

数理统计第五章测验题

考试时间 2022 年 4 月 10 日, 答卷时间 120 分钟, 总分 100 分, 出题人-李绍文

1. (15 分) 设总体 X 的分布列为

$$P\{X = -1\} = \frac{1-\theta}{2}, \quad P\{X = 0\} = \frac{1}{2}, \quad P\{X = 1\} = \frac{\theta}{2},$$

且 X_1, X_2, \dots, X_n 是样本, 用两种方法给出样本联合质量函数。

解: 方法一: 设 n_{-1}, n_0, n_1 分别是 x_1, x_2, \dots, x_n 中取 $-1, 0, 1$ 的样品个数, 则样本联合质量函数为

$$p(x_1, x_2, \dots, x_n; \theta) = \left(\frac{1-\theta}{2}\right)^{n_{-1}} \left(\frac{1}{2}\right)^{n_0} \left(\frac{\theta}{2}\right)^{n_1} = \frac{1}{2^n} (1-\theta)^{n_{-1}} \theta^{n_1}.$$

方法二: 总体 X 的质量函数为

$$p(x; \theta) = \left(\frac{1-\theta}{2}\right)^{\frac{1}{2}x(x-1)} \left(\frac{1}{2}\right)^{-(x+1)(x-1)} \left(\frac{\theta}{2}\right)^{\frac{1}{2}x(x+1)} = \frac{1}{2} (1-\theta)^{\frac{1}{2}(x^2-x)} \theta^{\frac{1}{2}(x^2+x)}, \quad x = -1, 0, 1,$$

故样本联合质量函数为

$$p(x_1, x_2, \dots, x_n; \theta) = \prod_{i=1}^n \frac{1}{2} (1-\theta)^{\frac{1}{2}(x_i^2-x_i)} \theta^{\frac{1}{2}(x_i^2+x_i)} = \frac{1}{2^n} (1-\theta)^{\frac{1}{2}\left(\sum_{i=1}^n x_i^2 - \sum_{i=1}^n x_i\right)} \theta^{\frac{1}{2}\left(\sum_{i=1}^n x_i^2 + \sum_{i=1}^n x_i\right)}$$

$$\text{或 } p(x_1, x_2, \dots, x_n; \theta) = \left(\frac{1-\theta}{2}\right)^{\frac{1}{2}\sum_{i=1}^n x_i(x_i-1)} \left(\frac{1}{2}\right)^{-\sum_{i=1}^n (x_i+1)(x_i-1)} \left(\frac{\theta}{2}\right)^{\frac{1}{2}\sum_{i=1}^n x_i(x_i+1)},$$

$$x_1, x_2, \dots, x_n = -1, 0, 1.$$

2. (10 分) 设总体 X 的 2 阶矩存在, X_1, X_2, \dots, X_n 是样本, \bar{X} 为样本均值,

试证 $E\left[\sum_{i=1}^n (X_i - \bar{X})^2\right] = (n-1)\text{Var}(X)$, 从而 $E(S^2) = \text{Var}(X)$, 但 $E(S^{*2}) \neq \text{Var}(X)$ 。

证明: 因

$$E\left[\sum_{i=1}^n (X_i - \bar{X})^2\right] = E\left[\sum_{i=1}^n X_i^2 - n\bar{X}^2\right] = \sum_{i=1}^n E(X_i^2) - nE(\bar{X}^2),$$

$$E(X_i^2) = \text{Var}(X_i) + [E(X_i)]^2 = \text{Var}(X) + [E(X)]^2,$$

$$E(\bar{X}^2) = \text{Var}(\bar{X}) + [E(\bar{X})]^2 = \frac{1}{n} \text{Var}(X) + [E(X)]^2,$$

则

$$E\left[\sum_{i=1}^n (X_i - \bar{X})^2\right] = n\{\text{Var}(X) + [E(X)]^2\} - n\left\{\frac{1}{n}\text{Var}(X) + [E(X)]^2\right\} = (n-1)\text{Var}(X),$$

故

$$E(S^2) = \text{Var}(X),$$

但

$$E(S^{*2}) = \frac{n-1}{n}\text{Var}(X) \neq \text{Var}(X)。$$

3. (15 分) 设总体 X 服从指数分布 $Exp(\lambda)$, X_1, X_2, \dots, X_n 是样本, $X_{(1)}$ 与 $X_{(n)}$ 分别是最小与最大顺序统计量, 求 $X_{(1)}$ 与 $X_{(n)}$ 密度函数 $p_1(x)$ 与 $p_n(x)$ 以及 $(X_{(1)}, X_{(n)})$ 的联合密度函数, 并问 $X_{(1)}$ 与 $X_{(n)}$ 是否独立?

