

§ 6.1 估计量优劣的评价标准（续）

对于总体分布中的同一个未知参数 θ ，若采用不同的估计方法，可能得到不同估计量 $\hat{\theta}$ 。例如：

均匀总体 $X \sim U(a, b)$ 的未知参数 a, b 的矩估计量和极大似然估计量分别为

$$\left\{ \begin{array}{ll} \hat{a}_{\text{矩}} = \bar{X} - \sqrt{3}S_n, & \hat{b}_{\text{矩}} = \bar{X} + \sqrt{3}S_n \\ \hat{a}_{MLE} = X_{\min}, & \hat{b}_{MLE} = X_{\max} \end{array} \right\}$$

二者
不同

究竟采用哪一个估计量更好呢？这就产生了如何评价与比较估计量的好坏的问题，自然与参数的真值偏离程度越小的估计量越好。

通常用偏差平方的期望 $E(\hat{\theta} - \theta)^2$ 来衡量估计量 $\hat{\theta}$ 的偏离程度，并称为均方误差(MSE)，记作：

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

如果存在一个估计量 $\hat{\theta}$ ，在所有估计量中， $\hat{\theta}$ 的均方误差最小，则称 $\hat{\theta}$ 是 θ 的最优估计量。

均方误差MSE可分解为两项：

$$\begin{aligned} MSE(\hat{\theta}) &= E(\hat{\theta} - \theta)^2 = E\left[(\hat{\theta} - E(\hat{\theta})) + (E(\hat{\theta}) - \theta)\right]^2 \\ &= E(\hat{\theta} - E(\hat{\theta}))^2 + (E(\hat{\theta}) - \theta)^2 \\ &= D(\hat{\theta}) + (E(\hat{\theta}) - \theta)^2. \end{aligned}$$

当上式右端的两项都很小, 左端才能很小. 通常同时控制均方误差右端的两项是很困难的, 但是控制第二项是可能的, 要求 $E(\hat{\theta}) - \theta = 0$, 或者 $E(\hat{\theta}) = \theta$ 是合理的, 也是容易办到的.

下面我们讨论估计量优劣性的三个标准

(1) 无偏性 (2) 有效性 (3) 相合性.

$$\left\{ \begin{array}{ll} \text{无偏性} & \text{有效性} \\ E(\hat{\theta}) = \theta & D(\hat{\theta}) \text{小者} \\ \text{相合性} & \text{渐进无偏性} \end{array} \right. \hat{\theta}_n \xrightarrow{P} \theta \Leftrightarrow \lim_{n \rightarrow \infty} E(\hat{\theta}_n) = \theta, \lim_{n \rightarrow \infty} D(\hat{\theta}_n) = 0.$$

一、无偏性

设 $\hat{\theta} = \hat{\theta}(X_1, X_2, \dots, X_n)$ 是参数 θ 的一个估计量,

如果 $E(\hat{\theta}) = \theta$, 则称 $\hat{\theta}$ 是 θ 的无偏估计(量).

如果 $E(\hat{\theta}) \neq \theta$, 则称 $\hat{\theta}$ 是 θ 的有偏估计(量).

