第二节 估计量的评价标准

对于未知参数θ，由于其估计量$\hat{\theta}$在一般情况下并不惟一，因此在实际问题中，选用合适的统计量以取得较好的效果，具有非常重要的意义.

一、无偏性

定义 1 设$\hat{\theta}$为θ的估计量，如果对任意的$\theta\in\Theta$，均有$E\hat{\theta}=\theta$，就称$\hat{\theta}$为θ的无偏估计．否则称为有偏估计．

无偏估计的直观意义：由于样本$(X_1,X_2,\cdots,X_n)$是随机的，利用$\hat{\theta}=\hat{\theta}(X_1,X_2,\cdots,X_n)$估计$\theta$时，有时会偏高，有时会偏低，但整体平均来说等于$\theta$．

讨论无偏性的关键在于计算$E\hat{\theta}$．

例 1 设总体 $X \sim N(\mu,\sigma^2)$，$(X_1, X_2,\cdots, X_n)$ $(n>1)$ 为来自总体 X 的一个简单随机样本，(1)问 $S_0^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2$ 是否是 σ^2 的无偏估计？(2)如果 $Y = c\sum_{i=1}^{n-1}(X_{i+1}-X_i)^2$ 为 σ^2 的无偏估计，求常数 c .

解 (1)由于 $E(S_0^2)=\sigma^2$，所以 S_0^2 是 σ^2 的无偏估计.

(2)由于 $E(X_{i+1}-X_i)=0, D(X_{i+1}-X_i)=2\sigma^2$，所以

$$E(X_{i+1}-X_i)^2 = 2\sigma^2+0^2 = 2\sigma^2，\quad i=1,2,\cdots,n-1，$$

故 $EY = c\sum_{i=1}^{n-1}E(X_{i+1}-X_i)^2 = c\sum_{i=1}^{n-1}2\sigma^2 = c2(n-1)\sigma^2$.

由 $EY=\sigma^2$，解得 $c=\dfrac{1}{2(n-1)}$.

定理 1 设总体 X 的数学期望 $EX=\mu$，方差 $DX=\sigma^2$，$(X_1, X_2, \cdots, X_n)$ $(n>1)$ 为来自总体 X 的样本，则

(1) $\overline{X}$ 是 μ 的无偏估计，即 $E\overline{X}=\mu$；

(2) S^2 是 σ^2 的无偏估计，即 $E(S^2)=\sigma^2$.

定理 2 设估计量 $\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_m$ 均为 θ 的无偏估计，$c_1, c_2, \cdots, c_m$ 为常数，且 $\sum_{i=1}^{m} c_i = 1$，则 $\sum_{i=1}^{m} c_i \hat{\theta}_i$ 仍为 θ 的无偏估计.

证 由于 $E(\sum_{i=1}^{m} c_i \hat{\theta}_i) = \sum_{i=1}^{m} c_i E(\hat{\theta}_i) = \sum_{i=1}^{m} c_i \theta = \theta \sum_{i=1}^{m} c_i = \theta$，

所以 $\sum_{i=1}^{m} c_i \hat{\theta}_i$ 为 θ 的无偏估计.

例 2 设总体 $X \sim P(\lambda)$，对任意的常数 $c \in (0,1)$，问 $c\bar{X}+(1-c)S^2$ 是否为 λ 的无偏估计？

解 由于 $E\bar{X}=EX=\lambda, E(S^2)=DX=\lambda$，故

$$E[c\bar{X}+(1-c)S^2]=cE\bar{X}+(1-c)E(S^2)=c\lambda+(1-c)\lambda=\lambda，$$

所以 $c\bar{X}+(1-c)S^2$ 为 λ 的无偏估计.

又解 由定理 1，$\bar{X}$ 和 S^2 均为 λ 的无偏估计. 且 $c+(1-c)=1$，再由定理 2 知，$c\bar{X}+(1-c)S^2$ 为 λ 的无偏估计.

例 3 设总体 $X \sim N(\mu, \sigma^2)$，由定理 1 知 $\overline{X}$ 是 μ 的无偏估计，问 $\overline{X}^2$ 是否为 μ^2 的无偏估计？

解 由于

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

所以 $\overline{X}^2$ 不是 μ^2 的无偏估计.

注： 本例表明：虽然 $E\overline{X} = \mu$，但 $E(\overline{X}^2) \neq \mu^2$.

一般地，虽然 $E\hat{\theta} = \theta$，但未必有 $Eg(\hat{\theta}) = g(\theta)$，即如果 $\hat{\theta}$ 为 θ 的无偏估计，但 $g(\hat{\theta})$ 未必为 $g(\theta)$ 的无偏估计.

