Genstat®

Anova and design

www.vsni.co.uk

VSNi

A Guide to Anova and Design in Genstat®
(22nd Edition)

by Roger Payne.

Genstat is developed by VSN International Ltd, in collaboration with practising statisticians at Rothamsted and other organisations in Britain, Australia, New Zealand and The Netherlands.

Published by: VSN International, 2 Amberside, Wood Lane,
Hemel Hempstead, Hertfordshire HP2 4TP, UK
E-mail: info@genstat.co.uk
Website: http://www.genstat.co.uk/

First published 2007, for GenStat *for Windows* 10th Edition
This edition published 2022, for Genstat *for Windows* 22nd Edition

Contents

Introduction 1

1 From t-test to one-way anova 2
- 1.1 Comparing two treatments: the two-sample t-test 3
- 1.2 Practical 8
- 1.3 One-way analysis of variance 8
- 1.4 Practical 15
- 1.5 One-way analysis of variance with several treatments 16
- 1.6 Polynomial contrasts 17
- 1.7 Practical 19
- 1.8 Multiple comparisons 20
- 1.9 Practical 21
- 1.10 Equivalence tests 22
- 1.11 Practical 24
- 1.12 Completely randomized designs 24

2 Blocking structures 25
- 2.1 Completely randomized designs 26
- 2.2 Randomized block designs 26
- 2.3 Practical 31
- 2.4 Blocking in two directions: Latin square designs 31
- 2.5 Practical 34

3 Treatment structure 35
- 3.1 Factorial designs with two treatment factors 36
- 3.2 Fitting contrasts 39
- 3.3 Practical 46
- 3.4 Syntax of model formulae 46
- 3.5 Factorial plus added control 48
- 3.6 Covariates 51
- 3.7 Practical 54
- 3.8 Summaries of results 55
- 3.9 Practical 56

4 Checking the assumptions 57
- 4.1 Homogeneity of variance 58
- 4.2 Normality and independence of the residuals 59
- 4.3 Additivity of the model 60
- 4.4 Outliers 60
- 4.5 Transformations 61
- 4.6 Automatic testing of the assumptions 65
- 4.7 Practical 68
- 4.8 Permutation and exact tests 68
- 4.9 Practical 69

5 Designs with several error terms 70
- 5.1 Split-plot design 71
- 5.2 Practical 73
- 5.3 Other stratified designs 74
- 5.4 Practical 77

6 Design and sample size 78
- 6.1 Designing an experiment 79
- 6.2 Practical 83
- 6.3 Control treatments 83
- 6.4 Practical 85

7 Balance and non-orthogonality 86
- 7.1 Confounding and efficiency factors 87
- 7.2 Balance 92
- 7.3 Practical 93
- 7.4 Unbalanced designs with two treatment factors 94
- 7.5 Practical 96
- 7.6 Unbalanced designs with several treatment factors 97
- 7.7 Practical 102

8 REML analysis of unbalanced designs 103
- 8.1 Linear mixed models: split-plot design 104
- 8.2 Practical 109
- 8.3 Linear mixed models: a non-orthogonal design 109
- 8.4 Practical 117
- 8.5 Analysis of variance by ANOVA, regression or REML 117
- 8.6 Practical 119

9 Commands for analysis of variance 120

Index 126

Introduction

Analysis of variance is one of the most widely used statistical techniques, with application areas that include biology, medicine, industry and finance. Genstat has a very powerful set of ANOVA techniques, that are nevertheless very straightforward and easy to use.

This book is designed to introduce you to these techniques, and give you the underlying knowledge and confidence to use them correctly and effectively. It also covers the basic principles of experimental design to help you plan effective experiments and investigations. It was written to provide the notes for VSN's course on anova and design in Genstat, but it can be used equally well as a self-learning tool.

Starting with the simplest situation, where two different treatments are compared by the standard t-test, straightforward examples will be used to introduce the following concepts.

- *Analysis* – covering simple to sophisticated situations, explaining ideas such as balance, and introducing advanced features like the use of REML for unbalanced designs
- *Interpretation* – explaining the results, producing relevant tables, graphs and figures for publication in reports and papers.
- *Design* – a range of experimental designs will be described, to cover the situations encountered by most Genstat users.
- *Blocking* – how to increase the accuracy of an experiment by forming the basic units (e.g. plots or subjects) into groups with similar properties.
- *Randomization* – how to avoid bias in the allocation of units to treatments, so that you can ensure that your results are reliable and unaffected by any systematic patterns in the units.
- *Replication* – determining how many replicates you need.
- *Treatments* – comparing several types of treatment in the same experiment
- *Covariates* – to improve precision by using additional background information about the experimental units, that was not used for blocking.

1 From t-test to one-way anova

In this chapter you will learn

- how to use the t-test to compare two treatments
- the mathematical equations that lie behind the t-test ★
- how to calculate a t-test by hand ★
- the T-Test menu
- how to use one-way analysis of variance to compare several treatments
- the model fitted in one-way anova
- the mathematical equations that lie behind one-way anova ★
- the statistical philosophy behind one-way anova
- the relationship between one-way anova and the t-test for two treatments
- how to use the One- and two-way ANOVA menu for one-way anova
- how to plot the means from one-way anova
- how to fit polynomial contrasts to quantitative treatments ★
- how to do multiple-comparison tests ★
- how to do equivalence tests ★

Note: the topics marked ★ are optional.

1.1 Comparing two treatments: the two-sample t-test

Suppose we have two sets of units, each of which has received a different treatment. For example, they might be animals that have been fed two different diets, or plots that have been given different fertilisers, or subjects with different drugs, or plants with different fungicides, or widgets that have been formed by different manufacturing methods, and so on.

In this first section, we assume that the units do not have any special structure – for example that the animals are all of the same breed, or that the plots are in a fairly uniform field, or that the subjects are of similar ages, weights and heights, and so on. So we have two sets of observations (one for each treatment), and we want to know if they differ by more than random variation.

The table shows data from an (unstructured) experiment to study yields from two different manufacturing methods.

We want to know whether the yields of the two methods differ by more than we would expect from the random variability in the experiment. We would also like to *estimate* the likely yields from each method. Data like this are often analysed using a two-sample t-test.

The assumption for the t-test is that each group has a Normal distribution. It is generally assumed that the distributions both have the same variance (this can be checked) and that they may have different means.

standard	23
new	24
new	21
standard	22
new	22
standard	19
standard	21
new	20
new	25
standard	20
standard	17
new	26
standard	18
new	24
new	22
standard	20

Example 1.1

We estimate the means by the averages of the observations with each treatment.

standard: (23 + 22 + 19 + 21 + 20 + 17 + 18 + 20) / 8 = 20

new: (24 + 21 + 22 + 20 + 25 + 26 + 24 + 22) / 8 = 23

If you'd like to see this in mathematical notation, the mean of the distribution of the data $\{y_{ij} : j = 1...n_i\}$ in group i is estimated by

$$m_i = (y_{i1} + y_{i2} + ... + y_{in_i}) / n_i$$

(If not, please ignore this and the later equations!) This calculation is usually written as

$$\hat{m}_i = \sum_{j=1}^{n_i} y_{ij} / n_i$$

where the $\sum$ symbol represents summation from the lower value 1 to the upper value n_i.

If the treatments have the same effect, the difference between the means, then

$d = m_1 - m_2$

should be zero. However, we have only an *estimate* of the difference. So, we need to know how variable this estimate might be. We can estimate the *standard error* of the distributions by the sum of the squares of the differences between each observation and the mean for the variety involved, divided by the *degrees of freedom* (essentially the number of "spare parameters" that we have left from our $n_1 + n_2$ observations after fitting the 2 means).

$$\{ 3^2 + 2^2 + (-1)^2 + 1^2 + 0^2 + (-3)^2 + (-2)^2 + 0^2$$
$$+ 1^2 + (-2)^2 + (-1)^2 + (-3)^2 + 2^2 + 3^2 + 1^2 + (-1)^2 \} / \{16 - 2\}$$

or, in mathematical notation,

$$\hat{s} = \sqrt{\{ \sum_{i=1}^{2} \sum_{j=1}^{n_i} (y_{ij} - \hat{m}_i)^2 / (n_1 + n_2 - 2) \}}$$

The standard error of the difference of the two means is

$\hat{s}_d = \hat{s} \times \sqrt{\{(n_1 + n_2) / (n_1 \times n_2)\}}$.

The t-statistic is simply the estimate of the difference divided by its standard error. So, to make a t-test for the *hypothesis* that there is no difference between the means, we just need to calculate $(\hat{m}_1 - \hat{m}_2) / \hat{s}_d$ or, in mathematical notation,

$$\frac{\sum_{j=1}^{n_1} y_{1j}/n_1 - \sum_{j=1}^{n_2} y_{2j}/n_2}{\sqrt{[\{ \sum_{i=1}^{2} \sum_{j=1}^{n_i} (y_{ij} - (\sum_{j=1}^{n_i} y_{ij}/n_i))^2 / (n_1+n_2-2)\} \times \{(n_1+n_2)/(n_1 n_2)\}]}}$$

We can then compare this with the appropriate value of the t-distribution for n_1+n_2-2 degrees of freedom.

To summarise, to do a t-test by hand:

- calculate the average of the observations in group 1 ($\hat{m}_1$)
- calculate the average of the observations in group 2 ($\hat{m}_2$)
- subtract the smaller from the larger ($\hat{d} = \hat{m}_1 - \hat{m}_2$)
- subtract the averages from the data values in the respective groups
- square the values (after subtracting the averages), add them up, divide by $\{ (n_1 + n_2 - 2) \times n_1 \times n_2 / (n_1 + n_2) \}$ and take the square root (this gives $\hat{s}_d$)
- finally, divide $\hat{d}$ by $\hat{s}_d$ and compare with the t distribution for n_1+n_2-2 degrees of freedom.

As in much experimental design, this is very much simpler if we have the same *replication* (that is, number of observations) for each treatment. Then $n_1=n_2=n$, and the t-statistic is

$$\left(\sum_{j=1}^{n} y_{1j}/n \; - \; \sum_{j=1}^{n} y_{2j}/n\right) \; / \; \sqrt{\left[\; \sum_{i=1}^{2} \sum_{j=1}^{n} \{y_{ij} - \left(\sum_{j=1}^{n} y_{ij}/n\right)\}^2 \; / \; \{(n-1)\times n\} \; \right]}$$

Complicated equations are less of a problem on a course like this, as we can use Genstat to do the calculations. However, another important consideration is that, with equal replication, we are estimating each mean with the same precision, and this may be important for example in drug and variety trials where we may need to show the originators of each drug or variety that it has been assessed fairly in comparison with the other drug or variety.

It is much simpler to analyse the experiment using Genstat. The data sets that are used in the examples and practicals in this Guide can be all be accessed from within Genstat. Click on File on the menu bar, and select the Open Example Data Sets option, as shown in Figure 1.1.

