Solutions to T7

Question 1

The data below is about a new analgesic drug and it is compared with aspirin and placebo for treatment of a simple headache. The measurements refer to the number of hours a patient is free from pain after taking the drug. In this small pilot study, two patients are given placebo, four are given the new drug and three are given aspirin.

Placebo	0.0	1.0		
New drug	2.3	3.5	2.8	2.5
Aspirin	3.1	2.7	3.8	

(i) Write down a model for the above study and specify the model assumptions.

Solution The model is as follows

$$y_{ij} = \theta_i + \varepsilon_{ij},\ \ i = 1, 2, 3,\ \ j = 1, \cdots, n_i.$$

From data we see that

$$k = 3,\ \ n_1 = 2, n_2 = 4, n_3 = 3, n = 9.$$

(ii) Construct an ANOVA table.

Solution From data we have

$$y_{1\cdot} = 1,\ \ y_{2\cdot} = 2.3 + 3.5 + 2.8 + 2.5 = 11.1,\ \ y_{3\cdot} = 3.1 + 2.7 + 3.8 = 9.6$$

$$y_{\cdot\cdot} = 1 + 11.1 + 9.6 = 21.7.$$

It follows that

$$SST = \sum_i y_{i\cdot}^2/n_i - y_{\cdot\cdot}^2/n = 62.0225 - 52.32111 = 9.7$$

and

$$S_{yy} = \sum_{i=1}^{k}\sum_{j=1}^{n_i} y_{ij}^2 - y_{..}^2/n = 63.97 - 52.32111 = 11.65.$$

Summarizing the above we obtain the ANOVA table

Source	SS	df	MS	F
Between groups	9.7	2	4.85	14.92
Within groups	1.95	6	0.325	

(iii) Is there any difference among the three drug effects at the 5% level of significance ?

Solution Note that $14.92 > F_{2,6}^{0.05} = 5.14$. Therefore we reject H_0, which implies there is difference among the drug effects at the 5% level of significance.

(iv) Suppose that one is interested in the comparison of the new drug with the average of the other two. Find the 95% confidence interval for this contrast.

Solution We estimate the contrast

$$\theta_2 - (\frac{\theta_1 + \theta_3}{2}).$$

The estimator is

$$\bar{y}_{2\cdot} - \frac{\bar{y}_{1\cdot} + \bar{y}_{3\cdot}}{2} = 2.775 - (0.5 + 3.2)/2 = 0.925.$$

It follows that the confidence interval is

$$0.925 \pm 2.47\sqrt{(0.325)(\frac{1/4}{2} + \frac{1}{4} + + \frac{1/4}{3})} = (-0.028, 1.88),$$

where $t_6^{0.025} = 2.47$.

Question 2

We consider one-way classification model to analyze a data set. The summary of the data set is as follows,

Level	1	2	3
Summary	(4,28,10)	(4,29,6)	(5,40.4,83.2)

For each level, we calculate the sample size n_i, sample mean $\bar{y}_{i.}$ and the SS $\sum(y_{ij}-\bar{y}_{i.})^2$. Put them into $(n_i, \bar{y}_{i.}, \sum(y_{ij} - \bar{y}_{i.})^2)$, and then form the above table. Please derive the ANOVA table and perform the F test to check whether this factor can affect the response.

Solution to Q 2

Obviously, $k = 3, n = 4 + 4 + 5 = 13$. From the table we also obtain

$$SSE = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2 = 10 + 6 + 83.2 = 99.2,$$

$$\bar{y}_{..} = \frac{n_1\bar{y}_{1.} + n_2\bar{y}_{2.} + n_3\bar{y}_{3.}}{n_1 + n_2 + n_3} = \frac{4 \times 28 + 4 \times 29 + 5 \times 40.4}{4 + 4 + 5} = 33.08.$$

$$SST = \sum_{j=1}^{3} n_i(\bar{y}_{i.} - \bar{y}_{..})^2 = 4(28 - 33.08)^2 + 4(29 - 33.08)^2 + 5(40.4 - 33.08)^2 = 437.7.$$

Thus the ANOVA table is below:

Source	SS	df	MS	F
Between groups	437.7	2	218.85	22.06
Within groups	99.2	10	9.92	
Total	536.9	21		

Since $22.06 > F_{2,10}^{0.95} = 4.10$ we conclude that the factor can affect the response.