习题(34)

34. 1 设$X_1,X_2,\cdots,X_n$是来自总体$N(\mu,\sigma^2)$的样本,其中μ,σ^2均未知,试确定常数k,使统计量$\frac{1}{k}\cdot\sum_{i=1}^{n-1}(X_{i+1}-X_i)^2$是$\sigma^2$的无偏估计.

34. 2 设总体X的概率密度函数

$$f(x;\theta)=\begin{cases}\frac{x}{\theta^2}\cdot\exp\{-\frac{x}{\theta}\} & , \quad x>0\\ 0 & , \quad x\le 0\end{cases},$$

其中$\theta>0$未知,$X_1,X_2,\cdots,X_n$为X的样本.

1) 试求θ的极大似然估计量$\hat{\theta}$;

2) 问$\hat{\theta}$是否为θ的无偏估计? 证明你的结论.

34. 3 设总体$X\sim U(0,\theta)$,其中$\theta>0$为未知参数,$X_1,X_2,\cdots,X_n$为X的样本.

1) 求θ的极大似然估计量$\hat{\theta}$;

2) 证明$\hat{\theta}$不是θ的无偏估计量;

3) 确定常数c,使得统计量$c\cdot\hat{\theta}$为θ的无偏估计;

34. 4 设二维总体(X,Y),样本为(X_1,Y_1),(X_2,Y_2),$\cdots$,(X_n,Y_n),记

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i, \quad \overline{Y}=\frac{1}{n}\sum_{i=1}^{n}Y_i.$$

求证: 统计量$T=\frac{1}{n-1}\cdot\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})$是$\mathrm{cov}(X,Y)$的无偏估计量.

习题(34)参考解答

34. 1 解: 由$X_{i+1}-X_i\sim N(0,2\sigma^2)$ $\qquad E(X_{i+1}-X_i)=0$,且

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

$$E[\frac{1}{k}\cdot\sum_{i=1}^{n-1}(X_{i+1}-X_i)^2] = \frac{1}{k}\cdot\sum_{i=1}^{n-1}E(X_{i+1}-X_i)^2 = \frac{1}{k}\cdot(n-1)\cdot 2\sigma^2 \overset{\text{要求}}{=} \sigma^2,$$

所以,当 $k=2(n-1)$ 时, $\frac{1}{k}\cdot\sum_{i=1}^{n-1}(X_{i+1}-X_i)^2$ 为 σ^2 的无偏估计. ♣

34.2 解: 1) 似然函数

$$L(\theta) = \prod_{i=1}^{n}[\frac{X_i}{\theta^2}\cdot\exp\{-\frac{X_i}{\theta}\}] = (\prod_{i=1}^{n}X_i)\cdot\theta^{-2n}\cdot\exp\{-\frac{1}{\theta}\cdot\sum_{i=1}^{n}X_i\}$$

$$\ln L(\theta) = \ln(\prod_{i=1}^{n}X_i) - 2n\ln\theta - \frac{1}{\theta}\cdot\sum_{i=1}^{n}X_i.$$

令

$$\frac{\partial}{\partial\theta}\ln L(\theta) = 0 \qquad -\frac{2n}{\theta} + \frac{1}{\theta^2}\cdot\sum_{i=1}^{n}X_i = 0,$$

解得 $\hat{\theta} = \frac{1}{2n}\cdot\sum_{i=1}^{n}X_i = \frac{\overline{X}}{2}$ 为 θ 的极大似然估计量.

2) 由

$$E(\hat{\theta}) = \frac{1}{2}E(\overline{X}) = \frac{1}{2}E(X) = \frac{1}{2}\int_{-\infty}^{+\infty}x\cdot f(x;\theta)\,dx = \frac{1}{2}\int_{0}^{+\infty}x\cdot\ \frac{x}{\theta^2}e^{-\frac{x}{\theta}}dx$$

$$= \frac{\theta}{2}\int_{0}^{+\infty}t^2\cdot e^{-t}dt = \frac{\theta}{2}\cdot\Gamma(3) = \theta,$$

所以, $\hat{\theta}$ 是 θ 的无偏估计. ♣

34.3 解: 1) 已知总体 X 的密度函数为

$$f(x;\theta) = \begin{cases} 1/\theta, & 0\le x\le\theta \\ 0, & \text{其他} \end{cases}.$$

若样本值为 $x_1, x_2, \cdots, x_n$,且记

$$x_{(1)} = \min\{x_1, x_2, \cdots, x_n\},\ x_{(n)} = \max\{x_1, x_2, \cdots, x_n\}.$$

