

# 多元统计分析课程作业 4

Phlinsia

2024 年 5 月 13 日

题目 1. 6.6.

利用练习 6.8 中处理 2 和处理 3 的数据：处理 2:  $\begin{bmatrix} 3 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 6 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}$

处理 3:  $\begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 5 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}$

- (a) 计算  $S_p$ .
- (b) 用双样本方法以  $\alpha = 0.01$  检验  $H_0: \mu_2 - \mu_3 = 0$
- (c) 构造  $\mu_{2i} = \mu_{3i}, i = 1, 2$  的 99% 联合置信区间

解答.

(a) 处理 2: 样本均值向量  $\begin{bmatrix} 2 \\ 4 \end{bmatrix}$ ; 样本协方差矩阵  $\begin{bmatrix} 1 & -\frac{3}{2} \\ -\frac{3}{2} & 3 \end{bmatrix}$

处理 3: 样本均值向量  $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ ; 样本协方差矩阵  $\begin{bmatrix} 2 & -\frac{4}{3} \\ -\frac{4}{3} & \frac{4}{3} \end{bmatrix}$

$\therefore S_{\text{pooled}} = \begin{bmatrix} 1.6 & -1.4 \\ -1.4 & 2 \end{bmatrix}$

(b)

$$\begin{aligned}
T^2 &= n(\bar{x}_1 - \bar{x}_2)^\top S_{\text{pooled}}^{-1}(\bar{x}_1 - \bar{x}_2) \\
&= \begin{bmatrix} 2-3 & 4-2 \end{bmatrix} \left[ \left( \frac{1}{3} + \frac{1}{4} \right) \begin{bmatrix} 1.6 & -1.4 \\ -1.4 & 2 \end{bmatrix} \right]^{-1} \begin{bmatrix} 2-3 \\ 4-2 \end{bmatrix} = 3.88 \\
&\frac{(n_1 + n_2 - 2)p}{n_1 + n_2 - p - 1} F_{p, n_1 + n_2 - p - 1}(0.01) = \frac{(5)2}{3} F_{2,4}(0.01) = 45 \\
&\therefore T^2 = 3.88 < 45 \therefore \text{接受 } H_0
\end{aligned}$$

(c) 99% 联合置信区间为

$$\begin{aligned}
\mu_{21} - \mu_{31} &: (2-3) \pm \sqrt{45} \sqrt{\left( \frac{1}{3} + \frac{1}{4} \right)} 1.6 = -1 \pm 6.5, \text{ 即 } [-7.5, 5.5] \\
\mu_{22} - \mu_{32} &: (4-2) \pm \sqrt{45} \sqrt{\left( \frac{1}{3} + \frac{1}{4} \right)} 2 = 2 \pm 7.2, \text{ 即 } [-5.2, 9.2]
\end{aligned}$$

**题目 2. 6.7.**

利用例 6.4 中的电力需求数据的汇总统计数字计算  $T^2$ ，并检验假设  $H_0: \mu_1 - \mu_2 = 0$ . (假定  $\mathbf{1} = \mathbf{2}, \alpha = 0.05$ ). 另外, 确定对拒绝  $H_0$  起关键作用的均值分量的线性组合。

**解答.**

$$\begin{aligned}
T^2 &= \begin{bmatrix} 74.4 & 201.6 \end{bmatrix} \left[ \left( \frac{1}{45} + \frac{1}{55} \right) \begin{bmatrix} 10963.7 & 21505.5 \\ 21505.5 & 63661.3 \end{bmatrix} \right]^{-1} \begin{bmatrix} 74.4 \\ 201.6 \end{bmatrix} = 16.1 \\
&\frac{(n_1 + n_2 - 2)p}{n_1 + n_2 - p - 1} F_{p, n_1 + n_2 - p - 1}(0.05) = 6.26 \therefore T^2 = 16.1 > 6.26 \therefore \text{拒绝 } H_0 \\
\text{线性组合: } \hat{\mathbf{a}} &\propto \left( \frac{1}{n_1} \mathbf{S}_1 + \frac{1}{n_2} \mathbf{S}_2 \right)^{-1} (\bar{x}_1 - \bar{x}_2) = S_{\text{pooled}}^{-1}(\bar{x}_1 - \bar{x}_2) = \begin{bmatrix} 0.0017 \\ 0.0026 \end{bmatrix}
\end{aligned}$$

**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:  $\begin{bmatrix} 6 \\ 7 \end{bmatrix}, \begin{bmatrix} 5 \\ 9 \end{bmatrix}, \begin{bmatrix} 8 \\ 6 \end{bmatrix}, \begin{bmatrix} 4 \\ 9 \end{bmatrix}, \begin{bmatrix} 7 \\ 9 \end{bmatrix}$

处理 2:  $\begin{bmatrix} 3 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 6 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}$

处理 3:  $\begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 5 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}$

- (a) 像式 (6-39) 一样, 把观测值分解为总平均、处理效应和残差三个部分. 对每个变量构造相应的矩阵 (见例 6.9)
- (b) 利用 (a) 中的信息, 构造单因子 MANOVA 表
- (c) 计算威尔克斯  $\Lambda$  统计量  $\Lambda^*$ , 并利用表 6.3 检验处理效应 (取  $\alpha = 0$  再利用带巴特利特修正的  $\chi^2$  近似重新检验上述假设 [见式 (6-43)], 并比较这两个结论

解答.

