QUEEN MARY, UNIVERSITY OF LONDON

MAS 314 Design of Experiments

Practical 5 7–13 February 2007

This practical covers the analysis of designed experiments when the plots are grouped into blocks, the treatments are unstructured and the design is an orthogonal block design.

1 (A complete-block design: analysis with blocks as fixed effects) An experiment was conducted in Mexico to compare seven varieties of rubber tree. The experimenter knew that the ground was quite variable, so he divided it up into five blocks. Each block was divided into seven plots. He chose his blocks so that the seven plots in each block would be as alike as possible. He then planted each variety of rubber tree on one of the plots in each block.

We are going to analyse the rubber tree data in the file `mexico.dat`. This contains 2 columns of 35 values. The first column is variety, coded 1–7. The second column is the yield of rubber in grams. All the plots for Block 1 come first, followed by by all the plots for Block 2, and so on. We are going to analyse this data using both the fixed-effects and the random-effects models, and compare the results.

Load the data into Genstat, calling the columns `variety` and `yield`. Then create the appropriate factor called `mexblock`.

In the anova **Dialogue Box**, the space

Treatment Structure: [ ]

is for declaring the expectation model. For the model with blocks as fixed effects, this should be $V_B + V_T$, so put

Treatment Structure: `mexblock + variety`

and analyse the data.

2 (Residuals and outliers) You should find that Genstat has given you a warning that two of the experimental units have large residuals. Identify these experimental units and see what the value of `yield` is there. Could there be a mistake in the data?

We can verify these warnings by looking at an appropriate graph of residuals. Before closing the anova Dialogue Box, choose Further Output... then Residual Plots... . Tick Half Normal and then click on OK. You should see a graph: what does it tell you?

Try replacing one or both of the outliers by a missing value. You can do this by editing the spreadsheet and putting $*$ in place of the outlier. Then analyse the data again. Do you get a very different result from before?

3 (A complete-block design: analysis with blocks as random effects) First restore all the original data values.

In the anova **Dialogue Box**, the space

Block Structure: [ ]

is for declaring the covariance part of the model. For the random-effects model, choose/type

Treatment Structure: `variety`
Block Structure: `mexblock`

and analyse the yield again. Compare the output from the two analyses.

4 (Alternative ways of giving the Block Structure) What does **Block Structure** mean? It is for telling Genstat what the strata are. Genstat always thinks that V_0 is a stratum, which it silently ignores. Here we have told Genstat that the **Block Structure** is `mexblock`, so that Genstat knows that another stratum is $W_{\text{mexblock}} = V_{\text{mexblock}} \cap V_0^{\perp}$. This does not use up the whole space of dimension 35, so Genstat makes a further stratum $V_{\text{mexblock}}^{\perp}$. What is its dimension? What does Genstat call it?

If you don't like leaving Genstat to make up its own name for the last stratum (also called the *bottom* stratum), you can add a column to the spreadsheet containing a factor `plot` numbered from 1 to 35. Now do the analysis again with

Treatment Structure:	`variety`
Block Structure:	`mexblock/plot`

Compare this with the previous analysis.

Finally, add another column `expunit`, which is a factor numbered from 1 to 7 in order within each block. Redo the analysis using `mexblock/expunit` as the **Block Structure**. You should find that it is the same as the previous two analyses apart from the name of the last stratum.

What is happening is that Genstat interprets `mexblock/expunit` to mean that experimental units are numbered *within* each block. In other words, there is no connection between experimental unit 1 in the first block and experimental unit 1 in the second block. So Genstat first creates the factor `mexblock` $\wedge$ `expunit`. This is the same as the factor `plot`. Genstat doesn't use the symbol $\wedge$, so it calls this factor `mexblock.expunit`. Then it knows that the three strata are, in order:

(a) V_0, which it ignores;

(b) $W_{\text{mexblock}} = V_{\text{mexblock}} \cap V_0^{\perp}$, which it calls `mexblock`;

(c) $V_{\texttt{mexblock}\wedge\texttt{expunit}} \cap V_{\text{mexblock}}^{\perp}$, which in this case is the same as $V_{\text{mexblock}}^{\perp}$ because $V_{\texttt{mexblock}\wedge\texttt{expunit}}$ is the whole space.

In fact, it then tries to create a fourth stratum $V_{\texttt{mexblock}\wedge\texttt{expunit}}^{\perp}$, but finds that it has dimension 0.

Even though the factors `plot` and `expunit` are different (they have different numbers of levels), the factors `mexblock` $\wedge$ `plot` and `mexblock` $\wedge$ `expunit` are the same. So Steps 1–3 above give exactly the same strata as we had before. That is why you can use either `mexblock/plot` or `mexblock/expunit` in the **Block Structure**.