解: 因总体密度函数与分布函数分别为

$$p(x) = \lambda e^{-\lambda x} I_{x>0}, \quad F(x) = \begin{cases} 1 - e^{-\lambda x}, & x > 0; \\ 0, & x \leq 0. \end{cases}$$

则最小与最大顺序统计量 $X_{(1)}$ 与 $X_{(n)}$ 的密度函数分别为

$$p_1(x) = n[1 - F(x)]^{n-1} p(x) = n\lambda e^{-n\lambda x} I_{x>0},$$

$$p_n(x) = n[F(x)]^{n-1} p(x) = n\lambda e^{-\lambda x} (1 - e^{-\lambda x})^{n-1} I_{x>0},$$

且 $(X_{(1)}, X_{(n)})$ 的联合密度函数为

$$p_{1n}(y, z) = n(n-1)[F(z) - F(y)]^{n-2} p(y)p(z)I_{y<z} = n(n-1)\lambda^2 e^{-\lambda(y+z)}(e^{-\lambda y} - e^{-\lambda z})^{n-2} I_{0<y<z}。$$

可见 $X_{(1)}$ 与 $X_{(n)}$ 不独立。

4. (15 分) 设总体 X 的 2 阶矩存在, X_1, X_2, \dots, X_n 是样本, 证明 $X_i - \bar{X}$ 与 $X_j - \bar{X}$ ($i \neq j$) 的相关系数为 $-(n-1)^{-1}$ 。

证明: 因 X_1, X_2, \dots, X_n 相互独立, 有 $\text{Cov}(X_l, X_k) = 0$, ($l \neq k$), 则

$$\begin{aligned} \text{Cov}(X_i - \bar{X}, X_j - \bar{X}) &= \text{Cov}(X_i, X_j) - \text{Cov}(X_i, \bar{X}) - \text{Cov}(\bar{X}, X_j) + \text{Cov}(\bar{X}, \bar{X}) \\ &= 0 - \text{Cov}(X_i, \frac{1}{n} X_i) - \text{Cov}(\frac{1}{n} X_j, X_j) + \text{Var}(\bar{X}) \\ &= -\frac{1}{n} \text{Var}(X_i) - \frac{1}{n} \text{Var}(X_j) + \text{Var}(\bar{X}) = -\frac{1}{n} \sigma^2 - \frac{1}{n} \sigma^2 + \frac{1}{n} \sigma^2 = -\frac{1}{n} \sigma^2, \end{aligned}$$

且

$$\begin{aligned}\text{Var}(X_i - \bar{X}) &= \text{Var}(X_i) + \text{Var}(\bar{X}) - 2\text{Cov}(X_i, \bar{X}) \\&= \sigma^2 + \frac{1}{n}\sigma^2 - 2\text{Cov}(X_i, \frac{1}{n}X_i) \\&= \sigma^2 + \frac{1}{n}\sigma^2 - \frac{2}{n}\sigma^2 = \frac{n-1}{n}\sigma^2 = \text{Var}(X_j - \bar{X}),\end{aligned}$$

故

$$\text{Corr}(X_i - \bar{X}, X_j - \bar{X}) = \frac{\text{Cov}(X_i - \bar{X}, X_j - \bar{X})}{\sqrt{\text{Var}(X_i - \bar{X})} \cdot \sqrt{\text{Var}(X_j - \bar{X})}} = \frac{-\frac{1}{n}\sigma^2}{\sqrt{\frac{n-1}{n}\sigma^2} \cdot \sqrt{\frac{n-1}{n}\sigma^2}} = -\frac{1}{n-1}.$$

5. (10分) 设 X_1, X_2, X_3 相互独立且都服从 $N(0, 1)$, 问 $\xi = \frac{3X_1^2}{X_1^2 + X_2^2 + X_3^2}$ 是否服从 F 分布。若是, 自由度是多少? 若不是, 构造一个服从 F 分布的随机变量 T , 指出 T 的自由度, 并将 ξ 表示为 T 的函数。

