多元统计分析期末考试

Phlinsia

2024 年 6 月 28 日

题目 1. 2.32（改）

给定了随机向量 $\mathbf{X}^\top = (X_1, X_2, X_3, X_4)$，均值向量 $\mu^\top = (3, 2, -2, 0)$，以及方差-协方差矩阵 $\mathbf{\Sigma}_{\boldsymbol{X}}$

$$\mathbf{\Sigma}_{\boldsymbol{X}} = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

又

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & -2 & 0 \\ 1 & 1 & 1 & -3 \end{bmatrix}$$

求 AX 的均值 $E(AX)$ 和协方差矩阵 $Cov(AX)$。

解答.

(a)

$$E(\mathbf{X}) = \mu$$

$$E(\mathbf{AX}) = \mathbf{A}\mu = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & -2 & 0 \\ 1 & 1 & 1 & -3 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ -2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 9 \\ 3 \end{bmatrix}$$

(b)

$$Cov(\mathbf{X}) = \mathbf{\Sigma}_{\boldsymbol{X}}$$

$$Cov(\mathbf{AX}) = \mathbf{A}\mathbf{\Sigma}_{\boldsymbol{X}}\mathbf{A}^\top = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & -2 & 0 \\ 1 & 1 & 1 & -3 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ 0 & -2 & 1 \\ 0 & 0 & -3 \end{bmatrix} = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 18 & 0 \\ 0 & 0 & 36 \end{bmatrix}$$

题目 2. 5.1（改）

$$\mathbf{X} = \begin{bmatrix} -5 & 6 \\ 1 & 3 \\ -1 & 3 \\ 1 & 4 \end{bmatrix}$$

在 $\alpha = 0.05$ 的显著性水平下检验原假设 $H_0 : \mu' = [0, 5]^\top$

解答.

$$\text{变量数} p = 2, \text{样本大小} n = 4, \bar{x} = \begin{bmatrix} \frac{-5+1+-1+1}{n} \\ \frac{6+3+3+4}{n} \end{bmatrix} = \begin{bmatrix} -1 \\ 4 \end{bmatrix},$$

$$\because \mathbf{S}_{ij} = \frac{1}{n-1} \sum_{k=1}^{n} (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j),$$

$$s_{11} = \frac{1}{3}((-5 - -1)^2 + (1 - -1)^2 + (-1 - -1)^2 + (1 - -1)^2) = 8,$$

$$s_{12} = s_{12} = \frac{1}{3}((-5 - -1)(6 - 4) + (1 - -1)(3 - 4)$$

$$+(-1 - -1)(3 - 4) + (1 - -1)(4 - 4)) = -\frac{10}{3},$$

$$s_{22} = \frac{1}{3}((6 - 4)^2 + (3 - 4)^2 + (3 - 4)^2 + (4 - 4)^2) = 2,$$

$$\therefore \mathbf{S} = \begin{bmatrix} 8 & -\frac{10}{3} \\ -\frac{10}{3} & 2 \end{bmatrix}, \text{又} \begin{bmatrix} a & b \\ b & c \end{bmatrix}^{-1} = \frac{1}{ac - b^2} \begin{bmatrix} c & -b \\ -b & a \end{bmatrix}$$

$$T^2 = n(\bar{x} - \mu')^\top \mathbf{S}^{-1} (\bar{x} - \mu') =$$

$$4 \begin{bmatrix} -1 - 0 & 4 - 5 \end{bmatrix} \frac{9}{44} \cdot \frac{1}{3} \begin{bmatrix} 6 & 10 \\ 10 & 24 \end{bmatrix} \begin{bmatrix} -1 - 0 \\ 4 - 5 \end{bmatrix} = 13.64$$

$$\because T^2 \sim \frac{(n-1)p}{(n-p)} F_{p,n-p}, \quad \therefore T^2 \sim 3F_{2,2} \quad H_0 \colon \mu' = [0, 5]$$

$$\because \alpha = 0.05 \therefore F_{2,2}(0.05) = 19$$

$$\because T^2 = 13.64 < F_{2,2}(0.05) = 57 \quad \therefore \alpha = 0.05 \text{时接受原假设} H_0$$

题目 3. 6.8.

