数理统计第五章测验题

考试时间 2023 年 4 月 9 日，答卷时间 120 分钟，总分 100 分，出题人-李绍文

1．（10 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 $X_1, X_2, \cdots, X_n$ 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 $x_1, x_2, \cdots, x_n$ 中取 $-1, 0, 1$ 的样品个数，则样本联合质量函数为

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

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

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

或 $p(x_1, x_2, \cdots, x_n; \theta)=\dfrac{1}{2^n}(1-\theta)^{\frac{1}{2}\left(\sum_{i=1}^{n}x_i^2-\sum_{i=1}^{n}x_i\right)}\theta^{\frac{1}{2}\left(\sum_{i=1}^{n}x_i^2+\sum_{i=1}^{n}x_i\right)},\quad x_1, x_2, \cdots, x_n=-1, 0, 1$。

2．（10 分）设总体 X 服从均匀分布 $U(0,\theta)$，$X_1, X_2, \cdots, X_n$ 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

$$p(x)=\frac{1}{\theta}I_{0<x<\theta},\quad F(x)=\begin{cases}0, & x<0;\\ \dfrac{x}{\theta}, & 0\le x<\theta;\\ 1, & x\ge\theta.\end{cases}$$

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

$$p_1(x)=n[1-F(x)]^{n-1}p(x)=\frac{n(\theta-x)^{n-1}}{\theta^n}I_{0<x<\theta},\quad p_n(x)=n[F(x)]^{n-1}p(x)=\frac{nx^{n-1}}{\theta^n}I_{0<x<\theta},$$

且 $(X_{(1)}, X_{(n)})$ 的联合密度函数为

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

可见 $X_{(1)}$ 与 $X_{(n)}$ 不独立。

3．（10 分）设总体 X 的 2 阶矩存在，$X_1, X_2, \cdots, X_n$ 是样本，求 $X_i-\bar{X}$ 与 $X_j-\bar{X}$ $(i\neq j)$ 的相关系数。

解：因 $X_1, X_2, \cdots, X_n$ 相互独立，有 $\text{Cov}(X_i, X_j)=0,\quad (i\neq j)$，且

$$\text{Cov}(X_i,\bar{X})=\text{Cov}\left(X_i,\frac{1}{n}X_i\right)=\frac{1}{n}\sigma^2\text{，}\quad \text{Var}(\bar{X})=\frac{1}{n}\sigma^2\text{，}$$

则

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

且

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

故

$$\text{Corr}(X_i-\bar{X},X_j-\bar{X})=\frac{\text{Cov}(X_i-\bar{X},X_j-\bar{X})}{\sqrt{\text{Var}(X_i-\bar{X})}\sqrt{\text{Var}(X_j-\bar{X})}}=-\frac{1}{n-1}\text{。}$$

4．（10 分）设 $X_1, X_2, \cdots, X_{15}$ 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

$$\frac{X_i-0}{0.2}=5X_i\sim N(0,1)\text{，}\quad \sum_{i=1}^{15}(5X_i)^2=25\sum_{i=1}^{15}X_i^2\sim\chi^2(15)\text{，}$$

故

$$P\left\{\sum_{i=1}^{15}X_i^2<1\right\}=P\left\{25\sum_{i=1}^{15}X_i^2<25\right\}\approx P\left\{25\sum_{i=1}^{15}X_i^2<\chi^2_{0.95}(15)\right\}=0.95\text{。}$$

5．（15 分）设 X_1, X_2, X_3, X_4 相互独立且都服从标准正态分布 $N(0,1)$，问 $\xi=\dfrac{2(X_1^2+X_2^2)}{X_1^2+X_2^2+X_3^2+X_4^2}$ 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 Y，指出 Y 的自由度，并将 ξ 表示为 Y 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$X_3^2+X_4^2$。根据F分布的构成可知

$$Y=\frac{(X_1^2+X_2^2)/2}{(X_3^2+X_4^2)/2}=\frac{X_1^2+X_2^2}{X_3^2+X_4^2}\sim F(2,2),$$