解: 虽然 ξ 是 χ^2 变量之商, 但分子分母不独立, 不能判断 ξ 服从 F 分布。

为了使得构造的 F 分布随机变量 T 分子分母独立, 在分母中只保留 $X_2^2 + X_3^2$ 。根据 F 分布的构成可知

$$T = \frac{X_1^2}{(X_2^2 + X_3^2)/2} = \frac{2X_1^2}{X_2^2 + X_3^2} \sim F(1, 2) \text{ 或 } T = \frac{X_2^2 + X_3^2}{2X_1^2} \sim F(2, 1),$$

并且可得

$$\xi = \frac{\frac{3X_1^2}{X_2^2 + X_3^2}}{\frac{X_1^2}{X_2^2 + X_3^2} + 1} = \frac{\frac{3}{2}T}{\frac{1}{2}T + 1} = \frac{3T}{T + 2} \text{ 或 } \xi = \frac{3}{1 + \frac{X_2^2 + X_3^2}{X_1^2}} = \frac{3}{1 + 2T}.$$

或

$$T = \frac{X_2^2 + X_3^2}{2X_1^2} \sim F(2, 1), \quad \xi = \frac{3}{1 + \frac{X_2^2 + X_3^2}{X_1^2}} = \frac{3}{1 + 2T}.$$

6. (10分) 设 X_1, X_2, \dots, X_{15} 相互独立且都服从 $N(0, 0.2^2)$, 求 $P\left\{\sum_{i=1}^{15} X_i^2 < 1\right\}$ 。

解: 因 $X_i \sim N(0, 0.2^2)$, 有

$$\frac{X_i - 0}{0.2} = 5X_i \sim N(0, 1),$$

则

$$\sum_{i=1}^{15} (5X_i)^2 = 25 \sum_{i=1}^{15} X_i^2 \sim \chi^2(15),$$

故

$$P\left\{\sum_{i=1}^{15} X_i^2 < 1\right\} = P\left\{25 \sum_{i=1}^{15} X_i^2 < 25\right\} \approx P\left\{25 \sum_{i=1}^{15} X_i^2 < \chi_{0.95}^2(15)\right\} = 0.95.$$

7. (15 分) 设总体 $X \sim N(-0.5, 4)$, 且 X_1, X_2, \dots, X_{16} 为样本, 求 $P\{\bar{X} > 0\}$,

$P\{S^2 > 6\}$ 。

解: 因 $\mu = -0.5$, $\sigma^2 = 4$, 样本容量 $n = 16$, 有

$$U = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = 2\bar{X} + 1 \sim N(0, 1),$$

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} = \frac{15S^2}{4} \sim \chi^2(15),$$

故

$$P\{\bar{X} > 0\} = P\{U = 2\bar{X} + 1 > 1\} = 1 - \Phi(1) = 0.1587,$$

$$P\{S^2 > 6\} = P\left\{\chi^2 = \frac{15S^2}{4} > 22.5\right\} = P\{\chi^2 > \chi_{0.9}^2(15)\} = 1 - 0.9 = 0.1.$$

8. (10 分) 设 X_1, X_2, \dots, X_n 是来自伽玛分布 $Ga(\alpha, \lambda)$, $\alpha > 0, \lambda > 0$ 的一个

样本, 伽玛分布密度函数为 $p(x; \alpha, \lambda) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} I_{x>0}$, 求 (α, λ) 的充分统计量。

解: 总体 X 的密度函数为

$$p(x; \alpha, \lambda) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} I_{x>0},$$

样本联合密度函数为

$$p(x_1, \dots, x_n; \alpha, \lambda) = \frac{\lambda^{n\alpha}}{[\Gamma(\alpha)]^n} (x_1 x_2 \cdots x_n)^{\alpha-1} e^{-\lambda \sum_{i=1}^n x_i} I_{x_1, \dots, x_n > 0},$$

令 $t_1 = x_1 x_2 \cdots x_n, t_2 = \sum_{i=1}^n x_i$, 取

$$g(t_1, t_2; \alpha, \lambda) = \frac{\lambda^{n\alpha}}{[\Gamma(\alpha)]^n} t_1^{\alpha-1} e^{-\lambda t_2}, \quad h(x_1, \dots, x_n) = I_{x_1, \dots, x_n > 0},$$

根据因子分解定理可知 $(T_1, T_2) = \left(X_1 X_2 \cdots X_n, \sum_{i=1}^n X_i \right)$ 是参数 (α, λ) 的充分统计量。