对于变量 1:

$$\begin{aligned}
 & \begin{bmatrix} 6 & 5 & 8 & 4 & 7 \\ 3 & 1 & 2 & & \\ 2 & 5 & 3 & 2 & \end{bmatrix}_{\text{观测}} = \begin{bmatrix} 4 & 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & & \\ 4 & 4 & 4 & 4 & \end{bmatrix}_{\text{均值}} + \begin{bmatrix} 2 & 2 & 2 & 2 & 2 \\ -2 & -2 & -2 & & \\ -1 & -1 & -1 & -1 & \end{bmatrix}_{\text{处理}} \\
 & + \begin{bmatrix} 0 & -1 & 2 & -2 & 1 \\ 1 & -1 & 0 & & \\ -1 & 2 & 0 & -1 & \end{bmatrix}_{\text{残差}}, SS_{\text{观测}} = 246, SS_{\text{均值}} = 192, SS_{\text{处理}} = 36, SS_{\text{残差}} = 18
 \end{aligned}$$

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

对于变量 2:

$$\begin{aligned}
 & \begin{bmatrix} 7 & 9 & 6 & 9 & 9 \\ 3 & 6 & 3 & & \\ 3 & 1 & 1 & 3 & \end{bmatrix}_{\text{观测}} = \begin{bmatrix} 5 & 5 & 5 & 5 & 5 \\ 5 & 5 & 5 & & \\ 5 & 5 & 5 & 5 & \end{bmatrix}_{\text{均值}} + \begin{bmatrix} 3 & 3 & 3 & 3 & 3 \\ -1 & -1 & -1 & & \\ -3 & -3 & -3 & -3 & \end{bmatrix}_{\text{处理}} \\
 & + \begin{bmatrix} -1 & 1 & -2 & 1 & 1 \\ -1 & 2 & -1 & & \\ 1 & -1 & -1 & 1 & \end{bmatrix}_{\text{残差}}, SS_{\text{观测}} = 402, SS_{\text{均值}} = 300, SS_{\text{处理}} = 84, SS_{\text{残差}} = 18
 \end{aligned}$$

$$SS_{\text{修正后总和}} = SS_{\text{观测}} - SS_{\text{均值}} = 402 - 300 = 102$$

为完整地给出 MANOVA 表, 还必须考虑它们的交叉乘积, 将两变量中的数据按行相乘, 可得交叉乘积的贡献

$$\begin{aligned}
 SCP_{\text{总}} &= 6 \times 7 + 5 \times 9 + 8 \times 6 + 4 \times 9 + 7 \times 9 + 3 \times 3 + 1 \times 6 + 2 \times 3 \\
 &+ 2 \times 3 + 5 \times 1 + 3 \times 1 + 2 \times 3 = 275
 \end{aligned}$$

$$SCP_{\text{均值}} = 4 \times 5 + 4 \times 5 + \dots + 4 \times 5 = 12 \times 4 \times 5 = 240$$

$$SCP_{\text{处理}} = 5(2 \times 3) + 3(-2 \times -1) + 4(-1 \times -3) = 48$$

$$SCP_{\text{残差}} = 0 \times -1 + -1 \times 1 + 2 \times -2 + \dots - 1 \times 1 = -13$$

$$SCP_{\text{修正后}} = \text{总交叉乘积} - \text{均值交叉乘积} = 35$$

已知公式：

总体均值向量数目/组数  $g = 3$ , 矩阵元素个数  $\sum_{l=1}^g n_l = 5 + 3 + 4 = 12$

$$B = \begin{bmatrix} SS_{\text{变量 1 处理}} & SCP_{\text{处理}} \\ SCP_{\text{处理}} & SS_{\text{变量 2 处理}} \end{bmatrix}, W = \begin{bmatrix} SS_{\text{变量 1 残差}} & SCP_{\text{残差}} \\ SCP_{\text{残差}} & SS_{\text{变量 2 残差}} \end{bmatrix}$$

