《概率论与数理统计》习题

第十九讲　点估计的评价方法 1

1. 设 $\hat{\theta}$ 是参数 θ 的无偏估计，且有 $\mathrm{Var}(\hat{\theta}) > 0$，试证 $(\hat{\theta})^2$ 不是 θ^2 的无偏估计。

   证明：

   由方差的定义可知：

$$Var(\hat{\theta}) = E(\hat{\theta}^2) - E^2(\hat{\theta}) > 0.$$

   且 $\hat{\theta}$ 是参数 θ 的无偏估计，即 $E(\hat{\theta}) = \theta$. 可得：

$$E(\hat{\theta}^2) = E^2(\hat{\theta}) + Var(\hat{\theta}) = Var(\hat{\theta}) + \theta^2 > \theta^2.$$

   所以，$(\hat{\theta})^2$ 不是 θ^2 的无偏估计。 ■

2. 设 $x_1, x_2, \cdots, x_n$ 是来自于下列总体的简单样本

$$p(x;\theta) = \begin{cases} 1, \theta - \dfrac{1}{2} \le x \le \theta + \dfrac{1}{2}, \\ 0, \text{其他}, \end{cases} , -\infty < \theta < \infty$$

   证明样本均值 $\bar{x}$ 及 $\frac{1}{2}(x_{(1)}) + x_{(n)})$ 都是 θ 的无偏估计，问哪一个更为有效？

   解：

   由题意可知：$X \sim U(\theta - \frac{1}{2}, \theta + \frac{1}{2})$，$E(X) = \theta$，$Var(X) = \frac{1}{12}$.

   可得 $E(\bar{x}) = \theta$，所以样本均值 $\bar{x}$ 是 θ 的无偏估计，且 $Var(\bar{x}) = \frac{1}{12n}$.

   接下来计算 $\frac{1}{2}(x_{(1)} + x_{(n)})$ 的均值和方差：

   令 $Y = X - (\theta - \frac{1}{2}) \sim U(0, 1)$，

   即：$y_i = x_i - (\theta - \frac{1}{2})$，$i = 1, 2, \ldots, n$.

   则：

$$\frac{1}{2}(x_{(1)} + x_{(n)}) = \frac{1}{2}(y_{(1)} + y_{(n)}) + \theta - \frac{1}{2}$$

   又因为 $y_{(i)} \sim Be(i, n-i+1)$，所以，$E(y_{(1)}) = \frac{1}{n+1}$，$E(y_{(n)}) = \frac{n}{n+1}$.

可得:

$$E(\frac{1}{2}(x_{(1)}+x_{(n)}))=\frac{1}{2}E(y_{(1)}+y_{(n)})+\theta-\frac{1}{2}=\theta.$$

所以 $\frac{1}{2}(x_{(1)}+x_{(n)})$ 是 θ 的无偏估计。

又因为 $y_{(n)}-y_{(1)}\sim Be(n-1,2)$，可得:

$$Var(y_{(1)})=Var(y_{(n)})=\frac{n}{(n+1)^2(n+2)}.$$

$$\begin{aligned}Var(y_{(n)}-y_{(1)})&=\frac{2(n-1)}{(n+1)^2(n+2)}.\\Cov(y_{(1)},y_{(n)})&=\frac{1}{2}\Big(Var(y_{(1)})+Var(y_{(n)})-Var(y_{(n)}-y_{(1)})\Big)\\&=\frac{1}{(n+1)^2(n+2)}.\end{aligned}$$

于是,

$$\begin{aligned}Var\left(\frac{1}{2}(x_{(1)}+x_{(n)})\right)&=\frac{1}{4}\Big(Var(y_{(1)})+Var(y_{(n)})+2Cov(y_{(1)},y_{(n)})\Big)\\&=\frac{1}{4}\left(\frac{2n}{(n+1)^2(n+2)}+\frac{2}{(n+1)^2(n+2)}\right)\\&=\frac{1}{2(n+1)(n+2)}.\end{aligned}$$

当 $n>2$ 时，有 $\frac{1}{12n}>\frac{1}{2(n+1)(n+2)}$. 说明在此情况下，$\frac{1}{2}(x_{(1)}+x_{(n)})$ 比 $\bar{x}$ 更有效。

3. 设从均值为 μ, 方差为 $\sigma^2>0$ 的总体中分布抽取容量为 n_1 和 n_2 的两独立样本, $\bar{x}_1$ 和 $\bar{x}_2$ 是这两个样本的均值。试证，对于任意常数 $a,b(a+b=1)$, $Y=a\bar{x}_1+b\bar{x}_2$ 都是 μ 的无偏估计，并确定常数 a,b 使得 $\mathrm{Var}(Y)$ 达到最小。

证明:

由于 $\bar{x_1}$ 和 $\bar{x_2}$ 分别为样本量 n_1 和 n_2 对应的均值,

则: $E(\bar{x_1})=\mu$, $E(\bar{x_2})=\mu$, $Var(\bar{x_1})=\frac{\sigma^2}{n_1}$, $Var(\bar{x_2})=\frac{\sigma^2}{n_2}$.

所以,

$$\begin{aligned}E(Y)&=E(a\bar{x_1}+b\bar{x_2})\\&=aE(\bar{x_1})+bE(\bar{x_2})\\&=(a+b)\mu\\&=\mu.\end{aligned}$$

即 Y 是 μ 的无偏估计。

由于 a+b = 1，Y 又可写成：$Y=a\bar{x_1}+(1-a)\bar{x_2}$.

$$\begin{aligned}Var(Y)&=\frac{a^2\sigma^2}{n_1}+\frac{(1-a)^2\sigma^2}{n_2}\\&=\sigma^2\left((\frac{1}{n_1}+\frac{1}{n_2})a^2+\frac{1-2a}{n_2}\right).\end{aligned}$$

对其求导，当 $a=\frac{n_1}{n_1+n_2}$ 时，$Var(Y)$ 可达到最小值，此时 $b=\frac{n_2}{n_1+n_2}$.

