M3S12 BIOSTATISTICS: ASSESSED COURSEWORK 1
SOLUTIONS

1. Using the code given (or by hand), the ANOVA table is

Source	D.F.	Sum of squares	Mean square	F	$p-$value
TYPE	4	1480.823	370.206	40.885	6.74×10^{-8}
Residual	15	135.822	9.055		
Total	19	1616.645			

where

$$\mathbf{TSS} = \sum_{k=1}^{K}\sum_{j=1}^{n_k}(y_{kj}-\overline{y}_{..})^2 = \sum_{k=1}^{5}\sum_{j=1}^{4}(y_{kj}-\overline{y}_{..})^2 = 1616.645$$

$$\mathbf{RSS} = \sum_{k=1}^{K}\sum_{j=1}^{n_k}(y_{kj}-\overline{y}_{k})^2 = \sum_{k=1}^{5}\sum_{j=1}^{4}(y_{kj}-\overline{y}_{k})^2 = 135.822$$

$$\mathbf{FSS} = \sum_{k=1}^{K}n_k(\overline{y}_{k}-\overline{y}_{..})^2 = \sum_{k=1}^{5}4(\overline{y}_{k}-\overline{y}_{..})^2 = 1480.823$$

The test statistic is

$$F = \frac{FSS/FDF}{RSS/RDF} = \frac{1480.823/4}{135.822/15} = 40.885.$$

Under the null hypothesis

$$H_0 : \mu_1 = ... = \mu_5$$

it follows that

$$F \sim Fisher\,(4, 15)$$

and thus the p-value is given by

$$p = P\left[F > 40.885 | H_0 \text{ is true}\right] = 1 - Fisher_{4,15}^{-1}(40.885)$$

which is computed using the SPLUS command `1-pf(40.885,4,15)`. The computed value, 6.74×10^{-8} is smaller than the conventionally chosen significance levels $\alpha = 0.05$ or 0.01, and so the null hypothesis is rejected; in fact the 0.95 critical value is C_R such that

$$P\left[F > C_R | H_0 \text{ is true}\right] = 0.05 \therefore C_R = \texttt{qf(0.95,4,15)} = 3.056$$

[8 MARKS]

2. The full ANOVA-table

Source	D.F.	Sum of squares	Mean square	F	p
DRUG DOSE	[1]	66.533	[6]	[9]	[12]
THERAPY	[2]	[4]	4.950	[10]	[13]
INTERACTION	6	156.800	[7]	[11]	[14]
Residual	[3]	518.983	[8]		
Total	59	[5]			

where we have $K = 3$ drug treatments and $L = 4$ therapies. Thus

$$\boxed{1} = K - 1 = 2 \qquad \boxed{2} = L - 1 = 3 \qquad \boxed{3} = (n-1) - (K-1) - (L-1) - 6 = 48$$

$$\boxed{4} = (L-1) \times 4.950 = 14.85 \qquad \boxed{5} = (66.533 + 14.85 + 156.800 + 518.893) = 757.166$$

and

$$\boxed{6} = \frac{66.533}{\boxed{1}} = \frac{66.533}{2} = 33.3277 \ , \ \boxed{7} = \frac{156.800}{6} = 26.133 \ , \ \boxed{8} = \frac{518.983}{\boxed{3}} = \frac{518.983}{48} = 10.812$$

For the F statistics,

$$\boxed{9} = \frac{\boxed{6}}{\boxed{8}} = \frac{33.327}{10.812} = 3.076 \ , \ \boxed{10} = \frac{4.950}{\boxed{8}} = \frac{4.950}{10.812} = 0.458 \ , \ \boxed{11} = \frac{\boxed{7}}{\boxed{8}} = \frac{26.133}{10.812} = 2.417$$

and the p-values are obtained by comparing the F statistics with the relevant F distribution:

$$\boxed{12} \ : \ \text{Compare } \boxed{9} \text{ with } Fisher\,(2,48) - p = 0.055 \ \ (C_R = 3.191)$$

$$\boxed{13} \ : \ \text{Compare } \boxed{10} \text{ with } Fisher\,(3,48) - p = 0.713 \ \ (C_R = 2.798)$$

$$\boxed{14} \ : \ \text{Compare with } \boxed{11} \ Fisher\,(6,48) - p = 0.040 \ \ (C_R = 2.295)$$

Thus we conclude that the null hypotheses for "no drug amount effect" and "no interaction" can be **rejected**, but that the null hypothesis for "no therapy effect" **cannot be rejected**.

The assumptions behind ANOVA require normality (here that is not verifiable), independence (plausible) and equal variance in each cross category. The feature of most concern is the apparently different variances in some of the cross-classifications; for example, see the **Absent/Nondir** and the **Low/Bmod** combinations.

[12 MARKS]

NOTE: The ANOVA results indicate only the interaction terms give a test statistic that indicates significance at the $\alpha = 0.05$ significance level. However, the "interaction only" model is not easy to interpret; the basic interaction hypothesis says that the mean in the $(i,j)^{th}$ cross-category, μ_{ij}, cannot be readily decomposed, and that each cross-category has a different mean. Thus, the interaction model has precisely $KL = 12$ parameters, and hence this is the same model as the full main effects plus interaction model. It is impossible in the "interaction only" model to assess the influence of the main effects. Thus, bearing in mind such potential "marginalization" problems, the final preferred fitted model would be the interaction model

$$DRUG * THERAPY = DRUG + THERAPY + DRUG \,.\, THERAPY.$$

Also, the hypothesis

$$H_0 : \text{All variances equal}$$

can in fact be formally tested using **Levene's test.**