论 算术均值比加权均值更有效.

例如 $X \sim N(\mu,\sigma^2)$, (X_1,X_2) 是一样本.

$$\left.\begin{aligned}
\hat{\mu}_1 &= \frac{2}{3}X_1 + \frac{1}{3}X_2 \\
\hat{\mu}_2 &= \frac{1}{4}X_1 + \frac{3}{4}X_2 \\
\hat{\mu}_3 &= \frac{1}{2}X_1 + \frac{1}{2}X_2
\end{aligned}\right\} \text{都是}\mu\text{的无偏估计量}$$

由例6(2) 知 $\hat{\mu}_3$ 最有效.

罗—克拉美（$Rao-Cramer$）不等式

若 $\hat{\theta}$ 是参数 θ 的无偏估计量, 则

$$D(\hat{\theta}) \ge \frac{1}{nE\left[\dfrac{\partial}{\partial\theta}\ln p(X,\theta)\right]^2} = D_0(\theta)$$

其中 $p(x,\theta)$ 是总体 X 的概率分布或密度函数，称 $D_0(\theta)$为方差的下界.

当 $D(\hat{\theta}) = D_0(\theta)$ 时, 称 $\hat{\theta}$ 为达到方差下界的无偏估计量, 此时称 $\hat{\theta}$ 为最有效的估计量, 简称有效估计量.

例7 设总体 X 的密度函数为

$$f(x;\theta)=\begin{cases}\dfrac{1}{\theta}e^{-\frac{x}{\theta}} & x>0, \\ 0 & x\le 0\end{cases}\qquad \theta>0 \text{ 为常数}$$

$(x_1,x_2,\cdots,x_n)$ 为 X 的一个样本值.

求 θ 的极大似然估计量, 并判断它是否达到方差下界的无偏估计量.

解 由似然函数

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

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

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{x}$$

θ的极大似然估计量为 $\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$

它是θ的无偏估计量.

$$D(\hat{\theta}) = D\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{\theta^2}{n}$$

而 $$\ln f(x,\theta) = -\ln\theta - \frac{x}{\theta}$$

$$\left[\frac{\partial}{\partial\theta}\ln f(x,\theta)\right]^2 = \left[-\frac{1}{\theta} + \frac{x}{\theta^2}\right]^2$$

$$E\left[\frac{\partial}{\partial\theta}\ln f(X,\theta)\right]^2 = E\left[-\frac{1}{\theta} + \frac{X}{\theta^2}\right]^2 = \frac{1}{\theta^2}$$

$$\Rightarrow \quad \frac{1}{nE\left[\frac{\partial}{\partial\theta}\ln f(X,\theta)\right]^2} = \frac{\theta^2}{n} = D(\overline{X})$$

故 $\overline{X}$ 是达到方差下界的无偏估计量.

一致性

定义 设 $\hat{\theta}=\hat{\theta}(X_1,X_2,\cdots,X_n)$ 是总体参数 θ 的估计量. 若对于任意的 $\theta\in\Theta$, 当 $n\to\infty$ 时, $\hat{\theta}$ 依概率收敛于 θ, 即 $\forall\varepsilon>0$,

$$\lim_{n\to\infty}P\left(\left|\hat{\theta}-\theta\right|\geq\varepsilon\right)=0$$

则称 $\hat{\theta}$ 是总体参数 θ 的一致(或相合)估计量.

一致性估计量仅在样本容量 n 足够大时, 才显示其优越性.

关于一致性的两个常用结论

1. 样本 k 阶矩是总体 k 阶矩的一致性估计量. } 由大数定律证明

2. 设 $\hat{\theta}$ 是 θ 的无偏估计量, 且 $\lim\limits_{n\to\infty} D(\hat{\theta}) = 0$, 则 $\hat{\theta}$ 是 θ 的一致估计量. } 用切贝雪夫不等式证明

矩法得到的估计量一般为一致估计量

在一定条件下, 极大似然估计具有一致性

例8 $X \sim f(x;\theta)=\begin{cases}\dfrac{1}{\theta}e^{-\frac{x}{\theta}} & x>0,\\ 0 & x\le 0\end{cases}$ $\theta>0$ 为常数

则 $\overline{X}$ 是θ的无偏、有效、一致估计量.

证 由例7 知 $\overline{X}$ 是θ的无偏、有效估计量.

$$\lim_{n\to\infty} D(\overline{X})=\lim_{n\to\infty}\frac{\theta^2}{n}=0$$

所以$\overline{X}$ 是 θ 的一致估计量, 证毕.

作业 P.231 习题七

15　　16

18　　20

补充题　设总体 $X \sim N(\mu, \sigma^2)$, $(X_1, X_2, \cdots, X_n)$ 为 X 的一个样本,常数 k 取何值可使 $k\sum_{i=1}^{n}|X_i - \bar{X}|$ 为 σ 的无偏估计量

第十四周　问　题

母亲嗜酒是否影响下一代的健康

美国的Jones医生于1974年观察了母亲在妊娠时曾患慢性酒精中毒的6名七岁儿童（称为甲组）.以母亲的年龄，文化程度及婚姻状况与前6名儿童的母亲相同或相近，但不饮酒的46名七岁儿童为对照租(称为乙组). 测定两组儿童的智商，结果如下:

智商 / 组别	人数 n	智商平均数 $\bar{x}$	样本标准差 S
甲 组	6	78	19
乙 组	46	99	16

由此结果推断母亲嗜酒是否影响下一代的智力？若有影响，推断其影响程度有多大?

提示 前一问题属假设检验问题

后一问题属区间估计问题