MAS programmes - Statistical Analysis (Autumn Term)

Exercises 8 - SOLUTIONS

1. From the question, $C_f = \dfrac{3799^2}{3 \times 3 \times 4} = 400900.0278$, and

$$SS_T = \sum_i \sum_j \sum_k y_{ijk}^2 - C_f = 478547 - 400900.0278 = 77646.9722$$

Using the *hint*, (the cell totals are $\sum_i \sum_j y_{ij.}$) we have

		Temperature 1	**2**	**3**	**Total**
Material Type	M_1	539	229	230	998
	M_2	623	479	198	1300
	M_3	576	583	342	1501
	Total	1738	1291	770	3799

We obtain (using the row and column totals respectively),

$$\begin{aligned} SS_A &= \frac{1}{3 \times 4} \sum_i y_{i..}^2 - C_f = \frac{1}{12}\{998^2 + 1300^2 + 1501^2\} - C_f \\ &= 411583.75 - 400900.0278 = 10683.7222 \end{aligned}$$

$$\begin{aligned} SS_B &= \frac{1}{3 \times 4} \sum_j y_{.j.}^2 - C_f = \frac{1}{12}\{1738^2 + 1291^2 + 770^2\} - C_f \\ &= 440018.75 - 400900.0278 = 39118.7222 \end{aligned}$$

Also, (from the cell totals)

$$\begin{aligned} SS_{Subtotals} &= \frac{1}{4} \sum_i \sum_j y_{ij.}^2 - C_f \\ &= \frac{1}{4}\{539^2 + 229^2 + 230^2 + 623^2 + 479^2 + 198^2 + 576^2 + 583^2 + 342^2\} - C_f \\ &= 460316.25 - 400900.0278 = 59416.2222 \end{aligned}$$

so that

$$SS_{AB} = SS_{Subtotals} - SS_A - SS_B = 59416.2222 - 10683.7222 - 39118.7222 = 9613.7778$$

and finally

$$SS_R = SS_T - SS_{Subtotals} = 77646.9722 - 59416.2222 = 18230.75$$

These values match those given in the output on p7 of the Lecture Notes.

[Check: $SS_T = SS_A + SS_B + SS_{AB} + SS_R$].

2.

(a)

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk}$$

where y_{ijk} is the k-th replicate taken under the j-th level of calcium and the i-th level of sodium,
μ is the 'overall' mean,
$\{\tau_i\}$ are the effects of the different levels of Sodium, with $\sum_{i=1}^{a} \tau_i = 0$,
$\{\beta_j\}$ are the effects of the different levels of Calcium, with $\sum_{j=1}^{b} \beta_j = 0$,
$\{(\tau\beta)_{ij}\}$ are the interaction terms between the i-th level of sodium and j-th level of calcium, with
$\sum_{j=1}^{b}(\tau\beta)_{ij} = 0$, $i=1,\ldots,a$,
$\sum_{i=1}^{a}(\tau\beta)_{ij} = 0$, $j=1,\ldots,b$, and
$\{\epsilon_{ijk}\} \sim \mathrm{NID}(0, \sigma^2)$.