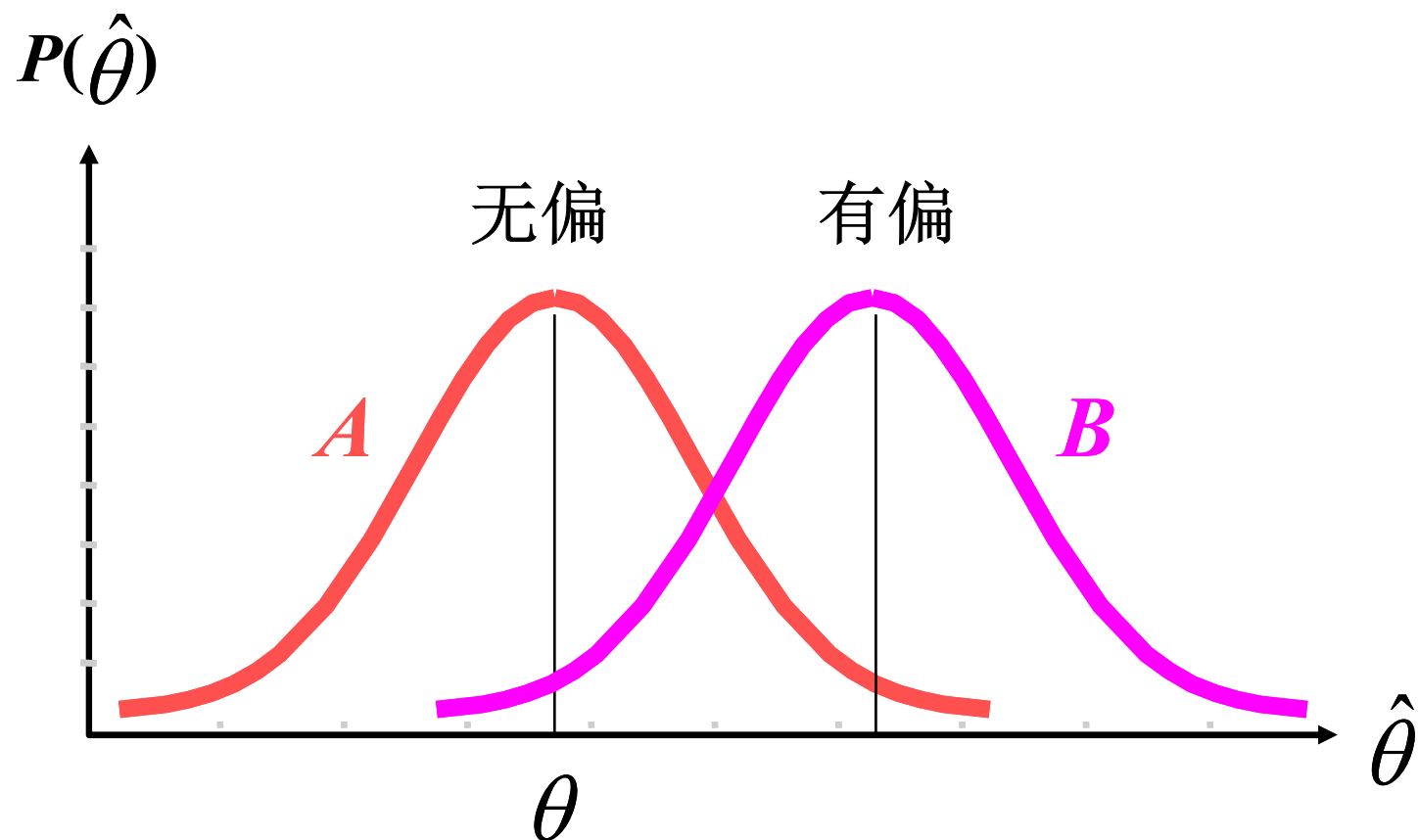
称 $E(\hat{\theta}) - \theta$ 为估计量 $\hat{\theta}$ 的偏差.

如果 θ 的一系列估计 $\hat{\theta}_n = \hat{\theta}_n(X_1, X_2, \dots, X_n)$
($n = 1, 2, \dots$), 满足关系式

$$\lim_{n \rightarrow \infty} E(\hat{\theta}_n) = \theta,$$

则称 $\hat{\theta}_n$ 是 θ 的渐进无偏估计(量).

注：无偏估计量的意义：估计量的观测值总是围绕着真值摆动，其多次试验的平均值等于参数的真值.



例1 设总体 X 的一阶和二阶矩存在, 记

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

则样本均值 \bar{X} 是 μ 的无偏估计, S_n^2 是 σ^2 的渐近无偏估计, 样本方差 S^2 是 σ^2 无偏估计.

证明 $E(\bar{X}) = E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n E(X_i) = \mu.$

无偏
估计

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

渐进
无偏

$$= \frac{1}{n} \left[\sum_{i=1}^n (\mu^2 + \sigma^2) - n\left(\mu^2 + \frac{1}{n} \sigma^2\right) \right] = \frac{n-1}{n} \sigma^2 \neq \sigma^2,$$

$$E(S^2) = \frac{n}{n-1} E(S_n^2) = \frac{n}{n-1} \times \frac{n-1}{n} \sigma^2 = \sigma^2.$$

无偏
估计

例2 设 (X_1, X_2, \dots, 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 将未知参数表示成总体原点矩的线性函数, 然后用样本矩作为总体矩的估计量. 这样即得无偏估计量. 令

$$\bar{X} = E(X) = np,$$

无偏估计

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

无偏估计

$$(n^2 - n)p^2 = \frac{1}{m} \sum_{i=1}^m X_i^2 - \bar{X}. \text{ 因此 } p^2 \text{ 的无偏估计量为}$$

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

无偏估计

例3 设 X_1, X_2, \dots, X_n 是来自总体 X 的一组样本, 总

体 X 的密度函数为
$$f(x; \theta) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

其中 $\theta > 0$ 为未知常数.

