資料分析方法-HW4

經研一 江彥亨 R11323040

1. Total sample variance :

$$TV(S_1) = tr(S_1) = 1+1+1 = 3 \#$$

$$TV(S_2) = tr(S_2) = 1+1+1 = 3 \#$$

Generalized sample variance :

$$GV(S_1) = |S_1| = 1 \times \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} + 0 \times \begin{vmatrix} 0 & 0 \\ 0 & 1 \end{vmatrix} + 0 \times \begin{vmatrix} 0 & 1 \\ 0 & 0 \end{vmatrix}$$

$$= 1 \#$$

$$GV(S_2) = |S_2| = 1 \times \begin{vmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{vmatrix} - (-\tfrac{1}{2}) \begin{vmatrix} -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{vmatrix} + (-\tfrac{1}{2}) \begin{vmatrix} -\frac{1}{2} & 1 \\ -\frac{1}{2} & -\frac{1}{2} \end{vmatrix}$$

$$= \frac{3}{4} - \frac{3}{8} - \frac{3}{8} = 0 \#$$

2.

$$|R| = \left| \begin{bmatrix} \frac{1}{\sqrt{S_{11}}} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sqrt{S_{pp}}} \end{bmatrix} S \begin{bmatrix} \frac{1}{\sqrt{S_{11}}} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sqrt{S_{pp}}} \end{bmatrix} \right|$$

$$= \left| \begin{bmatrix} \frac{1}{\sqrt{S_{11}}} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sqrt{S_{pp}}} \end{bmatrix} \right| |S| \left| \begin{bmatrix} \frac{1}{\sqrt{S_{11}}} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sqrt{S_{pp}}} \end{bmatrix} \right|$$

$$= \prod_{i=1}^{p} (S_{ii})^{-\frac{1}{2}} \, |S| \, \prod_{i=1}^{p} (S_{ii})^{-\frac{1}{2}} = |S| \prod_{i=1}^{p} (S_{ii})^{-1}$$

$$\Rightarrow |S| = |R| \prod_{i=1}^{p} S_{ii} \#$$

3. a.

$$\bar{Y}_1 = \bar{X}_1 + \bar{X}_2 + \bar{X}_3 + \bar{X}_4 = 0.766 + 0.508 + 0.438 + 0.161 = 1.873 \#$$

$$\widehat{Var}(Y_1) = \widehat{Var}(X_1 + X_2 + X_3 + X_4) = \widehat{Var}(X_1) + \widehat{Var}(X_2) + \widehat{Var}(X_3)$$

$$+ \widehat{Var}(X_4) + 2 \times [\widehat{Cov}(X_1, X_2) + \widehat{Cov}(X_1, X_3) + \widehat{Cov}(X_1, X_4)$$

$$+ \widehat{Cov}(X_2, X_3) + \widehat{Cov}(X_2, X_4) + \widehat{Cov}(X_3, X_4)]$$

$$= 0.856 + 0.568 + 0.171 + 0.043 + 2 [0.635 + 0.173 + 0.096$$

$$+ 0.128 + 0.067 + 0.039]$$

$$= 3.914 \#$$

3.b $\bar{Y}_2 = \bar{X}_1 - \bar{X}_2 = 0.766 - 0.508 = 0.258$ #

$\widehat{Var}(Y_2) = \widehat{Var}(X_1 - X_2) = \widehat{Var}(X_1) + \widehat{Var}(X_2) - 2\widehat{Cov}(X_1, X_2)$

$= 0.856 + 0.568 - 2 \times 0.635 = 0.154$ #

3.c $\widehat{Cov}(Y_1, Y_2) = \widehat{Cov}(X_1+X_2+X_3+X_4,\ X_1 - X_2)$

$= \widehat{Cov}(X_1, X_1) + \widehat{Cov}(X_2, X_1) + \widehat{Cov}(X_3, X_1) + \widehat{Cov}(X_4, X_1)$

$- \widehat{Cov}(X_1, X_2) - \widehat{Cov}(X_2, X_2) - \widehat{Cov}(X_3, X_2) - \widehat{Cov}(X_4, X_2)$