■

第二十讲　点估计的评价方法 2

1. 设总体密度函数为 $p(x;\theta)=\theta x^{\theta-1}, 0<x<1, \theta>0, x_1,x_2,\cdots,x_n$ 是样本. 求 $g(\theta)=1/\theta$ 的最大似然估计.

   解:

$$L(\theta)=\prod_{i=1}^{n}\theta x_i^{\theta-1}.$$

$$\begin{aligned}lnL(\theta)&=nln\theta+(\theta-1)\sum_{i=1}^{n}lnx_i\\&=-nlng(\theta)+\left(\frac{1}{g(\theta)}-1\right)\sum_{i=1}^{n}lnx_i.\end{aligned}$$

$$\frac{\partial lnL(\theta)}{\partial g(\theta)}=-\frac{\theta}{g(\theta)}-\frac{1}{g^2(\theta)}\sum_{i=1}^{n}lnx_i=0.$$

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

   (此题也可以先求关于 θ 的 MLE，再代入到 g(θ) 函数里)

2. 设总体 $X\sim Exp(1/\theta), x_1,x_2,\cdots,x_n$ 是样本。

(a) 验证 θ 的矩估计和最大似然估计都是 $\bar{x}$;

(b) 验证 $\bar{x}$ 也是 θ 的相合估计和无偏估计;

(c) 试证明在均方误差准则下存在优于 $\bar{x}$ 的估计。(提示:考虑 $\hat{\theta}_n = a\bar{x}$,找均方误差最小者)

解:

(a)

因为 $X \sim Exp(1/\theta)$,$f(x) = \frac{1}{\theta}e^{-\frac{x}{\theta}}$.

$E(X) = \theta = \bar{x}$,所以 θ 的矩估计为 $\bar{x}$;

$lnL(\theta) = -nln\theta - \frac{1}{\theta}\sum_{i=1}^{n} x_i$,令 $\frac{\partial lnL(\theta)}{\partial\theta} = 0$,得 $\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$,所以 θ 的 MLE 为 $\bar{x}$。

(b)

因为 $E(\bar{x}) = \frac{1}{n}\sum_{i=1}^{n} E(x_i) = \theta$,所以 $\bar{x}$ 也是 θ 的无偏估计;

$Var(\bar{x}) = \frac{1}{n^2}\sum_{i=1}^{n} Var(x_i) = \frac{\theta^2}{n}$,当 n 趋向于正无穷时,$Var(\bar{x})$ 趋向于 0,所以 $\bar{x}$ 也是 θ 的相合估计。

(c)

令 $\hat{\theta}_a = a\bar{x}$,则:

$$\begin{aligned}
MSE(\hat{\theta}_a) &= Var(\hat{\theta}_a) + (E\hat{\theta}_a - \theta)^2 \\
&= a^2 Var(\bar{x}) + (aE(\bar{x})_a - \theta)^2 \\
&= \frac{a^2\theta^2}{n} + \theta^2(a-1)^2.
\end{aligned}$$

$MSE(\hat{\theta}_a)$ 关于 a 求导,当 $a = \frac{n}{n+1}$ 时,$MSE(\hat{\theta}_a)$ 达到最小,最小值为 $\frac{\theta^2}{n+1} < MSE(\bar{x})$.

3. 设 $x_1, x_2, \ldots, x_n$ 是来自正态总体 $N(\mu_1, \sigma^2)$ 的样本, $y_1, y_2, \ldots, y_m$ 是来自正态总体 $N(\mu_2, \sigma^2)$ 的样本.

(a) 利用样本 $x_1, x_2, \ldots, x_n$ 求 σ^2 的估计;

(b) 利用样本 $x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m$ 求 σ^2 的估计;

(c) 请从充分性原则的角度评价这两个估计.

解:

(a)

$$L(\sigma^2) = (\frac{1}{\sqrt{2\pi}\sigma})^n e^{-\left(\frac{\sum_{i=1}^{n}(x_i-\mu_1)^2}{2\sigma^2}\right)}$$

$$lnL(\sigma^2) = -\frac{n}{2}ln2\pi - \frac{n}{2}ln\sigma^2 - \frac{\sum_{i=1}^{n}(x_i-\mu_1)^2}{2\sigma^2}$$

解得 $\hat{\sigma^2} = \frac{\sum_{i=1}^{n}(x_i-\mu_1)^2}{n}$

(b)

$$L(\sigma^2) = \frac{1}{(\sqrt{2\pi})^{n+m}(\sigma^2)^{\frac{n+m}{2}}}e^{-\left(\frac{\sum_{i=1}^{n}(x_i-\mu_1)^2+\sum_{i=1}^{m}(y_i-\mu_2)^2}{2\sigma^2}\right)}$$

$$lnL(\sigma^2) = -\frac{n+m}{2}ln2\pi - \frac{n+m}{2}ln\sigma^2 - \frac{\sum_{i=1}^{n}(x_i-\mu_1)^2+\sum_{i=1}^{m}(y_i-\mu_2)^2}{2\sigma^2}$$

解得 $\hat{\sigma^2} = \frac{\sum_{i=1}^{n}(x_i-\mu_1)^2+\sum_{i=1}^{m}(y_i-\mu_2)^2}{n+m}$

(c)

(b) 所得估计为充分统计量的函数, 而 (a) 所得估计不是充分统计量的函数。故从充分性原则角度评价，(b) 所得估计更优.