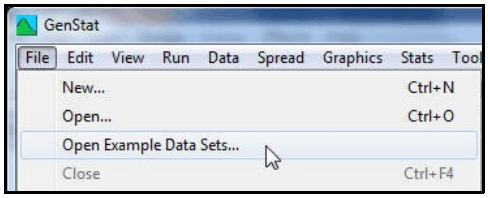


Figure 1.1

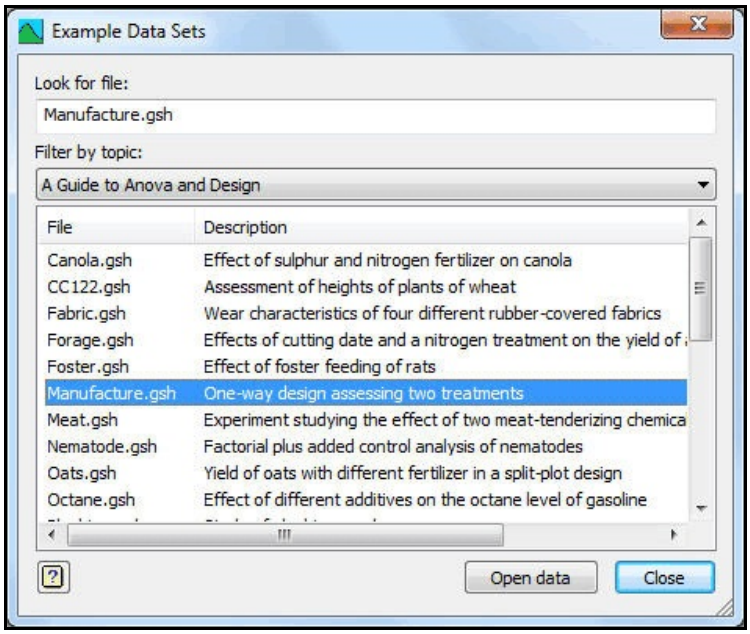


Figure 1.2

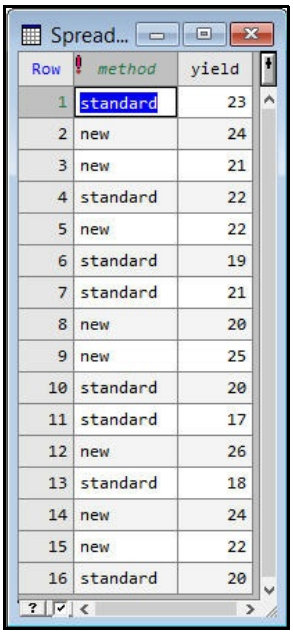


Figure 1.3

This opens the Example Data Sets menu, shown in Figure 1.2. It is easier to find the relevant file if you set the Filter by topic drop-down list to A Guide to Anova and Design. The data for the example in this section is available in the Genstat spreadsheet file `Manufacture.gsh`. So we select that file, and click on the Open data button.

The file is shown in Figure 1.3. There are two columns of data: the name *method* is in italics, showing that this column is a factor, and `yield` is a variate.

We can check some of our arithmetic by using the Summary Statistics menu, which you can open by clicking on the Summary Statistics sub-option of the Summary Statistics option of the Stats menu on the menu bar. The summary produced by the menu in Figure 1.4 is shown below.

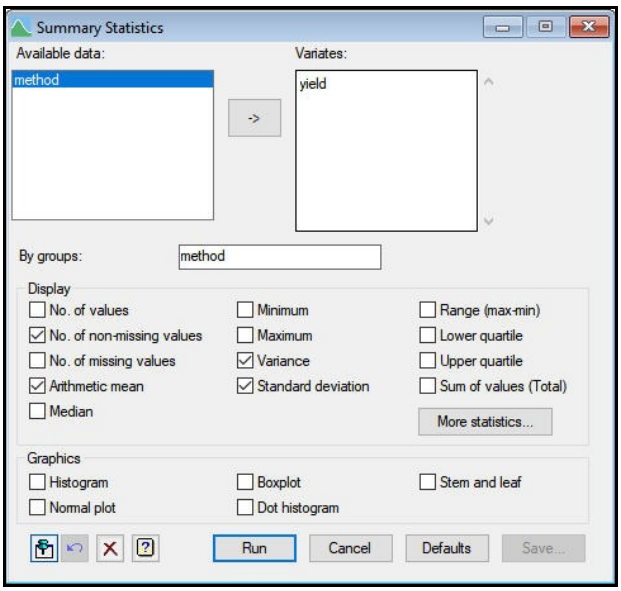


Figure 1.4

Summary statistics for yield: method new

Number of observations =	8
Mean =	23
Standard deviation =	2.070
Variance =	4.286

Summary statistics for yield: method standard

Number of observations =	8
Mean =	20
Standard deviation =	2
Variance =	4

To calculate the t-test directly, we open the T-Tests menu (Figure 1.5) by clicking on the One- and two-sample t-test sub-option of the Statistical-tests option of the Stats option on the menu bar. We select Two-sample in the Test drop-down list box, and One variate with group factor as the Data arrangement. We can then enter `yield` as the Data variate, and `method` as the Group factor defining the two groups. Clicking Run generates the output below.

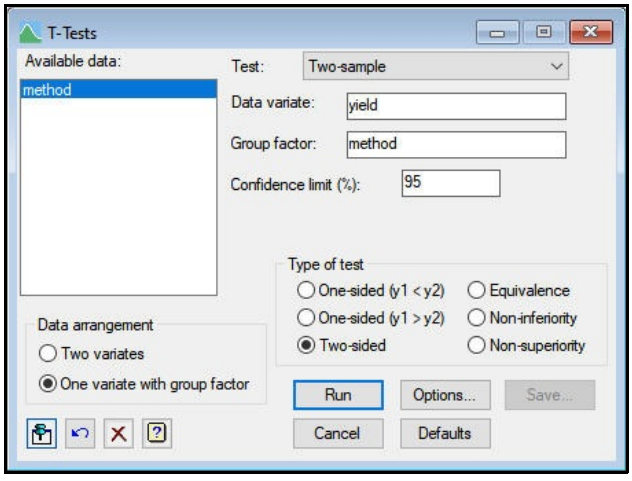


Figure 1.5

Two-sample t-test

Variate: yield
Group factor: method

Test for equality of sample variances

Test statistic F = 1.07 on 7 and 7 d.f.

Probability (under null hypothesis of equal variances) = 0.93

Summary

Sample	Size	Mean	Variance	Standard deviation	Standard error of mean
new	8	23.00	4.286	2.070	0.7319
standard	8	20.00	4.000	2.000	0.7071

Difference of means: 3.000
Standard error of difference: 1.018

95% confidence interval for difference in means: (0.8173, 5.183)

Test of null hypothesis that mean of yield with method = new is equal to mean with method = standard

Test statistic t = 2.95 on 14 d.f.

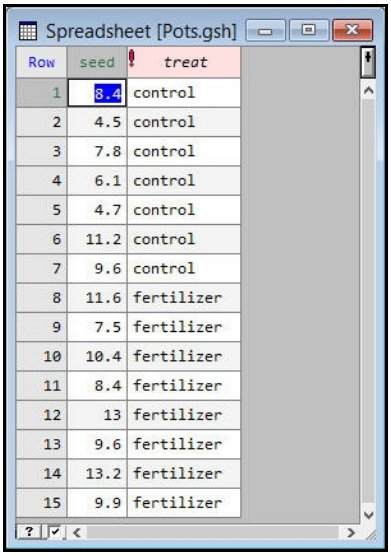
Probability = 0.011

The t-statistic is 2.95 on 14 degrees of freedom. Under the "null hypothesis" that there is no difference between the means, this would have a probability of 0.011. We can conclude that this is unlikely. So there is evidence that the manufacturing methods do differ.

1.2 Practical

Seven plants of wheat grown in pots and given no fertilizer treatment yield 8.4, 4.5, 7.8, 6.1, 4.7, 11.2 and 9.6g dry weight of seed. A further eight plants from the same source are grown in similar conditions but given a fertilizer treatment. These plants yield 11.6, 7.5, 10.4, 8.4, 13.0, 9.6, 13.2 and 9.9g dry weight respectively. The data are held in file `Pots.gsh` as two columns: the first holds the seed weights (variate `seed`) and the second holds factor `treat` indicating whether or not there was any fertilizer (control/fertilizer).

Read the data into Genstat, then look to see whether the fertilizer has an effect on seed production by carrying out a two-sample t-test using the T-Test menu.



Row	seed	treat
1	8.4	control
2	4.5	control
3	7.8	control
4	6.1	control
5	4.7	control
6	11.2	control
7	9.6	control
8	11.6	fertilizer
9	7.5	fertilizer
10	10.4	fertilizer
11	8.4	fertilizer
12	13	fertilizer
13	9.6	fertilizer
14	13.2	fertilizer
15	9.9	fertilizer

Figure 1.6

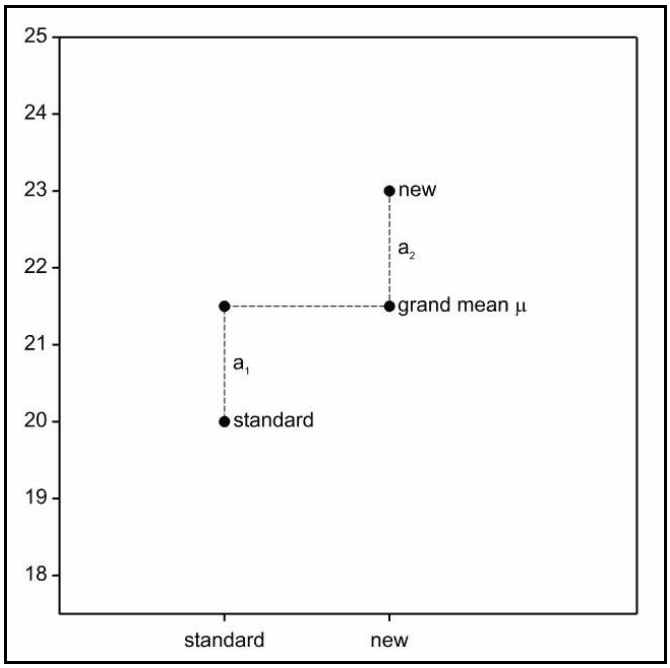
1.3 One-way analysis of variance

Another way of representing the situation, is that we have a *linear model*

$$y_{ij} = \mu + a_i + \varepsilon_{ij}$$

where each observation is represented by its mean m_i (which we have chosen to write as $\mu + a_i$) plus a *residual* ε_{ij} which represents the random variation in the situation.

For our example, it represents the data as follows:



standard	23 22 19 21 20 17 18 20	=	21.5	+	standard	−1.5	+	standard	3 2 −1 1 0 −3 -2 0
new	24 21 22 20 25 26 24 22				new	1.5		new	1 −2 −1 −3 2 3 1 −1
y_{ij}			$\hat{\mu}$		$\hat{a}_i$			ε_{ij}	

The residual variation can arise from many different causes, for example:

- the units may not be absolutely identical (and we shall discuss later how units should be allocated to treatments to take account of this),
- they may then experience slightly different conditions during the experiment,
- there may be measurement errors,
- they may be being dealt with by different people during the experiment.

The form of the model suggests another approach. If we were to assume that the treatments are both identical, then their *effects* a_1 and a_2 would be zero. Our model would simply be

$$y_{ij} = \mu + \varepsilon_{ij}$$

and we would estimate the *grand mean* μ by the average of all the data values: that is

$$\hat{\mu} = \sum_{i=1}^{2} \sum_{j=1}^{n_i} y_{ij} / (n_1 + n_2)$$

One way of measuring how well this model fits is to take the sum of squares of the residuals from this model (that is, to add up the squares of our estimates of the random variation on each observation for this model).

$$RSS_0 = \sum_{i=1}^{2} \sum_{j=1}^{n_i} (y_{ij} - \hat{\mu})^2$$

This has $n_1 + n_2 - 1$ degrees of freedom as we have fitted just one parameter, μ.