试证明 $\hat{\theta}_1 = \bar{X}$, $\hat{\theta}_2 = nX_{(1)}$ 都是 θ 的无偏估计量.

证明 $X \sim E(1/\theta)$, $E(X) = \theta$. 故 $E(\hat{\theta}_1) = E(\bar{X}) = E(X) = \theta$.

$X_{(1)}$ 的密度函数为
$$f_{X_{(1)}}(t) = \begin{cases} \frac{n}{\theta} e^{-\frac{nt}{\theta}}, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

$X_{(1)} \sim E(n/\theta)$, $E(X_{(1)}) = \theta/n$, 故 $E(\hat{\theta}_2) = E(nX_{(1)}) = \theta$.

练习题

设总体 $X \sim U(0, \theta)$, X_1, X_2, \dots, X_n 是来自总体的样本. 试证明

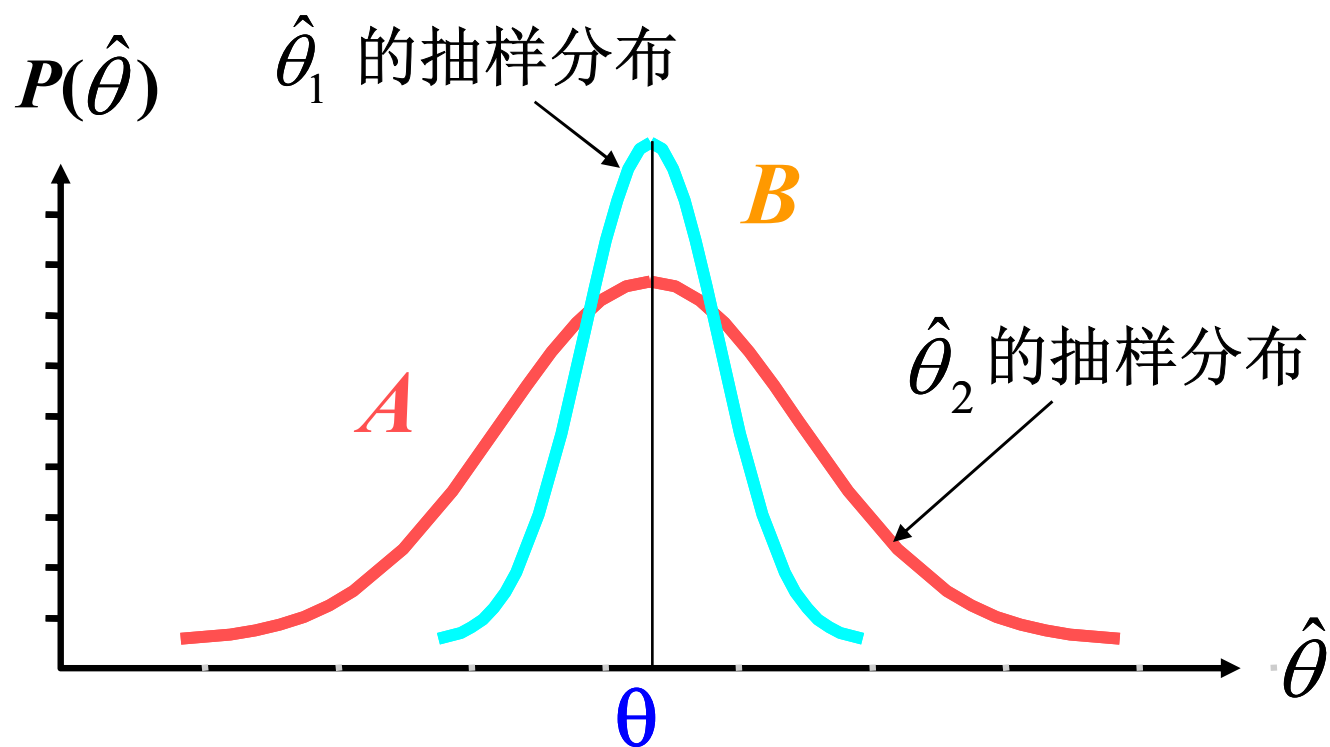
$$(n+1)X_{(1)} = (n+1)\min\{X_1, X_2, \dots, X_n\},$$

$$\frac{n+1}{n}X_{(n)} = \frac{n+1}{n}\max\{X_1, X_2, \dots, X_n\}.$$

都是 θ 的无偏估计.

