求导并设导数为零可得

$$\frac{\ln L(\theta,\mu)}{d\theta} = \frac{n}{\theta} - \sum_{i=1}^{n}(X_i-\mu) = 0 \quad \Rightarrow \quad \theta = \frac{1}{\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)},$$

另一方程为

$$\frac{\ln L(\theta,\mu)}{d\mu} = \theta\sum_{i=1}^{n}X_i = 0 \quad \Rightarrow \quad \theta = 0,$$

此时无解求解$\theta,\mu$的极大似然估计. 回到似然函数的定义

$$L(\theta,\mu) = \begin{cases} \theta^n e^{-\theta\sum_{i=1}^{n}(X_i-\mu)} & X_1,X_2,\cdots,X_n \geq \mu \\ 0 & 其它 \end{cases}$$

可以发现$\mu$越大, 似然函数$L(\theta,\mu)$越大, 但须满足$X_i \leq \mu$ $(i\in[n])$. 由此可得极大似然估计

$$\hat{\mu} = \min\{X_1,X_2,\cdots,X_n\},$$

以及

$$\hat{\theta} = \frac{1}{\frac{1}{n}\sum_{i=1}^{n}(X_i-\hat{\mu})}.$$

□

## 9.2 估计量的评价标准

前一节已经讲过不同的点估计方法, 不同的估计方法可能得到不同的估计值, 自然涉及到一个问题: 采用哪一种估计量更好, 或更好的标准是什么呢?

估计量的常用标准: 无偏性, 有效性, 一致性.

### 9.2.1 无偏性

**定义9.1.** $X_1,X_2,\cdots,X_n$是总体$X$的一样本, $\hat{\theta}=\hat{\theta}(X_1,X_2,\cdots,X_n)$是$\theta$的一个估计量, 若

$$E_{X_1,X_2,\ldots,X_n}[\hat{\theta}] = E[\hat{\theta}(X_1,X_2,\cdots,X_n)] = \theta,$$

称$\hat{\theta}$为$\theta$的无偏估计.

无偏估计不要求每次的估计值$\hat{\theta}$都与$\theta$相等, 但在期望下$E(\hat{\theta})=\theta$. 其意义在于无系统性偏差, 无偏性是一种对估计量常见而且重要的标准. 根据以前所学统计量的知识有

**例9.8** (即样本$k$阶矩为总体$k$阶矩的无偏估计)**.** $X_1,X_2,\cdots,X_n$是总体$X$的一个样本, 若$E[X^k]$存在, 则$A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k$是$a_k=E[X^k]$的无偏估计.

注意$\hat{\theta}$是$\theta$的无偏估计, 但并不一定有$g(\hat{\theta})$是$g(\theta)$的无偏估计, 如下例所示:

**例9.9.** 总体$X$的期望为$\mu$, 方差为$\sigma^2$, $X$的样本为$X_1, X_2, \cdots, X_n$, 则: 1) $S_0^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2$是$\sigma^2$的有偏估计; 2) $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$是$\sigma^2$的无偏估计.

**例9.10.** $X_1, X_2, \cdots, X_n$是总体$X$的一样本, $X$的概率密度

$$f(x;\theta) = \begin{cases} \frac{1}{\theta}e^{-\frac{x}{\theta}} & x \geq 0 \\ 0 & x < 0 \end{cases},$$

求证: $\bar{X}$和$n\min\{X_1, X_2, \cdots, X_n\}$是$\theta$的无偏估计.

*Proof.* 容易知道$\bar{X}$是$E[X]$的无偏估计. 设$Z = \min\{X_1, X_2, \cdots, X_n\}$, 则有

$$\begin{aligned} F_Z(z) &= \Pr[Z \leq z] = 1 - \Pr[Z > z] \\ &= 1 - \Pr[X_1 > z]\Pr[X_2 > z]\cdots\Pr[X_n > z] \\ &= 1 - \prod_{i=1}^n (1 - \Pr[X_i \leq z]) = \begin{cases} 0 & z < 0 \\ 1 - e^{-\frac{nz}{\theta}} & z \geq 0 \end{cases} \end{aligned}$$

于是得到$Z$的密度函数为

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

进一步有

$$E[Z] = \int_0^{+\infty} \Pr[Z > z]dz = \int_0^{+\infty} e^{-\frac{nz}{\theta}}dz = \frac{\theta}{n}.$$

所以$\theta = E[nZ]$, 从而完成证明. □

### 9.2.2 有效性

一个参数可能存在多个无偏估计, 若$\hat{\theta}_1, \hat{\theta}_2$都是$\theta$的无偏估计, 此时可以比较方差

$$\mathrm{Var}(\hat{\theta}_1) = E[(\hat{\theta}_1 - \theta)^2], \qquad \mathrm{Var}(\hat{\theta}_2) = E[(\hat{\theta}_2 - \theta)^2],$$

方差越小越好.