Now compare this with the full model above, in which the treatments are assumed to have different effects: we can estimate a_i by the mean of the observations that received treatment i, minus the overall mean, that is

$$\hat{a}_i = \sum_{j=1}^{n_i} y_{ij} / n_i - \hat{\mu}$$

and the residual sum of squares is given by

$$\begin{aligned} RSS_1 &= \sum_{i=1}^{2}\sum_{j=1}^{n_i}(y_{ij} - \hat{\mu} - \hat{a}_i)^2 \\ &= \sum_{i=1}^{2}\sum_{j=1}^{n_i}(y_{ij} - \hat{\mu})^2 - 2\sum_{i=1}^{2}\sum_{j=1}^{n_i}(y_{ij} - \hat{\mu})\hat{a}_i + \sum_{i=1}^{2}\sum_{j=1}^{n_i}\hat{a}_i^2 \\ &= \sum_{i=1}^{2}\sum_{j=2}^{n_i}(y_{ij} - \hat{\mu})^2 - \sum_{i=1}^{2} n_i\hat{a}_i^2 \end{aligned}$$

with $n_1 + n_2 - 2$ degrees of freedom. This takes a little thought as it may appear as though we have fitted three parameters but, in fact, there are really just the two means $\hat{m}_1$ and $\hat{m}_2$. Our use of the treatment effects a_1 and a_2 makes it easy to move from one model to the other (by setting them both to zero) but you can easily see that

$$\begin{aligned} \hat{\mu} &= (\hat{m}_1 \times n_1 + \hat{m}_2 \times n_2) / (n_1 + n_2) \\ &= \{(\hat{\mu} + \hat{a}_1) \times n_1 + (\hat{\mu} + \hat{a}_2) \times n_2\} / (n_1 + n_2) \end{aligned}$$

and so

$$\hat{a}_1 \times n_1 = -\hat{a}_2 \times n_2.$$

The difference between these two sums of squares is known as *the sum of squares due to the treatments*. This measures the effect of allowing for two different means, and has one degree of freedom. We can assess whether this exceeds the underlying level of variability by comparing it with RSS_1, but first we need to divide each one by its degrees of freedom to give the treatment and residual *mean squares*; this takes account of the different number of parameters that each one represents. By dividing the treatment mean square by the residual mean square, we obtain a statistic known as the *variance ratio*. If we assume that the residuals follow a Normal distribution, the variance ratio will have an F distribution on 1 and $(n_1 + n_2 - 2)$ degrees of freedom. (The degrees of freedom are the degrees of freedom for the nominator – that is the sum of squares due to treatments – and those for the denominator – that is the residual sum of squares.) The variance ratio is

$$VR = \sum_{i=1}^{2}\{ n_i\hat{a}_i^2 \} / \{ \sum_{i=1}^{2}\sum_{j=1}^{n_i}(y_{ij} - \hat{\mu} - \hat{a}_i)^2 / (n_1 + n_2 - 2) \}$$

It is interesting to note that, when there are only two treatments, the variance ratio is the square of the t-statistic. You can verify this in the example below, or see the proof in the following equations:

$$\begin{aligned} \sum_{i=1}^{2} n_i\hat{a}_i^2 &= \sum_{i=1}^{2} n_i(\hat{m}_i - \hat{\mu})^2 \\ &= n_1\{\hat{m}_1 - (n_1\hat{m}_1 + n_2\hat{m}_2)/(n_1 + n_2)\}^2 \\ &\quad - n_2\{\hat{m}_2 - (n_1\hat{m}_1 + n_2\hat{m}_2)/(n_1 + n_2)\}^2 \\ &= n_1 n_2(\hat{m}_1 - \hat{m}_2)^2/(n_1 + n_2) \end{aligned}$$

and so

$$VR = (\hat{m}_1 - \hat{m}_2)^2 / [\{(n_1 + n_2)/n_1 n_2)\} \times \{\sum_{i=1}^{2} \sum_{j=1}^{n_i} (y_{ij} - \hat{m}_i)^2 / (n_1 + n_2 - 2)\}]$$

The variance ratio, however, can be used if there are more than two treatments. Usually, the information is all laid out in an *analysis of variance* table. For our example this is:

Analysis of variance

Variate: yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
method	1	36.000	36.000	8.69	0.011
Residual	14	58.000	4.143		
Total	15	94.000			

Mathematically, when there are t treatments, the one-way analysis of variance can be calculated as follows:

Source	Sums of squares	Degrees of freedom	Mean square	Variance ratio
Treatments	$\sum_i n_i \hat{a}_i^2 =$ $\sum_i n_i \hat{m}_i^2 - (\sum_i n_i) \hat{\mu}^2$	$t - 1$	$(\sum_i n_i \hat{a}_i^2) / (t - 1)$	treatment mean square / residual mean square
Residual	$\sum_i \sum_j (y_{ij} - \hat{\mu} - \hat{a}_i)^2$ or as Total SS – Treat SS	$\sum_i n_i - t$	$\{\sum_i \sum_j (y_{ij} - \hat{\mu} - \hat{a}_i)^2\}$ $/ (\sum_i n_i - t)$	
Total	$\sum_i \sum_j (y_{ij} - \hat{\mu})^2$ $= \sum_i \sum_j y_{ij}^2 - (\sum_i n_i)\hat{\mu}^2$	$\sum_i n_i - 1$	$\{\sum_i \sum_j (y_{ij} - \hat{\mu})^2\}$ $/ (\sum_i n_i - 1)$	

Notice that the total sum of squares in the table is RSS_0. Usually there is no interest in assessing whether the observations have a non-zero overall mean, and so the table contains the total sum of squares "corrected for the grand mean". Also notice that two possible formulae are given for the Treatment and Total sums of squares. The second may be more convenient to calculate, but the first will be much more accurate if the accuracy of the representation is limited, as on computers or calculators.

Alternatively, we can ignore all this mathematics and use Genstat. The Analysis of Variance section of the Stats menu on the menu bar (Figure 1.7) offers two possibilities. One-way analysis of variance is easiest with the One- and two-way Analysis of Variance menu (Figure 1.8). Later in the Course, we will introduce the general Analysis of Variance menu, which accesses the full power of GenSat's analysis of variance facilities.

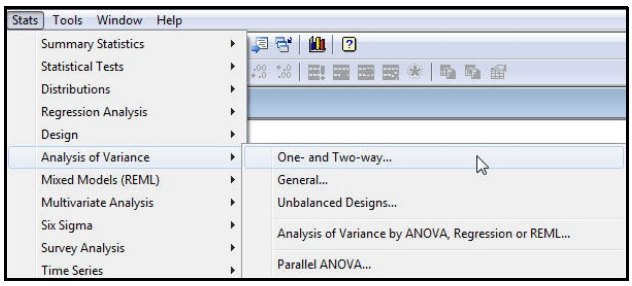


Figure 1.7

We select One-way as the Design, enter the name of the Y-variate (`yield`) and of the factor defining the Treatments (`method`), and then click on Run.

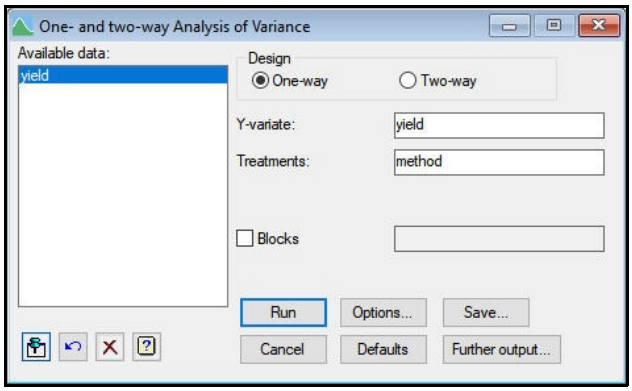


Figure 1.8

The output from the analysis is controlled by the ANOVA Options menu (Figure 1.9), obtained by clicking on the Options button on the One- and two-way Analysis of Variance menu.

With the analysis-of-variance table, we usually also present tables of means with associated standard errors or (more usefully) standard errors for differences between pairs of means (s.e.d's): for two means with replication n_1 and n_2,

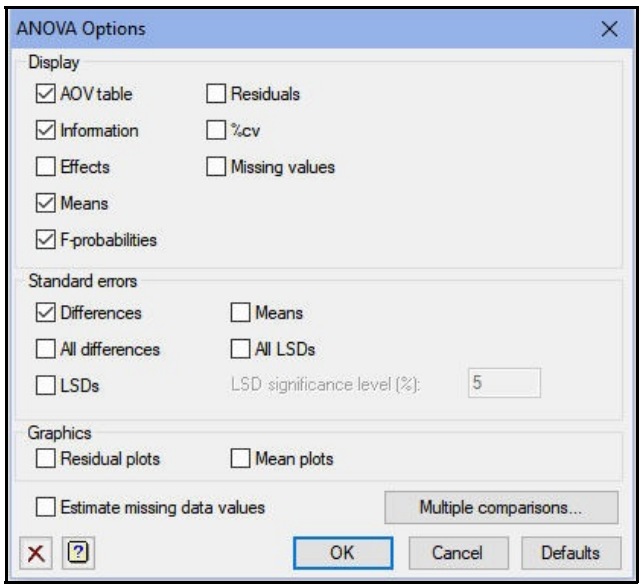


Figure 1.9

$$\text{s.e.d.} = \sqrt{\{\ (\text{residual-mean-square}) \times (1/n_1 + 1/n_2)\ \}}$$
$$= \sqrt{\{\ (\text{residual-mean-square}) \times (n_1 + n_2) / (n_1 \times n_2)\ \}}$$

You may recognise this as the denominator of the t-statistic from Section 1.1. In fact differences between means from analysis of variance, divided by their s.e.d., also follow t distributions (with degrees of freedom given by the residual d.f.).

Genstat can also produce *least significant differences*. These are s.e.d.'s multiplied by

the relevant t value, allowing a direct comparison with the difference between the means.

Tables of means

Variate: yield

Grand mean 21.50

method	new	standard
	23.00	20.00

Standard errors of differences of means

Table	method
rep.	8
d.f.	14
s.e.d.	1.018

Least significant differences of means (5% level)

Table	method
rep.	8
d.f.	14
l.s.d.	2.183

The philosophy then is that you first look at the variance ratio to assess whether there is any evidence of differences anywhere amongst the treatments; if so, the s.e.d. or the l.s.d. provides the necessary yardstick for comparing pairs of means. In published papers and reports, the analysis-of-variance table is usually omitted – although you would report that differences have been reported between the treatments (if they have!). Tables of means are presented, with their s.e's or s.e.d's.

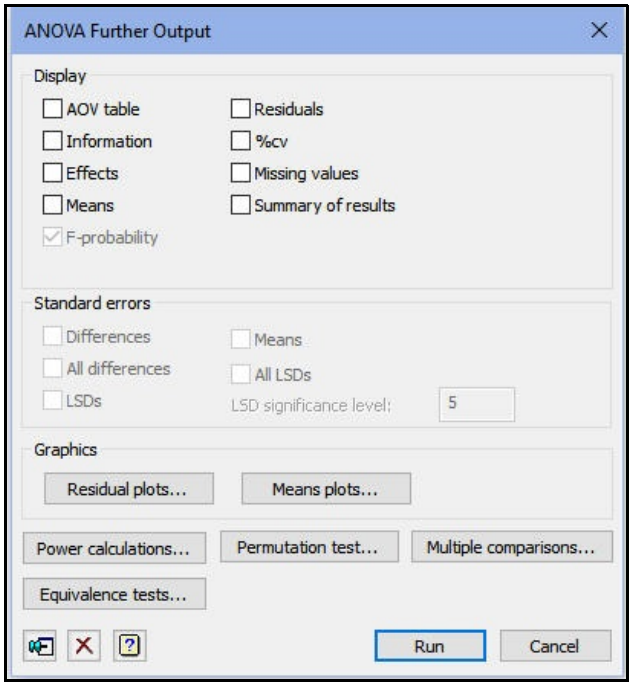


Figure 1.10

You do not need to decide on all your output before you do the analysis. You can obtain additional output by using the ANOVA Further Output menu (Figure 1.10), obtained by clicking on the Further output button on the One- and two-way Analysis of Variance menu.

You can click on the Means plots button to open the Means Plot menu. This allows you to choose how you want to plot the means, and how you want to represent their standard errors. In Figure 1.11 we have chosen to plot points for the means, with a bar to show the s.e.d. (see Figure 1.12). You would plot lines if the treatments represented different amounts of some quantity such as a fertilizer, a drug or a dietary supplement. Plotting the data values (as well as the means) can provide a visual confirmation of the significance (or non-significance) of the treatment effects reported in the analysis-of-variance table. The final possibility is to plot the means as a bar chart.

