§2 估计量的评选标准

对于同一个未知参数θ，可构造不同的估计量$\hat{\theta}$。究竟采用哪一个为好呢？这就产生了怎样去选择一个优良的估计量问题。

例如，为了估计总体均值$E(X)=\theta$，可以采用下面三个统计量中的任何一个作为θ的估计量。

$$\hat{\theta}_1(X_1,\ldots,X_n)=X_1,$$

$$\hat{\theta}_2(X_1,\ldots,X_n)=\frac{\max\limits_{1\le i\le n}\{X_i\}+\min\limits_{1\le i\le n}\{X_i\}}{2},$$

$$\hat{\theta}_3(X_1,\ldots,X_n)=\frac{1}{n}\sum_{i=1}^{n}X_i\,。$$

对于未知参数θ，由于其估计量θ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.

设θ为θ的估计量，则其误差为$\theta-\theta$．$\theta-\theta$的值由样本$(X_1,X_2,\cdots,X_n)$确定，因此$\theta-\theta$是随机变量．常选用平均绝对误差$E\left|\theta-\theta\right|$或均方误差$E[(\theta-\theta)^2]$，来衡量用$\theta$来估计$\theta$的估计效果．但由于绝对值函数$|\cdot|$的运算性质较弱，故一般采用均方误差$E[(\theta-\theta)^2]$更多一些.

定义1 设$\hat{\theta}$为θ的估计量，如果$E(\hat{\theta}-\theta)^2$存在，就称$E(\hat{\theta}-\theta)^2$为用$\hat{\theta}$估计θ时所产生的均方误差.

当$E(\hat{\theta}-\theta)^2$越小时, 表明在均方误差意义下, 用$\hat{\theta}$估计θ的效果越好.

由于 $E(\hat{\theta}-\theta)^2 = E[(\hat{\theta}-E\hat{\theta})-(\theta-E\hat{\theta})]^2$

$$= E[(\hat{\theta}-E\hat{\theta})^2-2(\hat{\theta}-E\hat{\theta})(\theta-E\hat{\theta})+(\theta-E\hat{\theta})^2]$$

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

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

一、无偏性

若对任何$\theta\in\Theta$有

$$E_\theta \hat{g}(X_1,\dots,X_n)=g(\theta)$$

则称$\hat{g}(X_1,\dots,X_n)$为$g(\theta)$的**无偏估计量**，否则称为有偏估计量。

定理 1 如果$\hat{\theta}_1,\hat{\theta}_2,\dots,\hat{\theta}_m$都是$\theta$的无偏估计量，则当$\sum_{i=1}^{m}C_i=1$时，

$\sum_{i=1}^{m}C_i\hat{\theta}_i$也是$\theta$的无偏估计量，即

$$E\left(\sum_{i=1}^{m}C_i\hat{\theta}_i\right)=\theta。$$

推论 样本均值$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$是总体均值$E(X)=\mu$的无偏估计量，

即$E(\overline{X})=\mu$。

例 1　试证样本方差 $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$ 是总体方差 $D(X)=\sigma^2$ 的无偏估计量。

$$\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)\ \Rightarrow\ E\left(\frac{(n-1)S^2}{\sigma^2}\right)=n-1$$

$$\Rightarrow\ E\left(S^2\right)=\sigma^2$$

?

例 2 设总体 $X \sim N(\mu,\sigma^2)$，试证 $\overline{X}$ 是 μ 的无偏估计量，但 $\overline{X}^2$ 不是 μ^2 的无偏估计量。

$$\bar{X} \sim \mathrm{N}(\mu,\frac{\sigma^2}{n}),\quad E(\bar{X})=\mu$$

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

如果 $\hat{\theta}$ 是参数 θ 的无偏估计量，除了 f 是线性函数外，一般并不能推出 $\hat{\theta}$ 的函数 $f(\hat{\theta})$ 也是 $f(\theta)$ 的无偏估计量。

