QUEEN MARY, UNIVERSITY OF LONDON

MAS 314 — Design of Experiments

Practical 11 — 21–28 March 2007

In this practical we see how to analyse data from a Graeco-Latin square. This introduces us to different ways of limiting which interactions are included in the treatment model.

After that, we begin looking at how to use the Hasse diagram to work out what GenStat syntax to use.

1 (A Graeco-Latin square) Four bottle-making machines were used for four days to test two treatment factors—mixture of ingredients for the glass, and type of colouring additive. Both treatment factors had four levels. It was assumed that there was no interaction between mixture and colour, and so a main-effects-only plan was used.

The data are in the file `bottle.dat`, using the same order conventions as in `wear.dat` and `apple.dat`, with machines as rows and days as columns (so the experimental units go machine 1, day 1, then machine 1, day 2, and so on). The first column shows mixture; the second column shows colour; the third contains the number of flaws in bottles made in a one-hour run of that machine on that day.

Analyse the data using

Treatment Structure:	`mixture + colour`
Block Structure:	`machine * day`

Use $+$ rather than $*$ to omit the interaction between `mixture` and `colour`.

Interpret the output.

2 (Including or omitting interactions) The menu for the general analyis of variance has a box which usually looks like this:

Interactions: `All interactions`

You can use the mouse to change this to

Interactions: `No interactions`

Analyse the data again, using

Treatment Structure: `mixture * colour`

and specifying no interactions. Do you get exactly the same output?

The third option is

Interactions: `Specify level of interaction` .

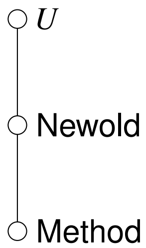
If you choose this, you also have to choose an number. The default is 3. This means that GenStat will fit factors like $A \wedge B \wedge C$ but not factors like $A \wedge B \wedge C \wedge D$. As we shall see in the next question, GenStat's idea of what an interaction is can be different from ours, so you sometimes need to be careful with the setting of this number. This option is not appropriate for this example unless you choose the number 1: anything else will give an error.

3 (A Latin square in which two letters have been merged) A greenhouse experiment was conducted to compare four methods of propagating lettuces. The square layout of the experiment is shown below, along with the weight (in kg) of lettuces produced on each plot.

C 20.1	*A* 14.3	*B* 25.1	*D* 21.6	*A* 17.8
B 22.0	*C* 21.0	*A* 12.9	*A* 16.2	*D* 19.1
A 14.0	*D* 21.5	*C* 22.4	*B* 20.6	*A* 19.1
A 12.8	*A* 16.4	*D* 19.9	*C* 21.5	*B* 22.8
D 21.3	*B* 20.0	*A* 15.5	*A* 13.4	*C* 21.8

Analyse the data.

In fact, method *A* was new while the other three methods were old. Introduce a two-level factor Newold to distinguish between the new and old methods. Now the Hasse diagram for the treatment factors is as follows.



To tell GenStat to analyse this, you need to use the

Treatment Structure: `Newold / Method`

As explained in Practical 5, this causes GenStat to decompose the treatment subspace into the following, in order:

(a) V_0, for the grand mean;

(b) $V_{\mathsf{Newold}} \cap V_0^{\perp}$, for the difference between new and old;

(c) $V_{\mathsf{Newold} \wedge \mathsf{Method}} \cap V_{\mathsf{Newold}}^{\perp}$, which is the same as $V_{\mathsf{Method}} \cap V_{\mathsf{Newold}}^{\perp}$, because $\mathsf{Newold} \wedge \mathsf{Method} = \mathsf{Method}$.

Thus we obtain the correct decomposition of the treatment space.

Now re-analyse the data. What do you deduce?

Unfortunately, GenStat may refer to the last part of the treatment space as the *interaction* between Newold and Method, because it has created the factor $\mathsf{Newold} \wedge \mathsf{Method}$. It thinks that this is an interaction with level 2. You can tell this by looking at the way it displays the table of means in the output. This does not cause any problem in this simple case, but in more complicated situations you may need to be aware of this different point of view.

4 (Scabby data with control) Look again at the output for Question 7 of Practical 3. You should now be able to fully understand the answer to Question 2 of Assignment 3.

5 (Crossover design on cows) Retrieve the spreadsheet for this that you created in Practical 8.

Now the experimenter tells you that the feeds are:

1	2	3
good hay	poor hay	straw.

Create a two-level factor `haystraw` to distinguish between hay and straw.

Reanalyse the data. Are all differences between the feeds explicable as the difference between hay and straw?