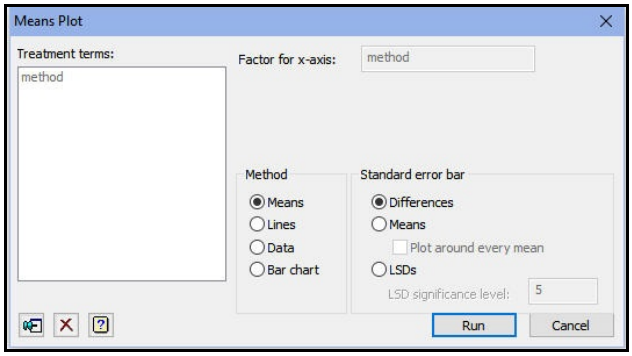


Figure 1.11

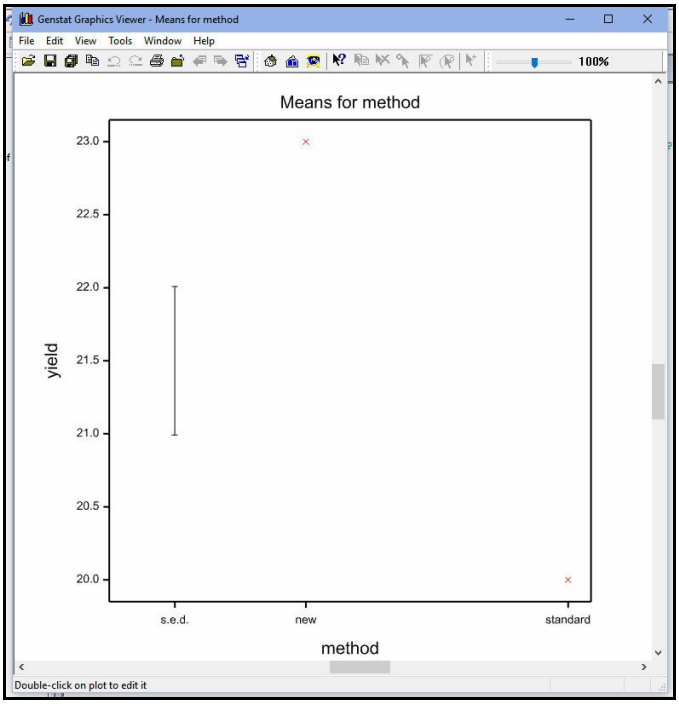


Figure 1.12

You can cut and paste results of the analysis, from the Output window to word processing systems like Microsoft Word. You can also save it into Genstat data structures or to external spreadsheet files. To do this, click on the Save button on the main the One- and two-way Analysis of Variance menu (Figure 1.8) to open the ANOVA Save Options menu, as shown in Figure 1.13. Section 4.5 shows how to use this menu to save a table of means in a Genstat spreadsheet (see Figure 4.10).

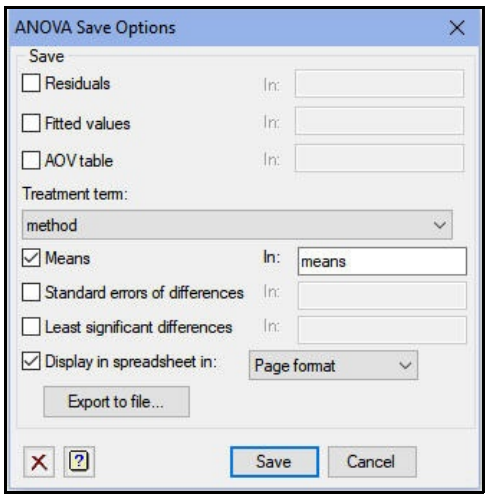


Figure 1.13

Alternatively, you can click on the Export to file button to open the Save ANOVA Results in a Spreadsheet File menu, which allows you to save the output to a spreadsheet file on your computer. Figure 1.14, shows the menu with the default output components selected in the check boxes, and the Save in file box filled in to save them in the Excel file ManufactureResults.xlsx.

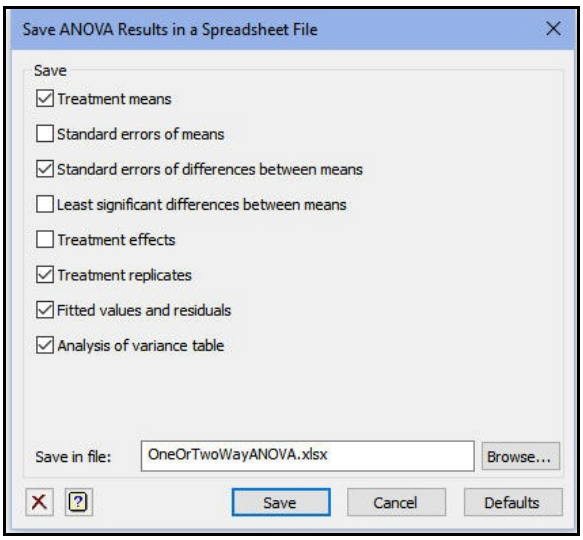


Figure 1.14

Each output component is saved on a separate page in the spreadsheet file. Figure 1.15 shows the page with the treatment means.

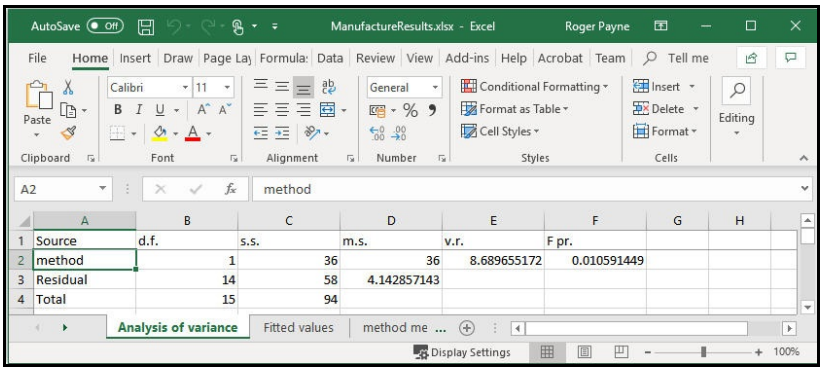


Figure 1.15

1.4 Practical

Do a one-way analysis of variance for the data in `Pots.gsh` and compare the results with those from the t-test. Plot the means, and also plot the data values. Does the plot with the data values confirm what you have found in the analysis of variance? Save the results to an Excel file. Open the file and compare them with the output in the Output window.

1.5 One-way analysis of variance with several treatments

Diet	Weight
a	81.5 80.7 80.3 79.8
b	81.6 81.9 80.4 80.4
c	83.5 81.6 82.2 81.3
d	82.4 83.1 82.8 81.8
e	83.2 82.8 82.1 82.1

The advantages of analysis of variance become clearer when there are more than two treatments.

Spreadsheet file `Rat.gsh` contains data from an experiment to study the effect of a dietary supplement on the gain in weight of animals. There were five different treatments (representing different amounts of the supplement) and twenty animals were allocated at random, four to each treatment. The data be analysed and we can plot the means, using the One- and two-way Analysis of Variance menu as before.

Analysis of variance

Variate: weight

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
diet	4	12.7930	3.1982	6.32	0.003
Residual	15	7.5925	0.5062		
Total	19	20.3855			

Tables of means

Variate: weight

Grand mean 81.76

diet	a	b	c	d	e
	80.58	81.08	82.10	82.53	82.55

Standard errors of differences of means

Table	diet
rep.	4
d.f.	15
s.e.d.	0.503

Least significant differences of means (5% level)

Table	diet
rep.	4
d.f.	15
l.s.d.	1.072

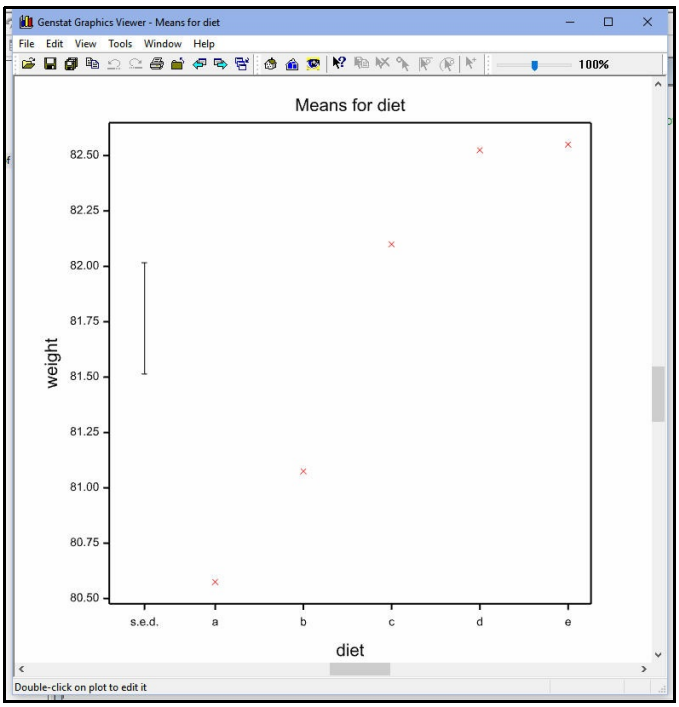


Figure 1.16

1.6 Polynomial contrasts

Suppose the treatments represent amounts 0, 1, 2, 3 and 4 of supplement. We might now be interested to see how linear the relationship is. The general Analysis of Variance menu (Figure 1.17) extends the facilities in the specialized One- and two-way Analysis of Variance menu, to allow you to estimate contrasts amongst the treatments.

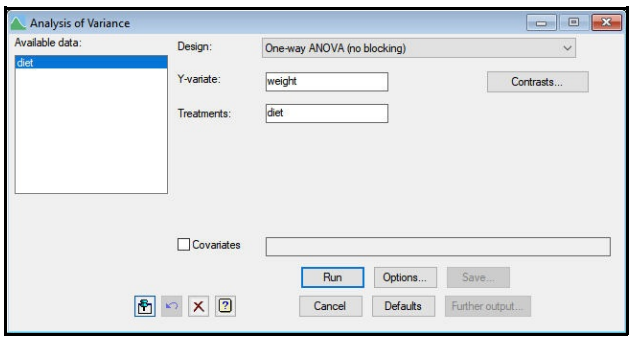


Figure 1.17

The menu is obtained by selecting the General sub-option of the Analysis of Variance option of the Stats menu on the menu bar, instead of the One- and Two-way sub-option (Figure 1.7). Setting One-way ANOVA (no blocking) for the Design provides similar controls to those in the One- and two-way Analysis of Variance menu (Figure 1.8), with the addition of a Contrasts button.

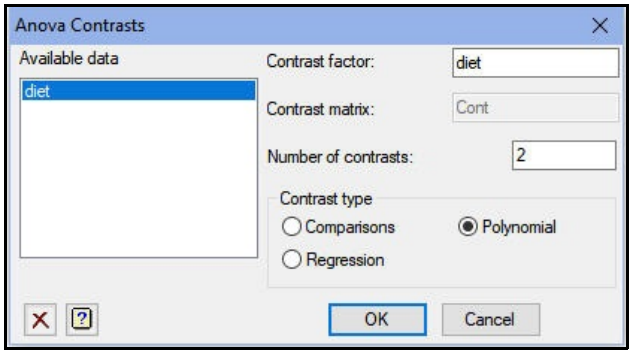


Figure 1.18

This button generates the Anova Contrasts menu (Figure 1.18), in which we have asked Genstat to fit two polynomial contrasts (i.e. linear and quadratic) between diet. The

analysis is now extended to examine the linear and quadratic effects of supplement.

Analysis of variance

Variate: weight

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
diet	4	12.7930	3.1982	6.32	0.003
Lin	1	11.6640	11.6640	23.04	<.001
Quad	1	0.6864	0.6864	1.36	0.262
Deviations	2	0.4426	0.2213	0.44	0.654
Residual	15	7.5925	0.5062		
Total	19	20.3855			

In the analysis of variance, the sum of squares for `diet` is partitioned into the amount that can be explained by a linear relationship of the yields with amount of supplement (the line marked `Lin`), the extra amount that can be explained if the relationship is quadratic (the line `Quad`), and the amount represented by deviations from a quadratic polynomial. A cubic term would be labelled as `Cub`, and a quartic as `Quart`. You are not allowed to fit more than fourth-order polynomials.