对三种处理收集了两种响应的观测值, 观测值向量 $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ 为

处理 1：6 , 5 , 8 , 4 , 7

处理 2：3 , 1 , 2

处理 3：2 , 5 , 3 , 2

(a) 对每个观测量做下述分解

$$x_{ij} = \bar{x} + (\bar{x_i} - \bar{x_i}) + (x_{ij} - \bar{x_i})$$

观测值、总均值、处理效应、残差效应。

(b) 计算处理效应平方和 $SS_{处理}$ 与残差平方和 $SS_{残差}$。

解答.

$$均值：\frac{6+5+8+4+7+3+1+2+2+5+3+2}{12} = 4$$

$$\therefore \begin{bmatrix} 4 & 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & & \\ 4 & 4 & 4 & 4 & \end{bmatrix}_{均值}$$

$$处理：\frac{6+5+8+4+7}{5} = 6 = 4+2, \quad \frac{3+1+2}{3} = 2 = 4+(-2)$$

$$\frac{2+5+3+2}{4} = 3 = 4+(-1) \therefore \begin{bmatrix} 2 & 2 & 2 & 2 & 2 \\ -2 & -2 & -2 & & \\ -1 & -1 & -1 & -1 & \end{bmatrix}_{处理}$$

$$\therefore \begin{bmatrix} 6-6 & 5-6 & 8-6 & 4-6 & 7-6 \\ 3-2 & 1-2 & 2-2 & & \\ 2-3 & 5-3 & 3-3 & 2-3 & \end{bmatrix} = \begin{bmatrix} 0 & -1 & 2 & -2 & 1 \\ 1 & -1 & 0 & & \\ -1 & 2 & 0 & -1 & \end{bmatrix}_{残差}$$

对于变量 1：

$$\begin{bmatrix} 6 & 5 & 8 & 4 & 7 \\ 3 & 1 & 2 & & \\ 2 & 5 & 3 & 2 & \end{bmatrix}_{\text{观测}} = \begin{bmatrix} 4 & 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & & \\ 4 & 4 & 4 & 4 & \end{bmatrix}_{\text{均值}} + \begin{bmatrix} 2 & 2 & 2 & 2 & 2 \\ -2 & -2 & -2 & & \\ -1 & -1 & -1 & -1 & \end{bmatrix}_{\text{处理}}$$

$$+ \begin{bmatrix} 0 & -1 & 2 & -2 & 1 \\ 1 & -1 & 0 & & \\ -1 & 2 & 0 & -1 & \end{bmatrix}_{\text{残差}},$$

计算对应平方和：$SS_{\text{处理}} = 5 \times 2^2 + 3 \times (-2)^2 + 4 \times (-1)^4 = 36$

$$SS_{\text{残差}} = (-1)^2 + 2^2 + (-2)^2 + 1^2 + 1^2 + (-1)^2 + (-1)^2 + 2^2 + (-1)^2 = 18$$

$$SS_{\text{修正后总和}} = SS_{\text{观测}} - SS_{\text{均值}} = 246 - 192 = 54$$

题目 4. 例 11.2【换数字】

设组别 π_1, π_2 的概率密度函数分别是 $f_1(x)$ 和 $f_2(x)$，又知 $c(1|2) = 12$，$c(2|1) = 4$，根据以往经验给出 $p_1 = 0.6, p_2 = 0.4$

(a) 给出将一个心得观测值分入两个总体之一的最小 ECM 规则的一般形式。

(b) 对一个新的观测值产生的密度函数 $f_1(x_0) = 0.36, f_2(x_0) = 0.24$，根据以上信息将新项目分到哪个组别？

解答. 最小 ECM 判别规则为：

$$\begin{cases} \text{若} \dfrac{f_1(x)}{f_2(x)} \geq \dfrac{c(1|2)p_2}{c(2|1)p_1} = \dfrac{12 \times 0.4}{4 \times 0.6} = 2 & \text{则} x \in \pi_1 \\ \text{若} \dfrac{f_2(x)}{f_2(x)} < \dfrac{c(1|2)p_2}{c(2|1)p_1} = \dfrac{12 \times 0.4}{4 \times 0.6} = 2 & \text{则} x \in \pi_2 \end{cases}$$