二、有效性

设 $\hat{g}_1=\hat{g}_1(X_1,\ldots,X_n)$，$\hat{g}_2=\hat{g}_2(X_1,\ldots,X_n)$ 都是 $g(\theta)$ 的无偏估计量。如果

$$D_\theta(\hat{g}_1)\le D_\theta(\hat{g}_2)，\quad 对一切\ \theta\in\Theta，$$

则称 $\hat{g}_1$ 比 $\hat{g}_2$ 有效。

如果无偏估计量 $\hat{g}_0=\hat{g}_0(X_1,\ldots,X_n)$ 使不等式

$$D_\theta(\hat{g}_0)\le D_\theta(\hat{g})，$$

对一切 $\theta\in\Theta$ 和一切无偏估计量 $\hat{g}=\hat{g}(X_1,\ldots,X_n)$ 都成立，则称 $\hat{g}_0$ 是 $g(\theta)$ 的

一致最小方差无偏估计量。

定理 2 设参数空间Θ是实数轴上一个开区间，总体X的概率函数或密度函数$f(x;\theta)$存在且满足一定的正则条件，则对θ的任一无偏估计量$\hat{\theta}$，必有

$$D_\theta(\hat{\theta}) \ge \frac{1}{nI(\theta)}，对一切\theta \in \Theta，$$

其中

$$I(\theta) = E_\theta\left(\frac{\partial \ln f(x;\theta)}{\partial \theta}\right)^2。$$

注：一般形式见教材定理2.2

例 3 设总体 $X \sim B(1,p)$，试验证参数 P 的无偏估计 $\hat{p}=\overline{X}=\sum_{i=1}^{n}\frac{X_i}{n}$ 是有效估计。

解：分布$B(1,p)$的概率函数为 $f(x;p)=p^x(1-p)^{1-x}, \qquad (x=0,1)$

所以

$$\frac{\partial}{\partial p}\ln f(X;p)=\frac{\partial}{\partial p}[X\ln p+(1-X)\ln(1-p)]=\frac{X}{p}-\frac{1-X}{1-p}=\frac{X-p}{p(1-p)}$$

$$I(p)=E_p\left[\frac{\partial \ln f(X;p)}{\partial p}\right]^2=E_p\left[\frac{X-p}{p(1-p)}\right]^2=\frac{E_p[(X-p)]^2}{[p(1-p)]^2}$$

$$=\frac{D_p(X)}{[p(1-p)]^2}=\frac{p(1-p)}{[p(1-p)]^2}=\frac{1}{p(1-p)}$$

于是，参数p的无偏估计量的方差下界是 $\frac{1}{nI(p)}=\frac{p(1-p)}{n}$

又无偏估计量$\hat{p}=\bar{X}$的方差$D_p(\hat{p})=D_p(\bar{X})=\frac{p(1-p)}{n}$

达到了方差下界，所以$\hat{p}=\bar{X}$是p的有效估计。

例 4 设总体 $X \sim N(\mu, \sigma^2)$，试证 μ 的无偏估计量 $\hat{\mu} = \overline{X}$ 是有效估计；而 σ^2 的无偏估计量 $\hat{\sigma}^2 = S^2$ 不是有效估计。

P55

不过，可以证明 S^2 是 σ^2 的一致最小方差无偏估计。这说明 $C-R$ 不等式中的方差下界不一定能达到。

例 5 设 $X_1,\ldots,X_n$ 是来自任一总体 X 的样本，则样本均值 $\overline{X}$ 是总体均值 $E(X)=\mu$ 的一致最小方差线性无偏估计量。

解：设 $\hat{g}=\sum_{i=1}^{n}C_iX_i$，由 $E_\mu(\hat{g})=\mu$ 知 $\sum_{i=1}^{n}C_i=1$。

由 $Cauchy-Schwarz$ 不等式可得

$$\left(\sum_{i=1}^{n}C_i\right)^2\le\left(\sum_{i=1}^{n}1^2\right)\left(\sum_{i=1}^{n}C_i^2\right)=n\sum_{i=1}^{n}C_i^2,$$

即

$$\sum_{i=1}^{n}C_i^2\ge\frac{1}{n}\left(\sum_{i=1}^{n}C_i\right)^2=\frac{1}{n}。$$

于是有 $D_\mu(\hat{g})=\left(\sum_{i=1}^{n}C_i^2\right)D_\mu(X)\ge\frac{1}{n}D_\mu(X)=D_\mu(\overline{X})$，对一切 $\mu\in\Theta$，

所以 $\overline{X}$ 是 μ 的一致最小方差线性无偏估计量。