MANOVA 表如下：

变化来源	平方和与交叉乘积和矩阵 SSP	自由度 d.f.
处理	$B = \begin{bmatrix} 36 & 48 \\ 48 & 84 \end{bmatrix}$	$g - 1 = 2$
残差	$W = \begin{bmatrix} 18 & -13 \\ -13 & 18 \end{bmatrix}$	$\sum_{l=1}^g n_l - g = 9$
修正后总和	$B + W = \begin{bmatrix} 54 & 35 \\ 35 & 102 \end{bmatrix}$	$\sum_{l=1}^g n_l - 1 = 2 + 9 = 11$

$$\Lambda^* = \frac{|W|}{|B + W|} = \frac{18 \times 18 - (-13) \times (-13)}{54 \times 102 - 35 \times 35} = \frac{155}{4283} = 0.0362 \text{ 且 } p = 2, g = 3 \geq 2$$

$$\therefore \text{查表 6.3 知威尔克斯分布} \left( \frac{\sum n_l - g - 1}{g - 1} \right) \left( \frac{1 - \sqrt{\Lambda^*}}{\sqrt{\Lambda^*}} \right) \sim F_{2(g-1), 2(\sum n_l - g - 1)}$$

$$\therefore F_{4,16}(0.01) = 4.77 < 17.02 = \left( \frac{8}{2} \right) \left( \frac{1 - \sqrt{0.0362}}{\sqrt{0.0362}} \right)$$

$\therefore$  拒绝  $H_0$ , 处理间的差异在  $\alpha = 0.01$  的显著性水平下存在。

此外, 公式 (6-43) 利用带巴特利特修正的  $\chi^2$  近似检验

$$- \left( n - 1 - \frac{p + g}{2} \right) \ln \Lambda^* = - \left( 12 - 1 - \frac{2 + 3}{2} \right) \ln(0.0362) = 28.209$$

$$\text{公式 (6-44) 中 } \chi_{p(g-1)}^2(\alpha) = \chi_4^2(0.01) = 13.28 < 28.209$$

$\therefore$  拒绝  $H_0$ , 处理间的差异在  $\alpha = 0.01$  的显著性水平下存在。

**题目 4. 6.13.**

无重复的双因子 MANOVA:

考虑以下双因子表给出的 (注意: 对因子水平的每一种组合中仅有一个观测值向量) 响应  $x_1, x_2$  的观测值:

		因子 2			
		水平 1	水平 2	水平 3	水平 4
因子 1	水平 1	6	4	8	2
		8	6	12	6
	水平 2	3	-3	4	-4
		8	2	3	3
	水平 3	-3	-4	3	-4
		2	-5	-3	-6

无重复的双因子 MANOVA 模型为

$$\mathbf{X}_{lk} = \boldsymbol{\mu} + \boldsymbol{\tau}_l + \boldsymbol{\beta}_k + \mathbf{e}_{lk}; \quad \sum_{l=1}^g \boldsymbol{\tau}_l = \sum_{k=1}^b \boldsymbol{\beta}_k = \mathbf{0}$$

其中  $\mathbf{e}_{lk}$  为独立的  $N_p(\mathbf{0}, \boldsymbol{\Sigma})$  随机向量

(a) 与例 6.9 类似, 将两变量中的每一个变量的观测值分解为

$$x_{lk} = \bar{x} + (\bar{x}_{l.} - \bar{x}) + (\bar{x}_{.k} - \bar{x}) + (x_{lk} - \bar{x}_{l.} + \bar{x}_{.k} + \bar{x})$$

对每一个响应, 这个分解会产生几个  $3 \times 4$  矩阵, 这里  $\bar{x}$  为总平均,  $\bar{x}_{l.}$  为因子 1 第  $l$  个水平下的平均, 而  $\bar{x}_{.k}$  为因子 2 第  $k$  个水平下的平均.

(b) 矩阵的每一行看成一个串成的长向量, 并计算平方和与交叉乘积和矩阵

(c) 构造 MANOVA 表

(d) 检验因子 1 和因子 2 的主效应

解答.

对于第一个变量:

$$\begin{bmatrix} 6 & 4 & 8 & 2 \\ 3 & -3 & 4 & -4 \\ 3 & -3 & 4 & -4 \end{bmatrix}_{\text{观测}} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}_{\text{均值}} + \begin{bmatrix} 4 & 4 & 4 & 4 \\ -1 & -1 & -1 & -1 \\ -3 & -3 & -3 & -3 \end{bmatrix}_{\text{因子 1 效果}} \\ + \begin{bmatrix} 1 & -2 & 4 & -3 \\ 1 & -2 & 4 & -3 \\ 1 & -2 & 4 & -3 \end{bmatrix}_{\text{因子 2 效果}} + \begin{bmatrix} 0 & 1 & -1 & 0 \\ 2 & -1 & 0 & -1 \\ -2 & 0 & 1 & 1 \end{bmatrix}_{\text{残差}}$$

$$SS_{\text{总}} = 220, SS_{\text{均值}} = 12, SS_{\text{因子 1}} = 104, SS_{\text{因子 2}} = 90, SS_{\text{残差}} = 14$$