The analysis shows that there is a strong linear effect, but no evidence of any curvature (as assessed by the quadratic contrast).

To fit polynomial contrasts, Genstat calculates *orthogonal polynomials* and does a multiple regression of the effects of factor using the polynomials as x-variates (see *Guide to the Genstat Command Language*, Part 2, Section 4.5 for details).

We can obtain additional output, as before, by using the ANOVA Further Output menu. When the menu is opened from the general Analysis of Variance menu it has some additional boxes. In Figure 1.19 we use the menu to print the regression coefficients of the polynomial contrasts, and the equation of the polynomial.

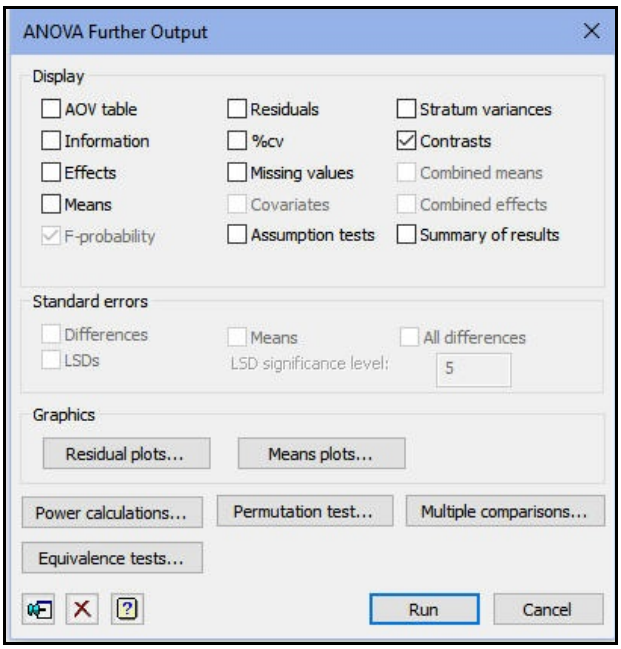


Figure 1.19

Tables of contrasts

Variate: weight

diet contrasts

Lin 0.54, s.e. 0.112, ss.div. 40.0

Quad -0.111, s.e. 0.0951, ss.div. 56.0

Deviations, e.s.e. 0.356, ss.div. 4.00

diet	a	b	c	d	e
	0.11	-0.26	0.11	0.11	-0.07

Equation of the polynomial for diet

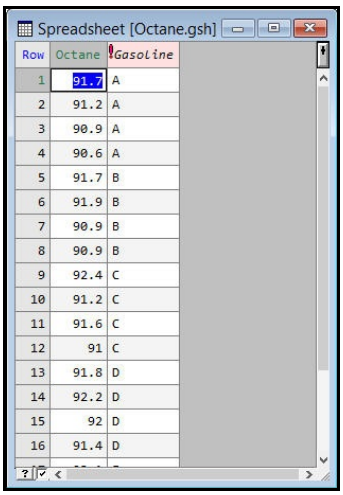
80.46 + 0.98 * diet - 0.11 * diet**2

The orthogonal polynomials cannot be printed from the menu, but they can be saved by the AKEEP directive, and printed by the PRINT directive; see Chapter 9 for more details.

1.7 Practical

Spreadsheet Octane.gsh contains data from an experiment to study the effect of different additives on the octane level of gasoline (P.W.M. John, Statistical Design and Analysis of Experiments, page 46). There were 5 types of gasoline (A-E), and 4 observations on each. Use analysis of variance to assess whether there are differences in octane level between the gasolines.

Suppose that gasolines A-E contain 0, 1, 2, 3 and 4 cc/gallon of additive, respectively (but are otherwise identical). Estimate the linear and quadratic effects of the additive.



Row	Octane	Gasoline
1	91.7	A
2	91.2	A
3	90.9	A
4	90.6	A
5	91.7	B
6	91.9	B
7	90.9	B
8	90.9	B
9	92.4	C
10	91.2	C
11	91.6	C
12	91	C
13	91.8	D
14	92.2	D
15	92	D
16	91.4	D

Figure 1.20

1.8 Multiple comparisons

Multiple-comparison tests are designed to take account of the fact that there may be many possible comparisons between pairs of treatment means in an analysis of variance (with t treatments there are $t \times (t-1)/2$). So, some researchers feel that their significance levels should be adjusted to take account of all the tests that they might make – and this can be achieved by use of a multiple-comparison test. Conversely, it has been pointed out that multiple-comparisons are unnecessary if you have only a small number of comparisons to make – either because there are few treatments, or because you should have identified beforehand the comparisons that you feel are likely to be of interest. Also, they are inappropriate if the treatments have any sort of structure. For example, the levels of a treatment factor may represent different amounts of a substance like a fertiliser or a drug. It would then be more sensible to assess the treatment effect over all its levels by fitting some sort of trend (like the polynomial contrasts that we fitted in Section 1.6), and illogical to assume that only some of the amounts might have an effect.

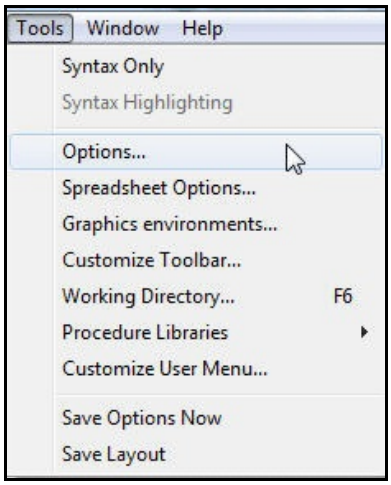


Figure 1.21

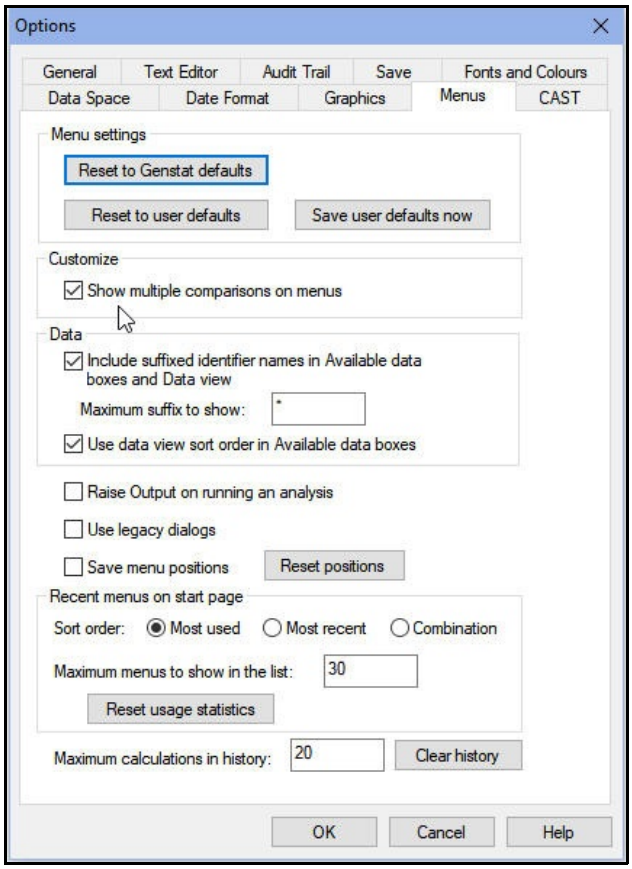


Figure 1.22

However, Genstat does have menus if you do need to use multiple-comparison tests. Because some organisations may want to discourage their use, these can be enabled and disabled through the Options menu. You open the menu by clicking on the Options option of the Tools menu on the menu bar (Figure 1.21). In the menu (Figure 1.22), you need to select the Menus tab, and check the box Show multiple comparisons on ANOVA menus.

There will then be a Multiple comparisons button on the ANOVA Options and Further Output menus, which you can use to open the Multiple Comparisons menu. The menu provides all the standard tests, ranging from Fisher's LSD tests (which simply compare the means using their least significant differences) to e.g. Duncan's, Scheffe's and Tukey's tests.

Genstat will not let us do a multiple comparison test on a treatment term where we have fitted contrasts, as this implies that we have more informative comparisons to make. So we also need to redo the analysis without the polynomial for `diet`.

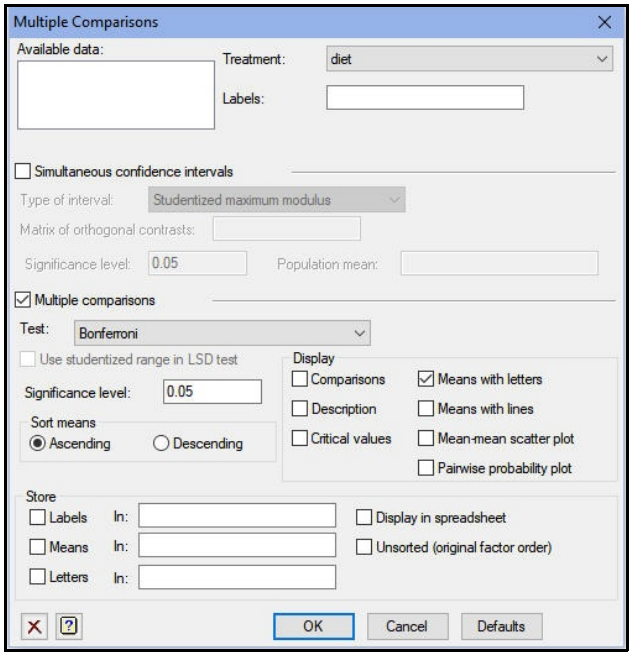


Figure 1.23

In Figure 1.23, we have selected Bonferroni test. If we now click on Run in the Further Output menu, we obtain the output below.

Bonferroni test

diet

Comparison-wise error rate = 0.0050

	Mean	
a	80.58	a
b	81.08	ab
c	82.10	ab
d	82.53	b
e	82.55	b

1.9 Practical

Do a Bonferroni multiple-comparison test to compare the types of gasoline in Practical 1.7.

1.10 Equivalence tests

It is generally accepted that you can use a statistical test to provide evidence that the means of two treatments differ, but it cannot prove that they are identical. A non-significant probability simply means that the results could have been obtained under the null-hypothesis that they have the same means. It does not mean that they *must* have the same means – there will be a range of differences between the means that could also provide non-significant probabilities for the results. This presents difficulties for investigations where you want to show that a new treatment can be used instead of a standard one without causing adverse effects. For example, you might want to show that the side-effects of a new drug are no worse than the current drug, or that your weight will be unaffected by switching to a new diet. The solution is to do an equivalence test. There are three types of test.

In the full equivalence test, you specify a lower and an upper limit for the difference between the mean of the new treatment and the mean of the control. These define the zone within which the new treatment can be regarded as equivalent to the control. The null hypothesis is that the treatment is *not* equivalent to the control i.e. that the difference in means lies outside that zone. The test calculates t-statistics for the distance of the difference above the lower limit, and its distance below the upper limit. Their probabilities provide the evidence to assess whether the difference lies within the equivalence zone, at the lower and upper end respectively. Genstat reports the larger (i.e. the less significant) of the two probabilities together with its t-statistic. You can also check the tests by printing or plotting the confidence limits. Both tests need to be significant, and thus both ends of the confidence interval must be within the zone, to conclude that the treatments are equivalent.

In the non-inferiority test, the difference between the mean of the treatment and the mean of the control must not be less than a (negative) limit. Any positive difference is acceptable, and a negative difference must be greater than the limit. The null hypothesis is that the treatment is *inferior* to the control i.e. that the difference is less than the limit. There is just one t-statistic, assessing whether the difference is greater than the limit, and the confidence interval is unbounded at the positive end.

Similarly, in the non-superiority test, the difference between the mean of the treatment and the mean of the control must not be greater than a (positive) limit. Any negative difference is acceptable, and a positive difference must be less than the limit. The null hypothesis is that the treatment is *superior* to the control i.e. that the difference greater than the limit. There is just one t-statistic, assessing whether the difference is less than the limit, and the confidence interval is unbounded at the negative end.