二、有效性

设 $\hat{\theta}_1 = \theta_1(X_1, X_2, \dots, X_n)$, $\hat{\theta}_2 = \theta_2(X_1, X_2, \dots, X_n)$ 都是 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4 (续例3) 由例4知: $\hat{\theta}_1 = \bar{X}$, $\hat{\theta}_2 = nX_{(1)}$ 都是 θ 的无偏估计量. 问哪个估计量更有效?

解
$$D(\bar{X}) = \frac{\theta^2}{n}, \quad D(nX_{(1)}) = \theta^2.$$

所以估计量 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 更有效.

例5 设总体 $X \sim U [0, \theta]$, 参数 $\theta > 0$, (X_1, X_2, \dots, X_n) 是来自总体 X 样本.

(1) 试证明: θ 的矩估计量 $\hat{\theta}_1 = 2\bar{X}$ 和修正的极大似然估计量 $\hat{\theta}_2 = \frac{n+1}{n} X_{(n)}$ 均是 θ 的无偏估计;

(2) 问: $\hat{\theta}_1$ 和 $\hat{\theta}_2$ 哪一个更有效? ($n \geq 2$).

证 (1) $E(\hat{\theta}_1) = E(2\bar{X}) = 2E(\bar{X}) = 2E(X) = 2 \times \frac{\theta}{2} = \theta,$

所以 $2\bar{X}$ 是 θ 的无偏估计量.

因为 $X \sim U[0, \theta]$, 密度函数 $f(x) = \begin{cases} 1/\theta, & x \in [0, \theta] \\ 0, & \text{其它} \end{cases}$

分布函数 $F(x) = \int_{-\infty}^x f(t) dt = \begin{cases} 0, & x < 0 \\ x/\theta, & 0 \leq x \leq \theta \\ 1, & x > \theta \end{cases}$

所以 $X_{(n)} = \max(X_1, X_2, \dots, X_n)$ 的概率密度为

$$f_{X_{(n)}}(x) = nF^{n-1}(x) \cdot f(x) = \begin{cases} \frac{nx^{n-1}}{\theta^n}, & 0 \leq x \leq \theta \\ 0, & \text{其它.} \end{cases}$$

$$E(X_{(n)}) = \int_{-\infty}^{+\infty} x f_{X_{(n)}}(x) dx = \int_0^{\theta} x \cdot \frac{nx^{n-1}}{\theta^n} dx = \frac{n}{n+1} \theta,$$

所以 $E(\frac{n+1}{n} X_{(n)}) = \theta$, 即 $\hat{\theta}_2$ 也是 θ 的无偏估计量.

(2)问: $\hat{\theta}_1 = 2\bar{X}$ 和 $\hat{\theta}_2 = \frac{n+1}{n} X_{(n)}$ 哪一个更有效?

由于

$$D(\hat{\theta}_1) = 4D(\bar{X}) = \frac{4}{n} D(X) = \frac{\theta^2}{3n},$$

$$D(\hat{\theta}_2) = D(\frac{n+1}{n} X_{(n)}) = (\frac{n+1}{n})^2 D(X_{(n)}).$$

$$\text{又因为 } E(X_{(n)}) = \frac{n+1}{n} \theta,$$

$$E(X_{(n)}^2) = \int_{-\infty}^{+\infty} x^2 f_{X_{(n)}}(x) dx = \int_0^\theta \frac{n}{\theta^n} x^{n+1} dx = \frac{n}{n+2} \theta^2,$$

$$D(X_{(n)}) = E(X_{(n)}^2) - [E(X_{(n)})]^2 = \frac{n}{(n+1)^2(n+2)} \theta^2,$$

$$D(\hat{\theta}_2) = D\left(\frac{n+1}{n} X_{(n)}\right) = \left(\frac{n+1}{n}\right)^2 D(X_{(n)}) = \frac{1}{n(n+2)} \theta^2.$$

又 $n \geq 2$, 所以 $D(\hat{\theta}_2) < D(\hat{\theta}_1)$, $\hat{\theta}_2$ 较 $\hat{\theta}_1$ 有效.

$$D(\hat{\theta}_1) = \frac{\theta^2}{3n}$$

练习题

设 X_1, X_2, \dots, X_n 是来自均匀总体 $X \sim U(\theta - \frac{1}{2}, \theta + \frac{1}{2})$ 的一组样本.