**定义9.2** (有效性)**.** 设$\hat{\theta}_1 = \hat{\theta}_1(X_1, X_2, \cdots, X_n)$和$\hat{\theta}_2 = \hat{\theta}_2(X_1, X_2, \cdots, X_n)$是$\theta$的无偏估计, 如果有$Var(\hat{\theta}_1) \leq Var(\hat{\theta}_2)$成立, 则称$\theta_1$比$\theta_2$有效.

**例9.11.** $X_1, X_2, \cdots, X_n$是总体$X$的一样本, $X$的概率密度

$$f(x;\theta) = \begin{cases} \frac{1}{\theta}e^{-\frac{x}{\theta}} & x \geq 0 \\ 0 & x < 0 \end{cases},$$

令$Z = \min\{X_1, X_2, \cdots, X_n\}$, 求证: 当$n > 1$时, $\bar{X}$较$nZ$有效.

*Proof.* 根据独立性有$Var(\bar{X})=\frac{1}{n^2}\sum_{i=1}^n Var(X_i)=\frac{\theta^2}{n}$. 由上例可知$Z$的概率密度为

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

从而得到$Var(Z)=\frac{\theta^2}{n^2}$以及$Var(nZ)=\theta^2$. 当$n\geq 1$ 时, $Var(nZ)\geq Var(\bar{X})$, 故$\bar{X}$比$nZ$更有效. □

**例9.12.** $X_1,X_2,\cdots,X_n$是总体$X$的一个样本, 且$E(X)=\mu$, $Var(X)=\sigma^2$. 设常数$c_1,c_2,\cdots,c_n\geq 0$且$\sum_{i=1}^n c_i=1, c_i\neq 1/n$, 求证: $\bar{X}$比$\sum_{i=1}^n c_iX_i$有效.

*Proof.* 根据独立同分布性质有$E[\bar{X}]=\mu$以及$Var(\bar{X})=\sigma^2 n$, 根据期望线性性有$E[\sum_{i=1}^n c_iX_i]=\mu$. 进一步有

$$Var(\sum_{i=1}^n c_iX_i)=\sum_{i=1}^n c_i^2 Var(X_i)=\sigma^2\sum_{i=1}^n c_i^2\geq\frac{\sigma^2}{n}$$

根据不等式$\sum_{i=1}^n c_i^2/n\geq(\frac{c_1+c_2+\cdots+c_n}{n})^2$, 所以$\mathrm{Var}(\sum_{i=1}^n c_iX_i)\geq\mathrm{Var}(\bar{X})$. □

**定义9.3** (Rao-Crammer不等式)**.** 设随机变量的概率密度$f(x;\theta)$或者分布函数$F(x;\theta)$, 令

$$Var_0(\theta)=\frac{1}{nE[\frac{\partial\ln f(X;\theta)}{\partial\theta}]^2}\ \text{或}\ Var_0(\theta)=\frac{1}{nE[\frac{\partial\ln F(X;\theta)}{\partial\theta}]^2},$$

对任意的无偏估计$\hat{\theta}$有

$$Var(\hat{\theta})\geq Var_0(\theta),$$

称$Var_0(\theta)$为方差的下界. 当$Var(\hat{\theta})=Var_0(\theta)$时, 称$\hat{\theta}$为达到方差下界的无偏估计量, 此时$\hat{\theta}$为最有效估计量, 简称有效估计量.

**例9.13.** 设总体$X$的密度函数为$f(x;\theta)=\begin{cases}\frac{1}{\theta}e^{-\frac{x}{\theta}} & x>0\\ 0 & x\leq 0\end{cases}$, $X_1,X_2,\cdots,X_n$为$X$的一样本值, 求$\theta$的极大似然估计.