To illustrate how this works, we might assume that the diets `b` - `e` in the Rat example represent different delicious "treats" added to the control diet `a`, and we want to check that these will not lead to an undue amount of extra weight. To open the menu we click on the Equivalence tests button on the ANOVA Further Output menu (Figure 1.19).

We have selected non-superiority as the Type of test, and decided that an increase of up to 2 would be acceptable. We are comparing diet means, and the control treatment is a.

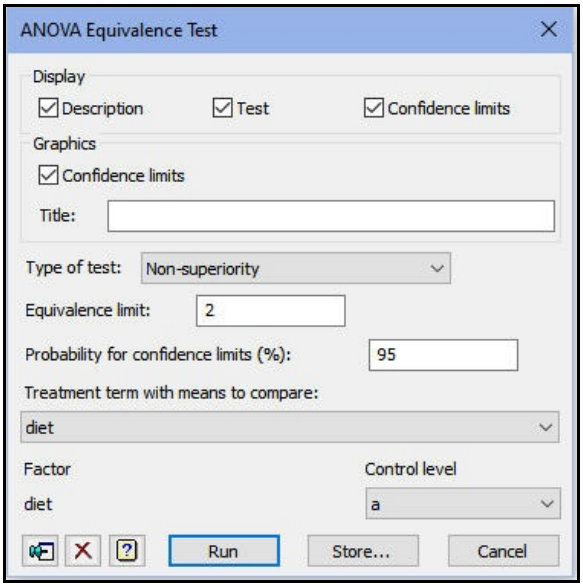


Figure 1.24

The output shows that the difference of 0.5 between the estimated mean of treatment b and that of the control a is significantly less than the limit. So it can be concluded that treatment b is not superior to the control. Alternatively, although the difference between the estimated mean of treatment c and that of control is less than 2, there is a probability of 0.18 under the null hypothesis that the difference is greater than 2. So we cannot come to the same conclusion for c (nor for the other two treatments). The confidence limits are plotted in Figure 1.25.

Test for non-superiority

Control: diet a.
Control mean: 80.58
Bound for equivalence: 2.00

diet	t statistic	Probability
a	Control	...
b	2.982	0.0047
c	0.944	0.1800
d	0.099	0.4611
e	0.050	0.4805

95% confidence intervals for difference from control

diet	Difference	Lower 95%	Upper 95%
a	Control	...	...
b	0.50	...	1.382
c	1.52	...	2.407
d	1.95	...	2.832
e	1.97	...	2.857

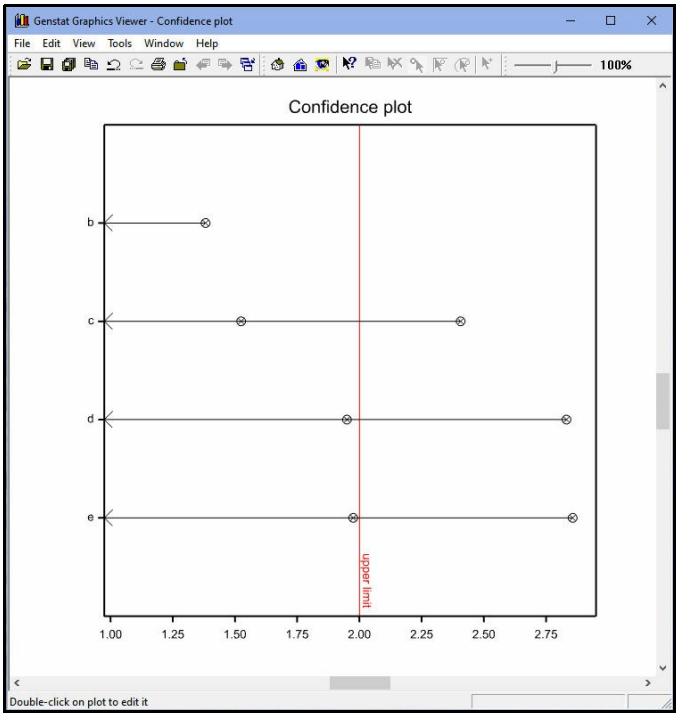


Figure 1.25

1.11 Practical

For the types of gasoline in Practical 1.7, do a non-inferiority test to assess whether gasolines A - D can be regarded as acceptably similar to gasoline E, assuming that we are willing to accept a difference of up to 1.5. (Hint: remember that, for a non-inferiority test, the limit must be negative.)

1.12 Completely randomized designs

The examples in this Chapter are analysed as though the data has come from a *completely randomized design*. In these designs, the units are assumed to have no special structure, and they are allocated at random to the sets to receive each treatment. This can be done, for example, using tables of randomized numbers: select $\sum n_i$ random numbers, allocate units with the n_1 smallest values to the first treatment, the units with the next n_2 smallest to treatment 2, and so on.

When considering how many replicates to use, it is useful to remember the formula for the standard error for the difference between two means:

s.e.d. $= \sqrt{\{}$ (residual-mean-square) $\times (n_1 + n_2) / (n_1 \times n_2) \}$

Usually it will be appropriate to have the same replication for each treatment. The main exception to this is that extra replicates are usually added for control treatments when the main interest is in comparing the other treatments with the control.

We explain later how to use Genstat's design and randomization menus to assess how many replicates are needed, and set up the design automatically.

2 Blocking structures

The *blocking structure* of an experiment is used to describe the underlying structure of the "experimental units", which are the smallest items on which the experiment is done. For example, the experimental units might be the subjects in a medical experiment, the plots of a field experiment, or the individual plants in a glasshouse experiment.

In this chapter you will learn

- how to improve the precision of an experiment by grouping the units into similar sets called "blocks"
- how randomization can avoid bias by guarding against unforeseen differences amongst the units
- how to design and analyse a complete randomized block design
- how to recognise situations that may require more than one type of blocking
- how to design and analyse a Latin square design ★

Note: the topics marked ★ are optional.

2.1 Completely randomized designs

In the simplest case, no formal structure is imposed on the units and treatments are just allocated to units at random (we will look later at how this is done in practice). This is called a *completely randomized* design.

One of the assumptions behind a completely randomized design is that the set of units to which the treatments are applied are effectively identical. For example:

- in a field experiment, that there are no systematic differences in the underlying fertility, drainage etc. of the plots;
- in a glasshouse, it assumes that the light and temperature are the same for each row of pots;
- in a factory, that the workforce behaves in essentially the same way at different times of day, days of the week and so on;
- in educational studies, that children in different schools are approximately the same, or students studying different subjects at Universities, or in different year groups etc.

Many of the designs that people use in practice are of this type. However, as we shall see, we can often obtain substantial improvements in precision and efficiency by studying the structure of the experimental units, and defining the block structure accordingly.

2.2 Randomized block designs

There are some situations where it is obvious that the units are non-uniform. For example, if a field experiment is laid out on a slope, plots at the top of the slope may be "better" than plots at the bottom. Several problems can then arise.

1. The random allocation of treatments to plots may not seem "fair". For example, all the replicates of treatment A may be allocated to "good" plots whilst all replicates of treatment B might be allocated to "bad" plots. If there was no difference between A and B, this allocation of plots could lead to treatment A appearing to be much better than treatment B.
2. The differences between plots will increase the residual sum of squares, and hence the estimate of the random variability (the variance σ^2). This means that the treatment differences must be larger to give a significant F-test and standard errors of differences between treatments will be larger, i.e. the experiment will give less precise results.

When you know that there are differences between units, you can avoid bias and improve precision by grouping (or *blocking*) the units into homogenous groups i.e. groups of units that are effectively identical. The simplest situation is the complete randomized-block design. Here

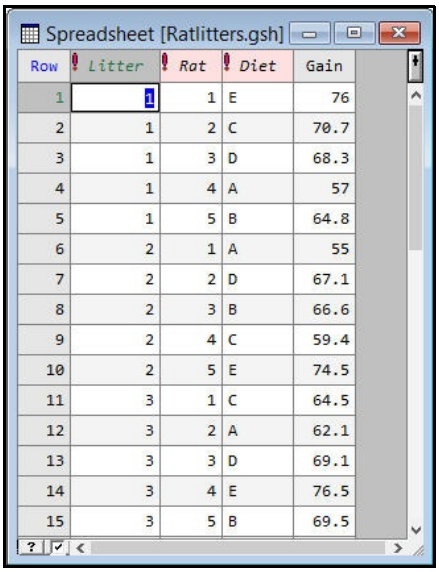
- there is a single grouping factor, usually known as *blocks*;
- each block has the same number of units, usually one for each treatment;
- within each block, the treatments are allocated randomly to the units.

Consider the field experiment described above. Suppose this experiment is designed to test the effect of four treatments A, B, C and D on the yield of winter wheat. The experiment is laid out in three rows along the side of a hill.

					↑
Block 1	D 4.6	A 7.3	C 5.5	B 6.3	U P
Block 2	A 6.6	C 5.4	D 4.1	B 5.9	H I
Block 3	B 5.6	D 3.5	C 4.9	A 6.0	L L

The treatment occurs exactly once in each block. So, provided the units within each block genuinely are similar, the allocation of treatments to units will be fair overall. Here the need for blocking seems clear: the yields from plots at the top of the slope can reasonably be expected to be larger than the yields from plots at the bottom of the slope.

Other situations may require more thought, while others may be more under your own control. For example you might decide to run an industrial experiment on several days, and use blocking to remove any systematic differences between days. You do not need to know exactly what these differences might be (temperature? humidity? motivation of the workforce?), merely that they are likely to occur – and be greater than those that occur within a day. As we shall see later, the analysis will show whether you have selected the criteria for blocking successfully.


Spreadsheet [Ratlitters.gsh]

Row	Litter	Rat	Diet	Gain
1	1	1	E	76
2	1	2	C	70.7
3	1	3	D	68.3
4	1	4	A	57
5	1	5	B	64.8
6	2	1	A	55
7	2	2	D	67.1
8	2	3	B	66.6
9	2	4	C	59.4
10	2	5	E	74.5
11	3	1	C	64.5
12	3	2	A	62.1
13	3	3	D	69.1
14	3	4	E	76.5
15	3	5	B	69.5

Figure 2.1

The easiest situation is when the grouping is an innate characteristic of the experimental units. Spreadsheet file Ratlitters.gsh contains data from another rat-feeding experiment (John & Quenouille, 1977, *Experiments Design and Analysis*, page 32). This has eight litters, each with five rats. Rats from the same litter can reasonably be assumed to be more similar than rats from different litters. So the experiment was set up with litters acting as blocks i.e. the five diets (A-E) were allocated at random to the five rats within each litter.

The advantage of the blocking can be demonstrated by comparing the analysis taking blocks into account with the analysis ignoring blocks. First we analyse the experiment ignoring blocks, and analyse the data as if the experiment were completely randomized (Figure 2.2).

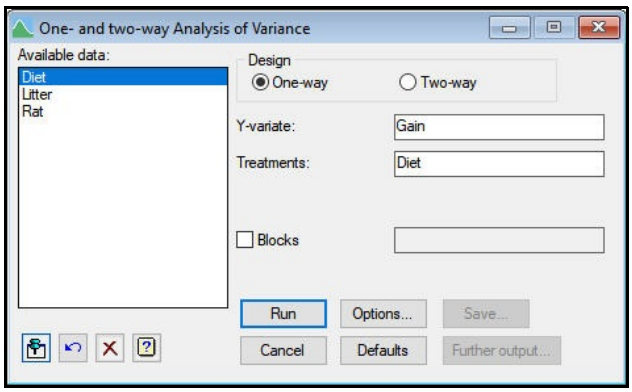


Figure 2.2

Analysis of variance

Variate: Gain

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Diet	4	346.9	86.7	0.42	0.794
Residual	35	7237.2	206.8		
Total	39	7584.1			

Tables of means

Variate: Gain

Grand mean 65.3

Diet	A	B	C	D	E
	62.6	65.4	64.2	63.3	70.9

Standard errors of differences of means

Table	Diet
rep.	8
d.f.	35
s.e.d.	7.19

Now we repeat the analysis, checking the Blocks box to show that there is a block factor, and entering specifying `Litter` in the box alongside.