试证明 $\hat{\theta}_1 = \bar{X}$, $\hat{\theta}_2 = \frac{1}{2}(X_{(1)} + X_{(n)})$ 都是 θ 的无偏估计量. 并指出哪一个更有效?

答案 $f_{X_{(1)}}(x) = n[1 - F(x)]^{n-1} f(x)$

$$= \begin{cases} n(\theta + 0.5 - x)^{n-1}, & \theta - 0.5 \leq x \leq \theta + 0.5 \\ 0, & \text{其他} \end{cases}$$

$$f_{X_{(n)}}(x) = n[F(x)]^{n-1} f(x)$$

$$= \begin{cases} n(x - \theta + 0.5)^{n-1}, & \theta - 0.5 \leq x \leq \theta + 0.5 \\ 0, & \text{其他} \end{cases}$$

$$E(X_{(1)}) = \theta - 0.5 + 1/(n+1), \quad E(X_{(n)}) = \theta + 0.5 - \frac{1}{n+1}$$

$$f_{X_{(1)}, X_{(n)}}(x, y) = n(n-1)[F(y) - F(x)]^{n-2} f(x)f(y)$$

$$= \begin{cases} n(n-1)(y-x)^{n-2}, & \theta - 0.5 \leq x \leq y \leq \theta + 0.5 \\ 0, & \text{其他} \end{cases}$$

答案(续)

$$E(X_{(1)} + X_{(n)})^2 = 4\theta^2 + \frac{2}{(n+1)(n+2)},$$

$$D(\hat{\theta}_1) = \frac{1}{12n}, \quad D(\hat{\theta}_2) = \frac{1}{(n+1)(n+2)}.$$

$n > 2$ 时

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

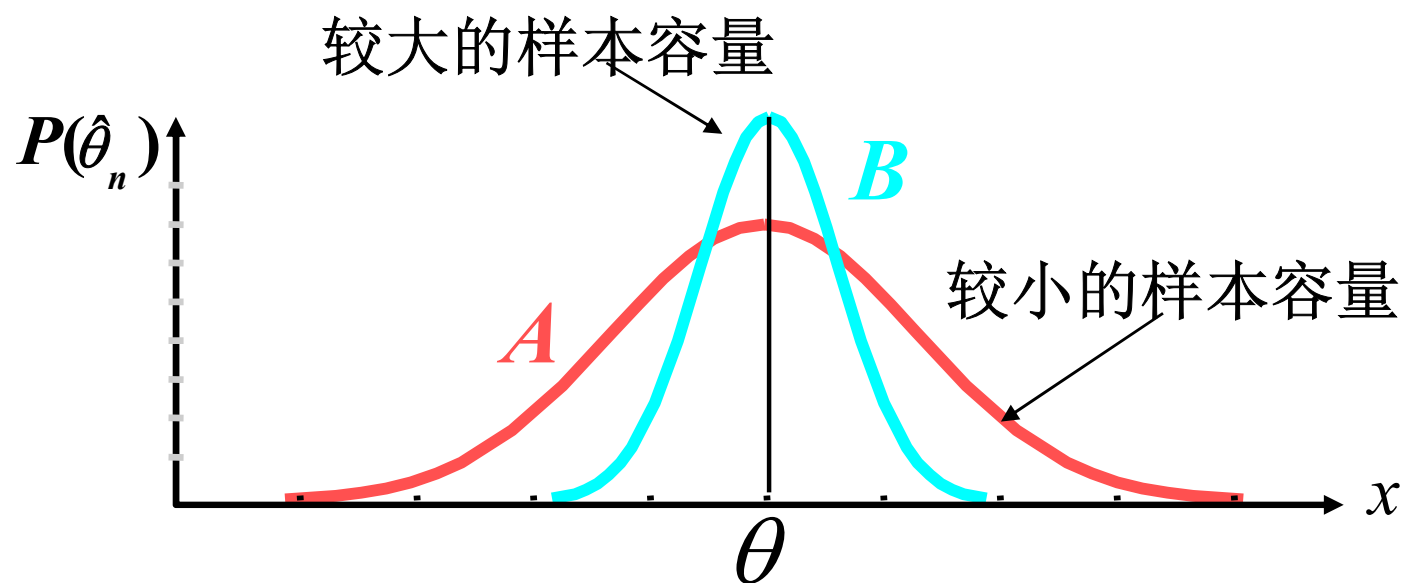
三、相合性

若 $\hat{\theta}_n = \hat{\theta}_n(X_1, X_2, \dots, X_n)$ 为参数 θ 的估计序列, 如果 $\hat{\theta}_n$ 依概率收敛于 θ , 即对任意 $\varepsilon > 0$, 有

