QUEEN MARY, UNIVERSITY OF LONDON

MAS 314 Design of Experiments

Practical 6 7–15 February 2007

This practical shows how to deal with simple factorial treatments. It also continues the investigation of block designs.

1 (Two treatment factors: no blocks) Read in the chick-feeding data from Practical 2 into a spreadsheet. Create the factors `protein` and `fishmeal` defined in lectures. Choose

General Analysis of Variance

in the **Analysis of Variance** menu. To fit the factors `protein` and `fishmeal` and the four-level factor `protein` ∧ `fishmeal`, declare

Treatment Structure: `protein*fishmeal`

The plots are unstructured, so leave the **Block Structure** box empty. Analyse the data and compare the output to the calculations in the lecture notes.

What labels does Genstat give to the lines for $W_{\texttt{protein}}$, $W_{\texttt{fishmeal}}$, and $W_{\texttt{protein}\wedge\texttt{fishmeal}}$?

2 (Some options in anova) The output from an analysis of variance can be changed by ticking and unticking boxes which become accessible when you click the Options ... box in the analysis-of-variance menu. Try ticking ☐ **Effects**. What is the extra output?

Try ticking ☐ **Mean plots**. What is the extra output?

3 (Analysis of Variance menu) You can fit the same model by choosing either

Completely Randomized Design

or

Two-way ANOVA (no Blocking)

in the **Analysis of Variance** menu. Try them both.

You should find that you can achieve the same output, but that you have to fill in the boxes rather differently. My advice is that it is simplest to *always* choose

General Analysis of Variance.

Then you consistently put the expectation part of the model into the **Treatment Structure** box, and the variance-covariance part into the **Block Structure** box.

From now on, these instructions assume that you always choose

General Analysis of Variance.

4 (Syntax of $*$) Genstat interprets the $*$ to mean two things. The first is that `protein` and `fishmeal` each make sense without the other. In other words, the question of whether or not the protein is groundnut or soya makes sense even if the level of `fishmeal` is different. Similarly, the question of whether or not fishmeal is added makes sense even if the level of `protein` is different.

The second part of the interpretation is that we want to know about the factor $\texttt{protein} \wedge \texttt{fishmeal}$ as well as the factors `protein` and `fishmeal`. When this is part of **Treatment Structure**, this gives the expectation part of the model. Now Genstat fits the expectation part of the model successively in

(i) V_0,

(ii) $V_{\texttt{protein}} \cap V_0^{\perp}$,

(iii) $\left(V_{\texttt{protein}} + V_{\texttt{fishmeal}}\right) \cap V_{\texttt{protein}}^{\perp}$,

(iv) $V_{\texttt{protein}\wedge\texttt{fishmeal}} \cap \left(V_{\texttt{protein}} + V_{\texttt{fishmeal}}\right)^{\perp}$.

In this case, the second space is $W_{\texttt{protein}}$ by definition; the third space is $W_{\texttt{fishmeal}}$ because we have ensured that $W_{\texttt{protein}}$ is orthogonal to $W_{\texttt{fishmeal}}$; and the final space is $W_{\texttt{protein}\wedge\texttt{fishmeal}}$ by definition.

5 (Another orthogonal block design: first look) The data for the eelworm experiment described in Question 4.1 are the in file `nematode.dat`. Load this into Genstat naming the columns `block`, `plot`, `dose`, `type`, `old` and `new`. The first four should be factors, the last two are the numbers of eelworm cysts in the spring and at harvest respectively. Calculate the variate `logcount` and put it into the spreadsheet.

We can describe each treatment uniquely by the combination of levels of `dose` and `type`. Genstat writes `dose.type` to denote what we have called `dose` $\wedge$ `type`. To create this new factor,

`Spread` $\rightarrow$ `Factor` $\rightarrow$ `Product/Combine`

Select the factors `dose` and `type`, call the resulting factor `treatment` and save it in the spreadsheet.

To get an idea of the differences between the treatments, draw a boxplot of the data for each treatment. Of course, these boxplots are distorted because of the effects of the blocks. We shall draw a second set of boxplots, removing this distortion.

To remove the effects of the blocks, we can do an analysis of variance in which we choose/type

Treatment Structure: `block`
Block Structure:

Before clicking on OK, see whether Genstat will allow you to click on Save...

(i) If you can click on Save... then do so, and choose to save the fitted values in a variate called `bmeans`.

(ii) If not, then just perform the anova. The output window will the show the block means. Make a new column in the spreadsheet, called `bmeans`, and copy the mean for the first block into all the rows for the first block, and so on for each block.

Then calculate a new variate `lcb` by

$$\texttt{lcb} = \texttt{logcount} - \texttt{bmeans}.$$

This should have the effect of subtracting the block means from the original values of `logcount`.

Now draw the boxplots for the variate `lcb`. How do they compare to the previous boxplots in terms of

- medians for each treatment
- relative positions of the medians
- spread for each treatment?

The boxplots give a good visual impression. You can give more precise answers to the above questions if you do

`Stats` → `Summary statistics` → `Summarize Contents of Variables`

to obtain summary statistics for both `logcount` and `lcb` grouped by `treatment`.

6 (Orthogonal block design continued: unstructured treatments) If we ignore the factorial structure of the treatments we can analyse the data `logcount` by declaring

Treatment Structure: `treatment`

and giving the appropriate **Block Structure**. Do this, and compare the output with what you calculated by hand in Question 5.1.

In fact, there is no need for the explicit calculation of the factor `treatment`. Do the analysis again using

Treatment Structure: `dose.type`

You should get exactly the same as before.

7 (Orthogonal block design continued: factorial treatments) To do a simple factorial analysis (fitting main effects of `dose` and `type` and also their interaction) we need (at this stage) to remove the experimental units with the control treatment. To do this:

`Spread` → `Restrict/Filter` → `To Groups (using factor levels)`

and choose *either* to exclude those plots where `dose` has level 0 *or* to include those plots where `dose` has levels 1 and 2. Then do a factorial analysis of the data.