对于第二个变量:

$$\begin{bmatrix} 8 & 6 & 12 & 6 \\ 8 & 2 & 3 & 3 \\ 2 & -5 & -3 & -6 \end{bmatrix}_{\text{观测}} = \begin{bmatrix} 3 & 3 & 3 & 3 \\ 3 & 3 & 3 & 3 \\ 3 & 3 & 3 & 3 \end{bmatrix}_{\text{均值}} + \begin{bmatrix} 5 & 5 & 5 & 5 \\ 1 & 1 & 1 & 1 \\ -6 & -6 & -6 & -6 \end{bmatrix}_{\text{因子 1 效果}} \\ + \begin{bmatrix} 3 & -2 & 1 & -2 \\ 3 & -2 & 1 & -2 \\ 3 & -2 & 1 & -2 \end{bmatrix}_{\text{因子 2 效果}} + \begin{bmatrix} -3 & 0 & 3 & 0 \\ 1 & 0 & -2 & 1 \\ 2 & 0 & -1 & -1 \end{bmatrix}_{\text{残差}}$$

$$SS_{\text{总}} = 440, SS_{\text{均值}} = 108, SS_{\text{因子 1}} = 248, SS_{\text{因子 2}} = 54, SS_{\text{残差}} = 30$$

交叉相乘之和:

$$SCP_{\text{总}} = 6 \times 8 + 4 \times 6 + \dots + (-4) \times (-6) = 227$$

$$SCP_{\text{均值}} = 1 \times 3 + 1 \times 3 + \dots + 1 \times 3 = 12 \times 1 \times 3 = 36$$

$$SCP_{\text{因子 1 处理}} = 4 \times 5 + 4 \times 5 + \dots + (-3) \times (-6) = 148$$

$$SCP_{\text{因子 2 处理}} = 1 \times 3 + (-2) \times (-2) + \dots + (-3) \times (-2) = 51$$

$$SCP_{\text{残差}} = 0 \times (-3) + 1 \times 0 + \dots + 1 \times (-1) = -8$$

$$SCP_{\text{修正后}} = 227 - 36 = 191, \text{验证: } 227 = 36 + 148 + 51 - 8$$

已知公式：

总体均值向量数目/组数  $g = 3$ , 列数  $b = 4$ ,

$$B_1 = \begin{bmatrix} SS_{\text{变量 1 因子 1 处理}} & SCP_{\text{因子 1 处理}} \\ SCP_{\text{因子 1 处理}} & SS_{\text{变量 2 因子 1 处理}} \end{bmatrix}, W = \begin{bmatrix} SS_{\text{变量 1 残差}} & SCP_{\text{残差}} \\ SCP_{\text{残差}} & SS_{\text{变量 2 残差}} \end{bmatrix}$$

MANOVA 表如下：

变化来源	平方和与交叉乘积和矩阵 SSP	自由度 d.f.
因子 1 处理	$B_1 = \begin{bmatrix} 104 & 148 \\ 148 & 248 \end{bmatrix}$	$g - 1 = 2$
因子 2 处理	$B_2 = \begin{bmatrix} 90 & 51 \\ 51 & 54 \end{bmatrix}$	$b - 1 = 3$
残差	$W = \begin{bmatrix} 14 & -8 \\ -8 & 30 \end{bmatrix}$	$(g - 1)(b - 1) = 6$
修正后总和	$B_1 + B_2 + W = \begin{bmatrix} 208 & 191 \\ 191 & 332 \end{bmatrix}$	$gb - 1 = 11$

验证因子 1 和因子 2 的主效应：

$$- \left[ (g - 1)(b - 1) - \left( \frac{p + 1 - (b - 1)}{2} \right) \right] \ln \Lambda^* = - \left[ 6 - \left( \frac{2 + 1 - 2}{2} \right) \right] \ln \frac{|W|}{|B_2 + W|}$$

$$= -6 \ln \frac{356}{6887} = 17.77 > \chi_4^2(0.05) = 9.49$$

$\therefore$  结论：在  $\alpha = 0.05$  的显著性水平下拒绝原假设  $H_0 : \tau_1 = \tau_2 = \tau_3 = 0$

$$- \left[ (g - 1)(b - 1) - \left( \frac{p + 1 - (g - 1)}{2} \right) \right] \ln \Lambda^* = - \left[ 6 - \left( \frac{2 + 1 - (4 - 1)}{2} \right) \right] \ln \frac{|W|}{|B_1 + W|}$$

$$= -5.5 \ln \frac{356}{13204} = 19.87 > \chi_6^2(0.05) = 12.59$$

$\therefore$  结论：在  $\alpha = 0.05$  的显著性水平下拒绝原假设  $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$