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

或
$$\lim_{n \rightarrow \infty} P\{|\hat{\theta}_n - \theta| \geq \varepsilon\} = 0$$

则称 $\hat{\theta}_n$ 为 θ 的相合估计量(或一致估计量).



定理2 设 $\hat{\theta}_n$ 是 θ 的一个估计量,若 $\lim_{n \rightarrow \infty} E(\hat{\theta}_n) = \theta$,
且 $\lim_{n \rightarrow \infty} D(\hat{\theta}_n) = 0$, 则 $\hat{\theta}_n$ 是 θ 的相合估计(或一致估计).

证明

$$\begin{aligned} 0 &\leq P\{|\hat{\theta}_n - \theta| \geq \varepsilon\} \leq \frac{1}{\varepsilon^2} E(\hat{\theta}_n - \theta)^2 \\ &= \frac{1}{\varepsilon^2} E\left[(\hat{\theta}_n - E\hat{\theta}_n) + (E\hat{\theta}_n - \theta)\right]^2 \\ &= \frac{1}{\varepsilon^2} [D(\hat{\theta}_n) + (E\hat{\theta}_n - \theta)^2] \end{aligned}$$

令 $n \rightarrow \infty$, 由定理的假设得

$$\lim_{n \rightarrow \infty} P\{|\hat{\theta}_n - \theta| \geq \varepsilon\} = 0$$

即 $\hat{\theta}_n$ 是 θ 的相合估计.

例如：样本 $k(k \geq 1)$ 阶矩是总体 X 的 k 阶矩 $E(X^k)$ 的相合估计量(辛钦大数定律). 因此，矩估计一般是相合估计.

例6 试证: (1)样本均值 \bar{X} 是总体均值 μ 的相合估计量; (2) 样本方差 $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ 及

$S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ 都是总体方差 σ^2 的相合估计量.

证明 (1) 由大数定律知, $\forall \varepsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{1}{n} \sum_{i=1}^n X_i - \mu \right| < \varepsilon \right\} = 1$$

所以 $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ 是 μ 的相合估计量.

$$\begin{aligned} (2) \text{ 又 } S_n^2 &= \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2 \\ &= A_2 - \bar{X}^2, \quad (A_2 \text{ 是样本二阶原点矩}) \end{aligned}$$

由大数定律知

$$A_2 = \frac{1}{n} \sum_{i=1}^n X_i^2 \text{ 依概率收敛于 } E(X^2),$$

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \text{ 依概率收敛于 } E(X),$$

故 $S_n^2 = A_2 - \bar{X}^2$ 依概率收敛于 $E(X^2) - [E(X)]^2 = \sigma^2$,

所以 S_n^2 是 σ^2 的相合估计量.

$$S^2 = \frac{n}{n-1} S_n^2 \text{ 也是 } \sigma^2 \text{ 的相合估计量.}$$

练习题

设 X_1, X_2, \dots, X_n 独立同分布,

$$EX_1 = \mu, \quad DX_1 = \sigma^2 < +\infty,$$

试证 $\hat{\mu} = \frac{2}{n(n+1)} \sum_{i=1}^n iX_i$ 是 μ 的相合估计.

四、小结

估计量的评价标准

$$\left\{ \begin{array}{ll} \text{无偏性} & \text{有效性} \\ E(\hat{\theta}) = \theta & D(\hat{\theta}) \text{小者} \\ \text{相合性} & \text{渐进无偏性} \\ \hat{\theta}_n \xrightarrow{P} \theta \Leftarrow \lim_{n \rightarrow \infty} E(\hat{\theta}_n) = \theta, \lim_{n \rightarrow \infty} D(\hat{\theta}_n) = 0. \end{array} \right.$$

相合性是对估计量的一个基本要求，不具备相合性的估计量是不予以考虑的。