$= 0.856 + 0.173 + 0.096 - 0.568 - 0.128 - 0.067$

$= 0.362$ #

4.a $\bar{X} = \begin{bmatrix} 6 \\ 10 \end{bmatrix} \Rightarrow S = \frac{1}{4-1}\begin{bmatrix} 2 & 8 & 6 & 8 \\ 12 & 9 & 9 & 10 \end{bmatrix}\left(\begin{bmatrix} 2 & 12 \\ 8 & 9 \\ 6 & 9 \\ 8 & 10 \end{bmatrix} - \begin{bmatrix} 6 & 10 \\ 6 & 10 \\ 6 & 10 \\ 6 & 10 \end{bmatrix}\right)$

$= \begin{bmatrix} 24 & -10 \\ -10 & 6 \end{bmatrix}$

$T^2 = 4\left(\begin{bmatrix} 6 \\ 10 \end{bmatrix} - \begin{bmatrix} 7 \\ 11 \end{bmatrix}\right)^T S^{-1}\left(\begin{bmatrix} 6 \\ 10 \end{bmatrix} - \begin{bmatrix} 7 \\ 11 \end{bmatrix}\right)$

$= 4\begin{bmatrix} -1 & -1 \end{bmatrix}\begin{bmatrix} \frac{3}{22} & \frac{5}{22} \\ \frac{5}{22} & \frac{6}{11} \end{bmatrix}\begin{bmatrix} -1 \\ -1 \end{bmatrix} = \frac{50}{11} \simeq 4.5455$

4.b $T^2 \sim T^2_{2,4-1} = \frac{2(4-1)}{4-2} F_{2,4-2} = 3 \cdot F_{2,2}$

4.c Given $\alpha = 0.05 \Rightarrow 3 \cdot F_{0.05,2,2} = 3 \times 19 = 57$

$T^2 \simeq 4.5455 \leq 57$

$\Rightarrow$ fail to reject $H_0 : \mu^T = [7, 11]$. We don't have enough information to reject H_0.

5. $H_0 : \mu_{male} = \mu_{female}$

$H_1 : \mu_{male} \neq \mu_{female}$

$\Psi = (\mu_{male} - \mu_{female})^T \left(\frac{1}{24} S_{male} + \frac{1}{24} S_{female}\right)^{-1} (\mu_{male} - \mu_{female})$

$\sim T^2_{3,23} \equiv \frac{23}{7} \times F_{3,21}$

$\alpha = 0.05$

$RR = \{\Psi \mid \Psi > \frac{23}{7} \times F_{0.05,3,21} \simeq 10.094\}$

$\Psi^* = 24 \times \begin{bmatrix} 22.6667 \\ 14.2917 \\ 11.3333 \end{bmatrix}^T \begin{bmatrix} 299.7222 & 183.8351 & 61.783 \\ 183.8351 & 116.6198 & 138.2882 \\ 61.783 & 138.2882 & 51.25 \end{bmatrix}^{-1} \begin{bmatrix} 22.6667 \\ 14.2917 \\ 11.3333 \end{bmatrix}$

$\simeq 41.9695$

$\Psi^* \simeq 41.9695 > 10.094 \Rightarrow \Psi^* \in RR$

$\Rightarrow$ reject H_0, $\mu_{male} \neq \mu_{female}$

6.a

```
          Df  Pillai approx F num Df den Df  Pr(>F)
species    2 1.67758   5.2031      4      4 0.06959 .
nutrient   1 0.68401   1.0823      2      1 0.56213
Residuals  2
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Based on the results, given a significance level of 0.05, species effect and nutrient effect are both insignificant. Therefore, we can conclude that neither species nor nutrient have significant effect on 560 cell means and 720 cell means.

6.b

```
 Response cm560 :
          Df Sum Sq Mean Sq F value  Pr(>F)
Species    2 47.476 23.7382 10.0551 0.09046 .
Nutrient   1  8.260  8.2603  3.4989 0.20232
Residuals  2  4.722  2.3608
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

 Response cm720 :
          Df  Sum Sq Mean Sq F value  Pr(>F)
Species    2 262.239 131.119 28.8196 0.03353 *
Nutrient   1   4.489   4.489  0.9867 0.42522
Residuals  2   9.099   4.550
---
```

Based on the results, given a significance level of 0.05, in 560 cells mean, both species effect and nutrient effect are both insignificant. However, in 720 cells mean, species effect is significant but nutrient effect is not significant. The reason why species effect is significant in 720 mean cells under the two-way ANOVA probably may be that two-way ANOVA does not take into account the covariance between 560 cells mean and 720 cells mean.