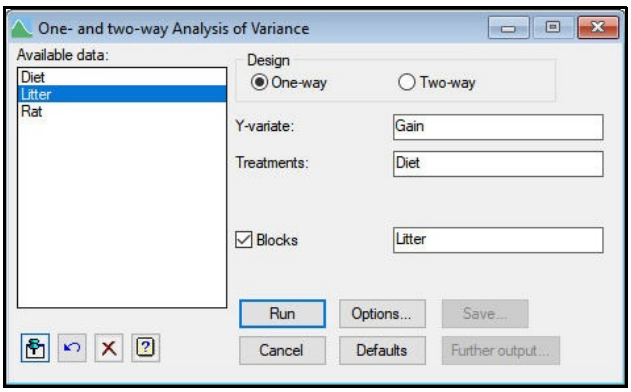


Figure 2.3

Analysis of variance

Variate: Gain

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Litter stratum	7	6099.47	871.35	21.44	
Litter.*Units* stratum					
Diet	4	346.87	86.72	2.13	0.103
Residual	28	1137.73	40.63		
Total	39	7584.07			

Tables of means

Variate: Gain

Grand mean 65.3

Diet	A	B	C	D	E
	62.6	65.4	64.2	63.3	70.9

Standard errors of differences of means

Table	Diet
rep.	8
d.f.	28
s.e.d.	3.19

The analysis of variance now has an additional line "`Litter stratum`" that records the variation between the complete litters of rats. (Diets are now estimated in the `Litter.*Units*` stratum, which represents the variation within litters.) The between-litter sum of squares (6099.47) has been subtracted from the original residual sum of

squares. So the residual sum of squares is now 7237.2 – 6099.47 = 1137.73. As a result, the residual mean square has decreased from 206.8 to 40.63, and the standard error for differences between the diet means has decreased from 7.19 to 3.19. This increase in *precision* means that we have a better chance of detecting differences between the diets. In fact, as you can see, the probability for the variance ratio of diet has decreased from 0.794 to 0.103 (still not significant, but getting closer!). You can see that the precision has improved from the fact that the variance ratio for the `Litter` stratum is greater than one – this indicates that the degrees of freedom that we have taken out of the original residual have more variability than those that are left.

Informally, blocking can be seen as a sort of insurance against large variation between groups of units which could increase your estimate of background variability, making it harder to detect treatment differences. In general, you don't have to know for certain that differences between groups will exist before you use blocks. If you suspect that certain groups of units may differ from each other, you should use those groups as a blocking factor. If the differences do appear, your estimated treatment effects will be more precise than if you had not used the blocks; if they don't, then generally you will be no worse off. Blocks most commonly correspond to position: units situated together will be subject to the same conditions and are therefore put into the same blocks.

You should also use your blocks to guard against differences introduced by the experimental procedure or husbandry of a field experiment. For example, you should make sure that the harvesting of a field experiment is done by blocks so that any differences due to harvesting time (or different machines) are accounted for by differences between blocks. Similarly, if subjective data (e.g disease scores) are to be collected by several observers, you should make sure that each observer collects data from a whole block so that differences between observers are accounted for by differences between blocks.

You will be at a disadvantage from using blocking only if you have got the blocks wrong, so that units within blocks are dissimilar. For example, if the field experiment discussed above had used blocks running down the hill rather than across the hill, units within blocks could not be considered identical. For this reason, care should be taken when forming blocks. If no obvious groups of similar units exist, a completely randomized design may be the best solution.

To generate a randomized-block design, you must first decide how many treatments are to be used in the experiment and then how many blocks, or replicates, are to be used for each treatment. Sometimes the size of your blocks may restrict the number of treatments you can test. You must use enough replicates to give a reasonable number of residual degrees of freedom, this ensures that you have a good estimate of the random error and your estimates of treatment effects will be more precise as replication increases. As a general rule, between 10 to 20 residual degrees of freedom is adequate.

Once you have decided on the number of blocks and treatments to be used, you must randomize the experiment. This means that for each block separately, you must generate a random ordering of the treatments to be applied to the units within each block. This randomization within blocks guards against any unsuspected sources of bias in the experiment. For example, for a medical experiment, it means that an experimenter could not introduce bias by giving the placebo treatment to the subjects who appeared to be least sick. If an unsuspected fertility trend ran across the hill in the field experiment we analysed earlier, then an unrandomized experiment with all blocks in order A, B, C, D

would give some treatments an unfair advantage. Randomization guards against this. However, remember that randomization should only be used to guard against *unsuspected bias* – if you have further information about differences between units within blocks, you should use this information to construct extra blocking factors.

Chapter 6 shows how this can all be done using the Genstat design menus.

2.3 Practical

Spreadsheet file `Wheatstrains.gsh` contains the results from a randomized block design to assess four strains of wheat (Snedecor, *Statistical Methods*, page 209). Analyse the experiment, and give your assessment of whether the blocking was worthwhile.

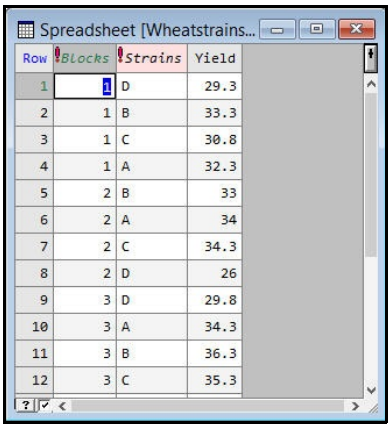


Figure 2.4

2.4 Blocking in two directions: Latin square designs

In some situations, we may need to consider blocking in two directions at once. Suppose that we want to run an experiment on pot plants in a glasshouse where there is a door in the east wall which may give rise to temperature differences. The experiment is arranged in rows facing the door. Suppose also that the glasshouse runs east-west, so that sunlight appears mainly from one side, the south.

| DOOR | North →

↓ ↓ Temperature gradient ↓ ↓

S	E	A	D	B	F	C
→	B	F	C	E	A	D
U	F	B	E	C	D	A
→	A	D	B	F	C	E
N	C	E	A	D	B	F
→	D	C	F	A	E	B

The pots on the south side of the glasshouse may receive more direct light than pots on the north side. So we need to have blocking in two directions: north-south and east-west.

One possibility here would be to use a *Latin square* design. This is

- a design for t treatments
- arranged in t rows and t columns (giving t^2 units)
- each treatment occurs exactly once in each row and once in each column

(You can check that the design above has these properties.)

Position effects that run in opposite directions are only one example of a situation where a Latin Square design is useful. Other situations include blocking for

- weekday × time-of-day,
- school × year-group,
- factory × weekday,
- time × location,

and so on.

Spreadsheet file `CC122.gsh` in Figure 2.5 contains data from an example on page 122 of Cochran & Cox (1957) *Experimental Designs* (second edition). In this experiment, six samplers were asked to assess the height of plants of wheat. The first blocking factor came about because there were six different areas to assess. The second was set up because it was felt that the accuracy of the samplers might vary during the experiment. So, the row factor of the square is `Areas`, and the column factor is `Orders`. The treatment factor is `Samplers`, and the variate for analysis `Height` is the difference between the sampler's assessment and the true mean height of the plants in the area concerned.

Spreadsheet [CC122.gsh]

Row	Areas	Orders	Samplers	Height
1	1	1	6	3.5
2	2	1	2	4.2
3	3	1	1	6.7
4	4	1	4	6.6
5	5	1	3	4.1
6	6	1	5	3.8
7	1	2	2	8.9
8	2	2	6	1.9
9	3	2	4	5.8
10	4	2	1	4.5
11	5	2	5	2.4
12	6	2	3	5.8
13	1	3	3	9.6
14	2	3	5	3.7
15	3	3	6	-2.7
16	4	3	2	3.7
17	5	3	4	6

Figure 2.5

The analysis can be produced by selecting the Latin square option for the Design drop-down list in the general Analysis of Variance menu (Figure 2.6). In the analysis of variance below, you can see that the variation between areas and between times of assessment have both been removed, thus increasing the precision with which the sampler effects are estimated.

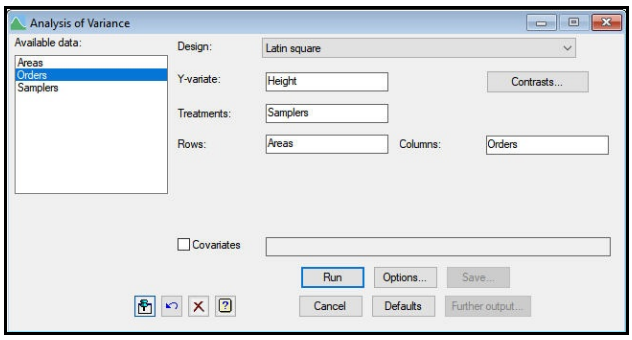


Figure 2.6

Analysis of variance

Variate: Height

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Areas stratum	5	78.869	15.774	4.74	
Orders stratum	5	28.599	5.720	1.72	
Areas.Orders stratum					
Samplers	5	155.596	31.119	9.35	<.001
Residual	20	66.563	3.328		
Total	35	329.627			

Message: the following units have large residuals.

Areas 5 Orders 6	3.40	s.e.	1.36

Tables of means

Variate: Height

Grand mean 4.76

Samplers	1	2	3	4	5	6
	6.07	5.58	6.12	6.92	2.67	1.20

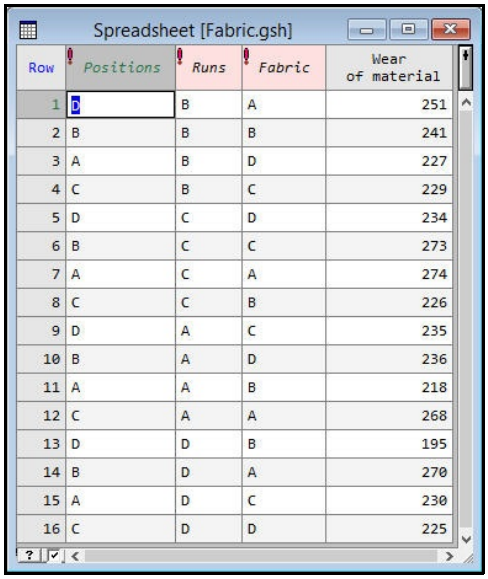
Standard errors of differences of means

Table	Samplers
rep.	6
d.f.	20
s.e.d.	1.053

The advantages of a Latin square design are similar to those of a randomized-block design, namely, you are able to estimate treatment effects more precisely by removing variation between blocking factors, while the structure of the design ensures that treatments are spread fairly over the different units. The difference is firstly that a Latin Square design allows you to take two independent blocking factors into account, and secondly, that the number of treatments is constrained to be the same as the numbers of rows and columns.

2.5 Practical

Spreadsheet file `Fabric.gsh` contains the results from an experiment that used a Latin square design to assess the wear characteristics of four different rubber-covered fabrics. The column factor of the square corresponds to four different runs, and the row factor corresponds to four positions on the testing machine used to generate wear under simulated natural conditions. (data from page 164 of Davies 1954, *Design and Analysis of Industrial Experiments*.) Analyse the results.



Row	Positions	Runs	Fabric	Wear of material
1	B	B	A	251
2	B	B	B	241
3	A	B	D	227
4	C	B	C	229
5	D	C	D	234
6	B	C	C	273
7	A	C	A	274
8	C	C	B	226
9	D	A	C	235
10	B	A	D	236
11	A	A	B	218
12	C	A	A	268
13	D	D	B	195
14	B	D	A	270
15	A	D	C	230
16	C	D	D	225

Figure 2.7

The variate `Wear` has a description "of material" associated with it. (You can see how to define one of these, by putting the cursor into the `Wear` column of the spreadsheet, and clicking on Spread on the menu bar, followed by Column and then Rename.) Notice how the description is appended to the variate name in the output, to provide additional annotation.