(b)
```
> growth <- c(107, 101, 97, 92, 104, 103, 92, 91, 100, 92, 88, 95,
97, 92, 91,
    85, 103, 101, 97, 79, 92, 88, 81, 81, 92, 75, 72, 67, 97, 81, 61, 57,
    89, 85, 66, 53)
> s <- rep(1:3, rep(12, 3))
> sodium <- factor(s)
> c <- rep(1:4, 9)
> calcium <- factor(c)
> plants <- data.frame(sodium, calcium, growth)
> plants.aov <- aov(growth ~ sodium + calcium + sodium:calcium, data = plants)
> summary(plants.aov)
               Df Sum Sq Mean Sq F value   Pr(>F)
sodium          2 3160.5  1580.2  55.018 1.09e-09 ***
calcium         3 2125.1   708.4  24.663 1.65e-07 ***
sodium:calcium  6  557.1    92.8   3.232    0.018 *
Residuals      24  689.3    28.7
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

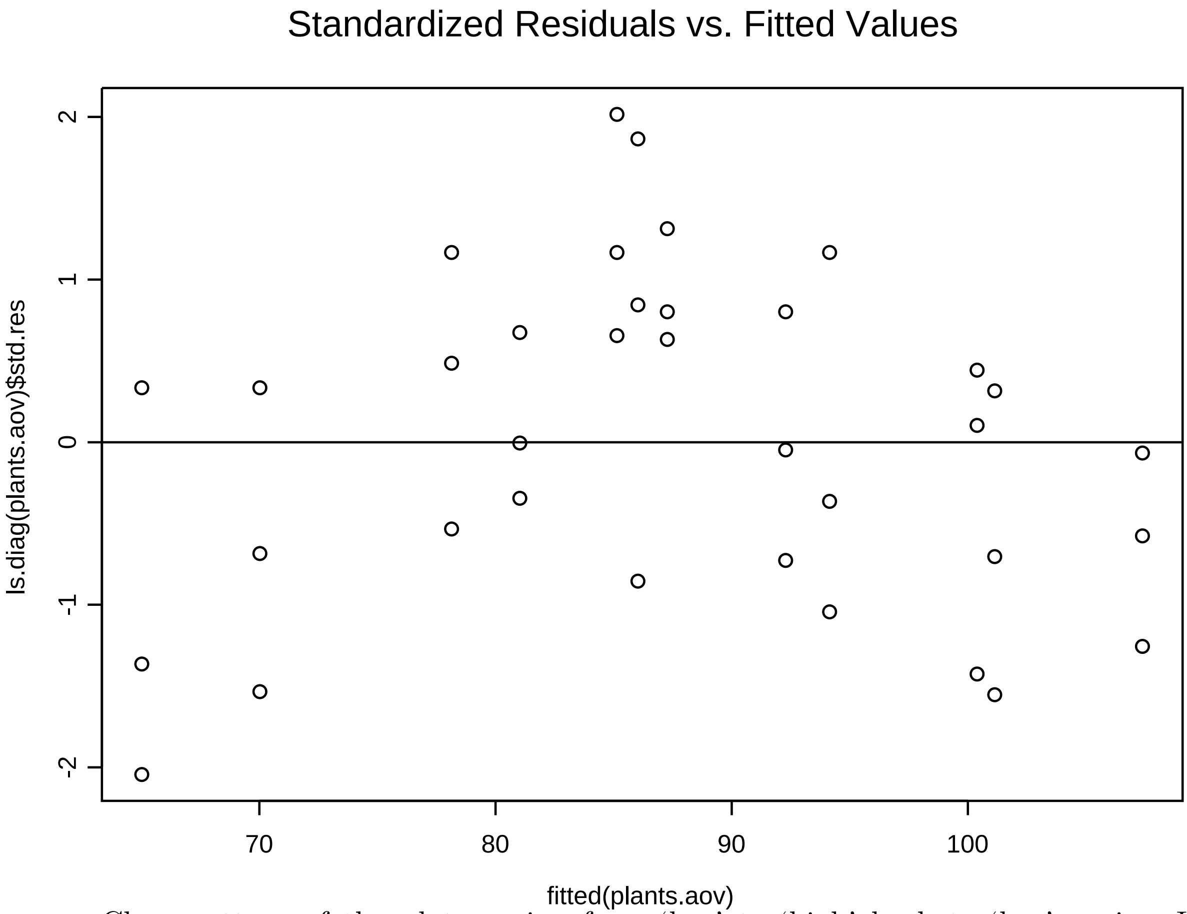
There are significant interaction effects, where we have an F-ratio of 3.23243 with a p-value 0.01797302 on the $F_{6,24}$ distribution. There is also significant variation according to the levels of sodium and of calcium, as evidenced by the significantly large F-ratios of 55.01838 and 24.66280 on the $F_{2,24}$ and $F_{3,24}$ distributions; indeed the corresponding p-values are significantly less than 0.01.

(c)
```
> plants.aov <- aov(growth ~ sodium + calcium, data = plants)
> summary(plants.aov)
           Df Sum Sq Mean Sq F value   Pr(>F)
sodium      2   3160  1580.2   38.04 5.93e-09 ***
calcium     3   2125   708.4   17.05 1.19e-06 ***
Residuals  30   1246    41.5
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

Without the interaction term, the main effects for the levels of sodium and calcium are still significant.

(d) Consider the plot of the (standardized) residuals against the fitted values in the 'no-interaction model'.

```
> plot(fitted(plants.aov), ls.diag(plants.aov)$std.res,
main = "Standardized Residuals vs. Fitted Values")
> abline(h=0)
```



Clear pattern of the plot moving from ‘low’ to ‘high’ back to ‘low’ again. It seems, therefore, that the no-interaction model should **not** be endorsed. Nothing untoward to report about the plots for the model that includes interaction. (Check).