由似然函数

$$L(\theta) = \prod_{i=1}^{n}f(x_i;\theta) = \frac{1}{\theta^n},\quad 0\le x_1, x_2, \cdots, x_n\le\theta$$

$$= \frac{1}{\theta^n},\quad 0\le x_{(1)}, x_{(n)}\le\theta \qquad (*)$$

对于确定的样本值$x_1, x_2, \cdots, x_n$,由(*)式可知,只有当$\theta = x_{(n)}$时,$L(\theta)$达最大.则θ的极大似然估计值为$\hat{\theta} = \max\{x_1, x_2, \cdots, x_n\}$.所以,$\theta$的极大似然估计量为

$$\hat{\theta} = \max\{X_1, X_2, \cdots, X_n\}.$$

2) 由总体X的分布函数

$$F(x) = \int_{-\infty}^{x} f(z;\theta)dz = \begin{cases} 0 & , \quad x < 0 \\ \dfrac{x}{\theta} & , \quad 0 \le x \le \theta \\ 1 & , \quad x > \theta \end{cases},$$

故$\hat{\theta} = \max\{X_1, X_2, \cdots, X_n\}$的分布函数和密度函数分别为

$$F_{\hat{\theta}}(z) = [F(z)]^n,$$

$$f_{\hat{\theta}}(z) = n \cdot [F(z)]^{n-1} \cdot f(z;\theta) = \begin{cases} \dfrac{n}{\theta^n} \cdot z^{n-1} & , \quad 0 \le z \le \theta \\ 0 & , \quad \text{其他} \end{cases}.$$

则

$$E(\hat{\theta}) = \int_{-\infty}^{+\infty} z \cdot f_{\hat{\theta}}(z)dz = \int_0^{\theta} z \cdot \frac{n}{\theta^n} \cdot z^{n-1}dz = \frac{n}{n+1} \cdot \theta \ne \theta.$$

所以,$\hat{\theta}$不是θ的无偏估计.

3) 由

$$E(c \cdot \hat{\theta}) = c \cdot E(\hat{\theta}) = c \cdot \frac{n}{n+1} \cdot \theta \overset{\text{令}}{=} \theta,$$

取$c = \dfrac{n+1}{n}$,则统计量$c \cdot \hat{\theta}$为θ的无偏估计. ♣

34.4 分析: 本问题的关键点是求统计量$T = \dfrac{1}{n-1} \cdot \sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y})$的数学期望$E(T)$,然后验证:$E(T) = \operatorname{cov}(X, Y)$.

证: 经化简,可得

$$T = \frac{1}{n-1} \cdot \sum_{i=1}^{n} X_i Y_i - \frac{n}{n-1} \cdot \overline{X} \cdot \overline{Y}.$$

再注意到

$$E(X_i) = E(\overline{X}) = E(X),\ E(Y_i) = E(\overline{Y}) = E(Y),$$

$$E(X_i Y_i) = E(XY),\ \operatorname{cov}(X_i, Y_i) = \operatorname{cov}(X, Y),\ i = 1,2,\cdots,n.$$

则

$$E(T)=E(\frac{1}{n-1}\cdot\sum_{i=1}^{n}X_iY_i-\frac{n}{n-1}\cdot\overline{X}\cdot\overline{Y})$$

$$=\frac{1}{n-1}\cdot\sum_{i=1}^{n}E(X_iY_i)-\frac{n}{n-1}\cdot E(\overline{X}\cdot\overline{Y})=\frac{n}{n-1}\cdot[E(XY)-E(\overline{X}\cdot\overline{Y})]$$

$$=\frac{n}{n-1}\cdot[\,(E(XY)-E(X)\cdot E(Y))-(E(\overline{X}\cdot\overline{Y})-E(X)\cdot E(Y))\,]$$

$$=\frac{n}{n-1}\cdot[\,(E(XY)-E(X)\cdot E(Y))-(E(\overline{X}\cdot\overline{Y})-E(\overline{X})\cdot E(\overline{Y}))\,]$$

$$=\frac{n}{n-1}\cdot[\mathrm{cov}(X,Y)-\mathrm{cov}(\overline{X},\overline{Y})].$$

又由

$$\mathrm{cov}(\overline{X},\overline{Y})=\mathrm{cov}(\frac{1}{n}\sum_{i=1}^{n}X_i\,,\ \frac{1}{n}\sum_{j=1}^{n}Y_j)=\frac{1}{n^2}\cdot\mathrm{cov}(\sum_{i=1}^{n}X_i\,,\sum_{j=1}^{n}Y_j)$$

$$=\frac{1}{n^2}\cdot\sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{cov}(X_i,Y_j)=\frac{1}{n^2}\cdot\sum_{i=1}^{n}\mathrm{cov}(X_i,Y_i)$$

$$=\frac{1}{n}\cdot\mathrm{cov}(X,Y),$$

$$E(T)=\frac{n}{n-1}\cdot[\mathrm{cov}(X,Y)-\frac{1}{n}\cdot\mathrm{cov}(X,Y)]=\mathrm{cov}(X,Y)\ .$$

所以,统计量$T=\frac{1}{n-1}\cdot\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})$是$\mathrm{cov}(X,Y)$的无偏估计量. ♣