并且可得

$$\xi=\frac{2\dfrac{X_1^2+X_2^2}{X_3^2+X_4^2}}{\dfrac{X_1^2+X_2^2}{X_3^2+X_4^2}+1}=\frac{2Y}{Y+1}。$$

6．（15 分）设总体$X\sim N(-0.5,4)$，且$X_1,X_2,\cdots,X_{16}$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

$$U=\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}=2\bar{X}+1\sim N(0,1),\quad \chi^2=\frac{(n-1)S^2}{\sigma^2}=\frac{15S^2}{4}\sim\chi^2(15),$$

故

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

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

7．（15 分）设总体$X\sim N(0,2)$，$Y\sim N(3,4)$且相互独立，$X_1,X_2,\cdots,X_{10}$与$Y_1,Y_2,\cdots,Y_5$分别为X与Y的样本，求$P\{\bar{X}>\bar{Y}\}$，$P\{S_x^2>3S_y^2\}$。

解：因$\mu_1=0$，$\sigma_1^2=2$，$\mu_2=3$，$\sigma_2^2=4$，样本容量$n_1=10$，$n_2=5$，有

$$U=\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}=\bar{X}-\bar{Y}+3\sim N(0,1),\quad F=\frac{S_x^2/\sigma_1^2}{S_y^2/\sigma_2^2}=2\frac{S_x^2}{S_y^2}\sim F(9,4),$$

故

$$P\{\bar{X}>\bar{Y}\}=P\{U=\bar{X}-\bar{Y}+3>3\}=1-\Phi(3)=0.0013,$$

$$P\{S_x^2>3S_y^2\}=P\left\{F=2\frac{S_x^2}{S_y^2}>6\right\}=P\{F>f_{0.95}(9,4)\}=0.05。$$

8．（15 分）设$X_1,X_2,\cdots,X_n$是来自正态分布$N(\mu,\sigma^2)$的样本，求(μ,σ^2)的

一个充分统计量。并问 $(\bar{X}, S^2)$ 是否 (μ, σ^2) 的一个充分统计量，为什么？

解：总体 X 的密度函数为

$$p(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

样本联合密度函数为

$$p(x_1, \cdots, x_n; \mu, \sigma^2) = \frac{1}{(\sqrt{2\pi}\sigma)^n} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2} = \frac{1}{(\sqrt{2\pi}\sigma)^n} e^{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n}x_i^2 - 2\mu\sum_{i=1}^{n}x_i + \mu^2\right)},$$

令 $t_1 = \sum_{i=1}^{n} x_i, t_2 = \sum_{i=1}^{n} x_i^2$，取

$$g(t_1, t_2; \mu, \sigma^2) = \frac{1}{(\sqrt{2\pi}\sigma)^n} e^{-\frac{1}{2\sigma^2}(t_2 - 2\mu t_1 + n\mu^2)}, \quad h(x_1, \cdots, x_n) = 1,$$

根据因子分解定理可知 $(T_1, T_2) = \left(\sum_{i=1}^{n} X_i, \sum_{i=1}^{n} X_i^2\right)$ 是参数 (μ, σ^2) 的充分统计量。

又令 $t_1 = \bar{x}, t_2 = s^2$，有

$$\sum_{i=1}^{n} x_i = nt_1, \quad \sum_{i=1}^{n} x_i^2 = (n-1)s^2 + n\bar{x}^2 = (n-1)t_2 + nt_1^2,$$

则

$$g(t_1, t_2; \mu, \sigma^2) = \frac{1}{(\sqrt{2\pi}\sigma)^n} e^{-\frac{1}{2\sigma^2}[(n-1)t_2 + nt_1^2 - 2n\mu t_1 + n\mu^2]}, \quad h(x_1, \cdots, x_n) = 1,$$

根据因子分解定理可知 $(T_1, T_2) = (\bar{X}, S^2)$ 是参数 (μ, σ^2) 的充分统计量。