对于新的观测值 x_0，有

$$\frac{f_1(x_0)}{f_2(x_0)} = \frac{0.36}{0.24} = 1.5 < 2, \quad \text{将新项目分到组别} \pi_2$$

题目 5. 确定下列协方差矩阵的总体主成分 $\zeta_1, \zeta_2, \zeta_3$

$$\Sigma = \begin{bmatrix} 5 & 2 & 0 \\ 2 & 4 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

并计算总体方差中第一主成分所解释的比例。

解答. 根据特征方程 $|\lambda I - E| = 0$，有特征值和特征向量对为：

$$\begin{cases} \lambda_1 = 6.562, & e_1^\top = \begin{bmatrix} 0.788 & 0.615 & 0 \end{bmatrix} \\ \lambda_2 = 2.438, & e_2^\top = \begin{bmatrix} 0.615 & -0.788 & 0 \end{bmatrix} \\ \lambda_3 = 2, & e_3^\top = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \end{cases}$$

因此，主成分为

$$\begin{cases} \zeta_1 = 0.788X_1 + 0.615X_2 \\ \zeta_2 = 0.615X_1 - 0.788X_2 \\ \zeta_3 = X_3 \end{cases}$$

第一主成分所 ζ_1 可以验证：

$$\begin{aligned} Var(\zeta_1) &= Var(0.788X_1 + 0.615X_2) \\ &= 0.788^2 Var(X_1) + 0.615^2 Var(X_2) + 2 \cdot 0.788 \cdot 0.615 Cov(X_1, X_2) \\ &= 0.788^2 \cdot 5 + 0.615^2 \cdot 4 + 2 \cdot 0.788 \cdot 0.615 \cdot 2 \\ &= 6.557 \end{aligned}$$

因此，第一主成分所解释的比例为:

$$\frac{Var(\zeta_1)}{Var(X_1) + Var(X_2) + Var(X_3)} = \frac{6.56}{6.56 + 2.43 + 2} = 59.65\%$$

题目 5 的注记. 特征方程与特征向量：

$$|\lambda I - \Sigma| = \begin{vmatrix} \lambda - 5 & -2 & 0 \\ -2 & \lambda - 4 & 0 \\ 0 & 0 & \lambda - 2 \end{vmatrix} = 0$$

$$(\lambda-5)(\lambda-4)(\lambda-2) - (-2)^2(\lambda-2) = 0 \quad \text{此处直接敲计算器}$$

$$(\lambda-2)(\lambda^2 - 9\lambda + 16) = 0$$

$$\lambda_{1,2} = \frac{-b \pm \sqrt{b^2-4ac}}{2a} = \frac{9 \pm \sqrt{81-64}}{2} = \frac{9 \pm \sqrt{17}}{2}$$

$$\lambda_1 = \frac{9+\sqrt{17}}{2} \approx 6.562,\ \lambda_2 = \frac{9-\sqrt{17}}{2} \approx 2.438,\ \lambda_3 = 2$$

解线性方程组 $(\lambda I - \Sigma)e = 0$ 得到特征向量，如：

$$\begin{bmatrix} \lambda_1 - 5 & -2 & 0 \\ -2 & \lambda_1 - 4 & 0 \\ 0 & 0 & \lambda_1 - 2 \end{bmatrix} = \begin{bmatrix} 1.562 & -2 & 0 \\ -2 & 2.562 & 0 \\ 0 & 0 & 4.562 \end{bmatrix}$$

设特征向量 $e_1 = (x, y, z)^\top$，我们有：

$$\begin{cases} 1.562x - 2y & = 0 \\ -2x + 2.562y & = 0 \\ 4.562z & = 0 \end{cases}$$

解得 $e_1 = (1, 0.781, 0)^\top = (0.788, 0.615, 0)^\top$，以此类推。