解. 首先得到对数似然函数

$$\ln L(\theta)=-n\ln\theta-\frac{\sum_{i=1}^n X_i}{\theta}\quad\Rightarrow\quad\hat{\theta}=\frac{1}{n}\sum_{i=1}^n X_i,$$

进一步得到统计量的方差

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

同时考察

$$\ln f(X;\theta)=-\ln\theta-\frac{X}{\theta},\quad\frac{\partial}{\partial\theta}\ln f(X;\theta)=-\frac{1}{\theta}+\frac{X^2}{\theta^2}$$

所以

$$E[\frac{\partial}{\partial\theta}\ln f(X;\theta)]^2=E[(-\frac{1}{\theta}+\frac{X^2}{\theta^2})^2]=\frac{1}{\theta^4}E[(X-E(X))^2]=\frac{1}{\theta^2},$$

从而得到$\mathrm{Var}_0(X)=1/n\theta^2=\mathrm{Var}(\hat{\theta})$, $\bar{X}$为达到方差下界的无偏估计量. □

### 9.2.3 一致性

**定义9.4** (一致性)**.** 设$\theta \in \Theta$为未知参数, 若$\hat{\theta}_n = \hat{\theta}_n(X_1, X_2, \cdots, X_n)$是$\theta$的一个估计量, 当$n \to \infty$时, 有

$$\hat{\theta}_n \xrightarrow{P} \theta$$

即对任意$\epsilon > 0$, 有$\lim_{n\to 0} \Pr[|\hat{\theta}_n - \theta| > \epsilon] = 0$成立, 则称$\hat{\theta}_n$为$\theta$的一致性估计量.

一致性估计量$\hat{\theta}$在足够多样本情形下有效逼近真实$\theta$, 是对估计的一个基本要求. 不满足一致性的估计量一般不予考虑.

**定理9.1.** 设$\hat{\theta}_n = \hat{\theta}_n(X_1, X_2, \cdots, X_n)$是$\theta$的一估计量, 若$\lim_{n\to\infty} E[\hat{\theta}_n] = \theta$, $\lim_{n\to\infty} Var(\hat{\theta}_n) = 0$, 则$\hat{\theta}_n$为$\theta$的一致性估计.

**定理9.2.** 设$\hat{\theta}_{n_1}, \hat{\theta}_{n_2}, \cdots, \hat{\theta}_{n_k}$为$\theta_1, \theta_2, \cdots, \theta_k$的一致性估计: $g : \mathbb{R}^n \to \mathbb{R}$的连续函数, 则函数$\widehat{\eta}_n = g(\hat{\theta}_{n_1}, \hat{\theta}_{n_2}, \cdots, \hat{\theta}_{n_k})$ 是$\eta = g(\theta_1, \theta_2, \cdots, \theta_k)$ 的一致性估计.

由大数定理: 样本$k$阶矩是总体$k$阶矩的一致性估计. 矩估计得到的估计量一般为一致性估计量. 极大似然估计量在一定条件下为一致性估计量.

**例9.14.** 总体$X$的密度函数为$f(x;\theta) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}} & x > 0 \\ 0 & x < 0 \end{cases}$, 则$\bar{X}$为$\theta$的无偏, 有效, 一致性估计量.

解.

$$\lim_{n\to\infty} Var(\bar{X}) = \lim_{n\to\infty} \frac{\theta^2}{n} = 0.$$

□

**例9.15.** 设$X_1, X_2, \cdots, X_n$是来自总体$X \sim U(0, \theta)$的一个样本, 求$\theta$的极大似然估计量是一致性估计量.

*Proof.* 首先有$\theta$的极大似然估计为$\hat{\theta} = \max(X_1, X_2, \cdots, X_n)$. 令$Z = \max(X_1, X_2, \cdots, X_n)$, 则$Z$的分布函数

$$F_Z(z) = \Pr[Z \le z] = \Pr[\max(X_1, X_2, \cdots, X_n) \le z] = \prod_{i=1}^{n} \Pr[X_i \le z] = (\frac{z}{\theta})^n.$$

由此得到密度函数$f_Z(z) = n\frac{z^{n-1}}{\theta^n}$ $(z \in [0, \theta])$. 可以发现

$$E[\hat{\theta}] = E[Z] = \int_0^\theta \frac{nz^n}{\theta^n} dz = \frac{n}{n+1}\theta,$$

因此$\hat{\theta}$是$\theta$的有偏估计. 另一方面有

$$E[Z^2] = \int_0^\theta \frac{nz^{n+1}}{\theta^n} dz = \frac{n}{n+2}\theta^2.$$

从而得到

$$Var(Z) = E[Z^2] - (E[Z])^2 = \frac{n}{n+2}\theta^2 - (\frac{n\theta}{n+1})^2 = \frac{n}{(n+1)^2(n+2)}\theta^2 \longrightarrow 0 \qquad n \to +\infty.$$

由此可得$\hat{\theta}$ 是$\theta$的有偏, 但一致性估计量. □