# ST4241: Design and Analysis of Clinical Trials

2009/2010: Semester I

Tutorial 11

1. The linear model for the balanced incomplete block design with design parameters $g, k, r$ and $\lambda$ is as follows:

$$X_{ij} = \mu + s_i + \alpha_j + \epsilon_{ij},$$

where $\sum_j \alpha_j = 0$, $s_i$'s are i.i.d. random variables with mean zero and variance $\sigma_s^2$, $\epsilon_{ij}$'s are i.i.d. random errors with mean zero and variance $\sigma_e^2$ and are independent of $s_i$'s. Let $a_j$ be the unbiased estimate of $\alpha_j$ as defined in the lecture notes.

**(i)** Show that the $a_j$'s are least squares estimates of $\alpha_j$'s; that is they are minimizers of

$$\sum_i \sum_j (X_{ij} - \mu - s_i - \alpha_j)^2, \text{ subject to } \sum_j \alpha_j = 0.$$

**(ii)** Show that

$$\text{Var}(a_j) = \frac{\sigma_e^2}{r\text{EFF}} \frac{g-1}{g}, \quad \text{Cov}(a_j, a_k) = -\frac{\sigma_e^2}{r\text{EFF}} \frac{1}{g},$$

where $\text{EFF} = \frac{g(k-1)}{k(g-1)}$.

**(iii)** Show that, for any contrast $\sum_{j=1}^{g} c_j a_j$,

$$\text{Var}(\sum_{j=1}^{g} c_j a_j) = \frac{\sigma_e^2}{r\text{EFF}} \sum_j c_j^2.$$

2. The data from the interexaminer reliability study considered in the lecture notes is given on the next page.

**(i)** Using the linear model approach, derive the $F$ ratio for testing the significance of the differences among the examiners.

| Patient | Examiner 1 | 2 | 3 | 4 | 5 | 6 | Mean |
|---|---|---|---|---|---|---|---|
| 1 | 10 | 14 | 10 | | | | 11.33 |
| 2 | 3 | 3 | | 1 | | | 2.33 |
| 3 | 7 | | 12 | | | 9 | 9.33 |
| 4 | 3 | | | 8 | 5 | | 5.33 |
| 5 | 20 | | | | 26 | 20 | 22.00 |
| 6 | | 20 | 14 | | 20 | | 18.00 |
| 7 | | 5 | | 8 | | 14 | 9.00 |
| 8 | | 14 | | | 18 | 15 | 15.67 |
| 9 | | | 12 | 17 | 12 | | 13.67 |
| 10 | | 18 | 19 | | | 13 | 16.67 |
| Mean | 8.6 | 11.2 | 13.2 | 10.6 | 16.2 | 14.2 | 12.33 |

**(ii)** By looking at the data, it seems that Examiner 5 and 6 have higher mean scores than the other examiners. One would like to see whether the following contrasts are significant:

$$\frac{\alpha_1+\alpha_2+\alpha_3+\alpha_4}{4} - \frac{\alpha_5+\alpha_6}{2},$$
$$\frac{\alpha_1+\alpha_2+\alpha_3+\alpha_4}{4} - \alpha_5,$$
$$\frac{\alpha_1+\alpha_2+\alpha_3+\alpha_4}{4} - \alpha_6.$$

Using an appropriate multiple comparison criterion, test the significance of the above contrasts by controlling the overall error rate at 0.05.

3. The data form the study comparing four formulations considered in the lecture notes is given at the end of this question. Using the R code provided in the lecture notes, fit a linear model to the data. Using the information in the fitted object, do the following:

**(i)** By computing the means for each formulation, it is obvious form that the estimated effect of Formulation 3 is different from all the others. Check that the value of the contrast $C = a_3 - (a_1 + a_2 + a_4)/3$ is $-0.6489$ with an estimated standard error 0.0883. Test the significance of the contrast at level 0.05 using Scheffe's criterion. (Why should Scheffe's criterion be used?)

**(ii)** Check that, of the six pairwise differences $a_j - a_{j'}$, only the three differences involving Formulation 3 are statistically significant by the Tukey criterion.

| | Formulation | | | |
|---|---|---|---|---|
| Patient | 1 | 2 | 3 | 4 |
| 1 | -1.0894 (A) | -1.3200 (B) | | |
| 2 | | | -1.7577 (B) | -0.9817 (A) |
| 3 | -1.0771 (B) | | -1.7531 (A) | |
| 4 | | -0.9381 (A) | -1.6769 (B) | |
| 5 | -1.2044 (B) | | | -0.7795 (A) |
| 6 | | -1.0395 (B) | | -1.0426 (A) |
| 7 | -1.0991 (B) | -0.8092 (A) | | |
| 8 | | | -2.0245 (A) | -1.3374 (B) |
| 9 | -0.9846 (A) | | -1.4712 (B) | |
| 10 | | -1.1395 (B) | -1.6683 (A) | |
| 11 | -0.8069 (A) | | | -1.1913 (B) |
| 12 | | -0.7789 (A) | | -1.1694 (B) |