例 4 设总体 X 的密度函数为

$$f(x;\theta)=\begin{cases} e^{-(x-\theta)}, & x\geq\theta, \\ 0, & x<\theta, \end{cases}$$

$(X_1,X_2,\cdots,X_n)$ 为来自总体 X 的样本．试分别讨论未知参数 θ 的估计量 $\hat{\theta}_M=\overline{X}-1$ 和 $\hat{\theta}_L=\min\limits_{1\leq i\leq n}X_i$ 的无偏性．

解 (1)由于 $EX=\int_{\theta}^{+\infty}xe^{-(x-\theta)}dx=1+\theta$，得 $E\overline{X}=1+\theta$，故

$$E\hat{\theta}_M=E(\overline{X}-1)=E\overline{X}-1=(1+\theta)-1=\theta,$$

所以 $\hat{\theta}_M=\overline{X}-1$ 是 θ 的无偏估计．

续解 (2) X 的分布函数为 $F(x)=\begin{cases}1-e^{-(x-\theta)}, & x\geq\theta,\\ 0, & x<\theta.\end{cases}$ 故

$\hat{\theta}_L=\min\limits_{1\leq i\leq n}X_i$ 的密度函数为

$$f_{\hat{\theta}_L}(x)=n[1-F(x)]^{n-1}f(x)=\begin{cases}ne^{-n(x-\theta)}, & x\geq\theta,\\ 0, & x<\theta.\end{cases}$$

由于 $$E\hat{\theta}_L=\int_{\theta}^{+\infty}x\cdot ne^{-n(x-\theta)}dx=\theta+\frac{1}{n}\neq\theta,$$

所以 $\hat{\theta}_L=\min\limits_{1\leq i\leq n}X_i$ 不是 θ 的无偏估计，即为有偏估计.

二、有效性

定义 2 设 $\hat{\theta}_1, \hat{\theta}_2$ 均为 θ 的无偏估计，如果 $D\hat{\theta}_1 < D\hat{\theta}_2$，就称 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 有效.

讨论有效性的关键在于计算 $D\hat{\theta}$.

例 5 设 $(X_1, X_2, \cdots, X_n)$ $(n>1)$ 为来自总体 X 的样本，且 $EX=\mu$，$DX=\sigma^2$ $(\sigma>0)$，问 μ 的估计量 $\mu_1=X_1$ 和 $\mu_2=\overline{X}$ 中，哪个更有效?

解 由于

$$E\mu_1=EX_1=\mu, E\mu_2=E\overline{X}=\mu,$$

故 $\mu_1=X_1$ 和 $\mu_2=\overline{X}$ 均为 μ 的无偏估计，又由于

$$D\mu_1=DX_1=\sigma^2, \quad D\mu_2=D\overline{X}=\frac{\sigma^2}{n},$$

有 $D\mu_2<D\mu_1$，所以 $\mu_2=\overline{X}$ 比 $\mu_1=X_1$ 更有效.

例 6 设总体 $X \sim N(\mu,\sigma^2)$，其中 μ 已知，σ^2 未知. $(X_1, X_2, \cdots, X_n)$ $(n>1)$ 为来自总体 X 的样本，问 σ^2 的估计量

$$\sigma_1^2 = S_0^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2 \text{ 和 } \sigma_2^2 = S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

中，哪个更有效?

【简解】 由第六章例题知,

$$E(S_0^2) = \sigma^2, E(S^2) = \sigma^2, D(S_0^2) = \frac{2\sigma^4}{n}, D(S^2) = \frac{2\sigma^4}{n-1},$$

所以 $\sigma_1^2 = S_0^2$ 和 $\sigma_2^2 = S^2$ 均为 σ^2 的无偏估计. 且 $D(S_0^2) < D(S^2)$，

所以 $\sigma_1^2 = S_0^2$ 比 $\sigma_2^2 = S^2$ 更有效.

三、一致性（相合性）

定义 5 设 $\hat{\theta}$ 为 θ 的估计量，如果对任意的 $\varepsilon>0$，均有

$$\lim_{n\to\infty} P\{|\hat{\theta}-\theta|<\varepsilon\}=1,$$

就称 $\hat{\theta}$ 为 θ 的一致估计量或相合估计量.

定理 4 设 $(X_1, X_2, \cdots, X_n)$ 是来自总体 X 的样本，且 $E|X|^r$ 存在，其中 r 为正整数. 则 $\frac{1}{n}\sum_{i=1}^{n} X_i^k$ 为 $E(X^k)$ 的相合估计，$k=1,2,\cdots,r$.