3 Treatment structure

So far we have considered only very straightforward situations, where the treatments do not have any special structure. More interesting investigations may have several different *types* of treatment. For example, we may have several different drugs to study, and we may also want to try a range of different doses; or we may want to try the effect of varying the amounts of several different types of fertiliser; or we may wish to study different varieties of wheat using a range of different types of fungicide to control eyespot. Each of these types of treatment should be represented by a different treatment *factor*, with *levels* defined to represent the various possibilities. For example:

Drug – levels *Morphine*, *Amidone*, *Phenadoxone*, *Pethidine*;
Dose – levels 2.5, 5, 10, 15;
Nitrogen – levels 0, 50, 100, 150;
Phosphate – levels 50, 100;
Fungicide – levels *Carbendazim*, *Prochloraz*;
Amount – levels 2, 3, 4.

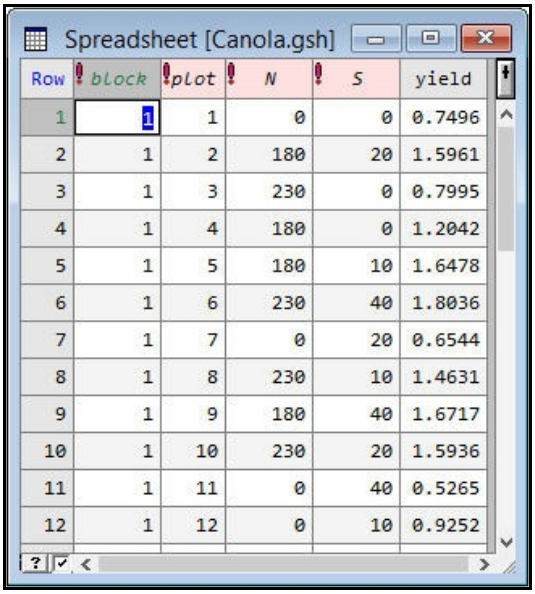
In this chapter you will learn

- how to recognise the need for more than one treatment factor
- how to analyse designs with two treatment factors using the One- and two-way Analysis of Variance menu
- how to define and interpret interactions between factors
- how to analyse designs with two treatment factors using the general Analysis of Variance menu ★
- how to use the Anova Contrasts menu ★
- how to estimate comparisons between the levels of a treatment factor ★
- how to interpret interactions between treatment contrasts ★
- the use of *model formulae* to define the treatment terms to be fitted
- how to include control treatments in a factorial experiment ★
- the use of covariates to improve precision by using additional background information about the experimental units, that was not used for blocking ★

Note: the topics marked ★ are optional.

3.1 Factorial designs with two treatment factors

One of the great advantages of analysis of variance is that it allows you to examine several different treatment factors at once. Suppose that we have an experiment on canola (oil-seed rape) with two treatment factors, `N` (nitrogen) and `S` (sulphur), in a randomized-block design (factor `block`) with three blocks and twelve plots (factor `plot`) per block. The data are available in Genstat spreadsheet file `Canola.gsh` (Figure 3.1).



Row	block	plot	N	S	yield
1	1	1	0	0	0.7496
2	1	2	180	20	1.5961
3	1	3	230	0	0.7995
4	1	4	180	0	1.2042
5	1	5	180	10	1.6478
6	1	6	230	40	1.8036
7	1	7	0	20	0.6544
8	1	8	230	10	1.4631
9	1	9	180	40	1.6717
10	1	10	230	20	1.5936
11	1	11	0	40	0.5265
12	1	12	0	10	0.9252

Figure 3.1

This is a two-way analysis of variance in randomized blocks, which can be analysed by the One- and two-way Analysis of Variance menu. Figure 3.2 shows the menu with all the relevant fields filled in, and the resulting output is shown below.

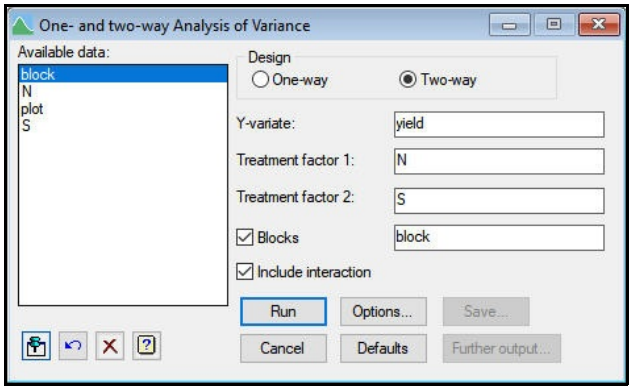


Figure 3.2

Analysis of variance

Variate: yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
block stratum	2	0.30850	0.15425	3.44	
block.*Units* stratum					
N	2	4.59223	2.29611	51.22	<.001
S	3	0.97720	0.32573	7.27	0.001
N.S	6	0.64851	0.10808	2.41	0.061
Residual	22	0.98625	0.04483		
Total	35	7.51269			

Tables of means

Variate: yield

Grand mean 1.104

N	0	180	230
	0.601	1.313	1.398

S	0	10	20	40
	0.829	1.155	1.167	1.266

N	S	0	10	20	40
0		0.560	0.770	0.524	0.552
180		0.894	1.289	1.525	1.545
230		1.032	1.404	1.454	1.700

Standard errors of differences of means

Table	N	S	N S
rep.	12	9	3
d.f.	22	22	22
s.e.d.	0.0864	0.0998	0.1729

Genstat has represented the grain yield y, recorded on the experimental plots, by the model

$$y_{ijk} = \mu + \beta_i + n_j + s_k + ns_{jk} + \varepsilon_{ijk}$$

This model is an extension of the one-way analysis discussed earlier except that now we have a term

β_i to represent the effect of blocks (`block stratum` in the aov table),

and three *terms* to represent the effects of the treatments. The parameters

n_j represent the *main effect* of nitrogen (`N`)

s_k represent the *main effect* of sulphur (`S`), and

ns_{jk} represent the *interaction* between nitrogen and sulphur (`N.S`).

Just as in the one-way analysis, the analysis of variance essentially fits each term in turn, to allow you decide how complicated a model is required to describe the results of the experiment. The analysis-of-variance table has a line for each of these, to allow you to assess whether the corresponding parameters are needed in the model. The full model, above, will estimate the *fitted values* for sulphur and nitrogen (the values predicted by the model) as

S×N means	N0	N180	N230
S0	0.560	0.894	1.032
S10	0.770	1.289	1.404
S20	0.524	1.525	1.454
S40	0.552	1.545	1.700

=

μ
1.104

+

S	
S0	−0.276
S10	0.051
S20	0.063
S40	0.162

+

N: N0	N180	N230
−0.503	0.209	0.294

+

N.S	N0	N180	N230
S0	0.234	−0.144	−0.090
S10	0.118	−0.075	−0.044
S20	−0.141	0.148	−0.007
S40	−0.211	0.071	0.141

A model like this, where you are fitting factors and their interactions, is called a *factorial* model. Here we have a 4×3 factorial.

It will be much easier to describe what is happening if there is no interaction. The model will then be

$$y_{ijk} = \mu + \beta_i + n_j + s_k + \varepsilon_{ijk}$$

leading to fitted values

N×S means	N0	N180	N230
S0	0.326	1.038	1.122
S10	0.652	1.364	1.448
S20	0.665	1.377	1.461
S40	0.763	1.475	1.559

=

μ
1.104

+

S	
S0	−0.276
S10	0.051
S20	0.063
S40	0.162

+

N: N0	N180	N230
−0.503	0.209	0.294

and you will see that we can decide on the best level of nitrogen without needing to consider how much sulphur is to be applied, and on the best level of sulphur without needing to think about the level of nitrogen on the plot. This is what we mean by saying that the two factors do not interact: the *interaction* assesses the way in which the changes in yield caused by the various levels of nitrogen differ according to the amount of sulphur or, equivalently, the way in which the response to amount of sulphur differs according to the level of nitrogen. Figure 3.3 plots the means for the model with an interaction, and Figure 3.4 plots those for the model with no interaction. When there is no interaction the lines are "parallel".

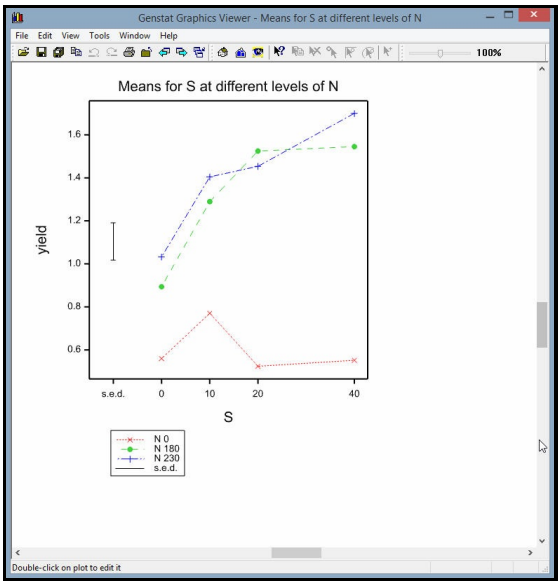


Figure 3.3

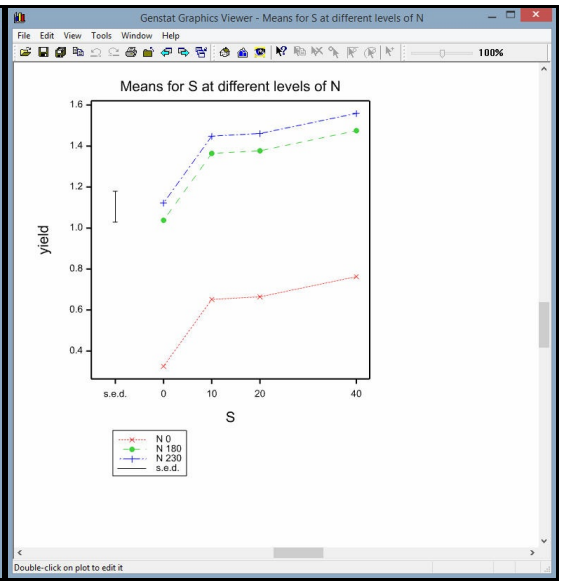


Figure 3.4

This affects the way the conclusions of the experiment are described in a resulting paper or report: if there was an interaction you might need to write, for example "for low and high levels of sulphur, the yields improved linearly with increasing levels of nitrogen, whereas for sulphur at 10kg they seemed to level off above 180kg of nitrogen". If there was no interaction this might become "application of 10kg sulphur improved yields but there seemed to be no further benefit from higher amounts; yields increased linearly with nitrogen, irrespective of the amount of sulphur". It also affects the tables or figures that should be presented. If there is an interaction, you will need to present the two-way table of means (nitrogen × sulphur); that is, you will need to present their effects jointly. If there is no interaction, you can simply present the one-way table for each of the main effects that is needed in the model.

A plot like Figure 3.3 may help to explain the interaction, or even suggest a way of modelling it. We shall explore these ideas further in the next section.

3.2 Fitting contrasts

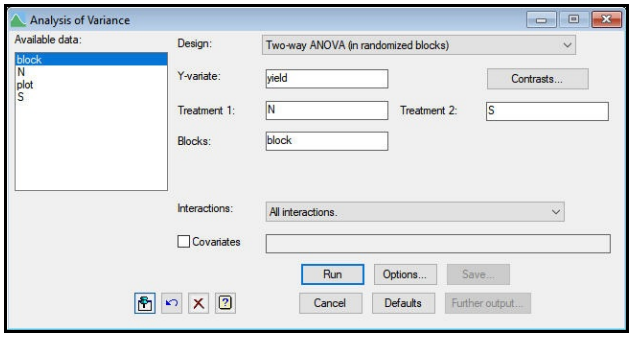


Figure 3.5

Sometimes there may be comparisons between the levels of a treatment factor that you are particularly keen to assess. For example, you might have had an initial suspicion that there would be little difference between the 180 and 230 levels of nitrogen in the previous section, but similar (and larger) differences between 0 and 180, and between 0 and 230. You might then want to fit a single mean for the 180 and 230 levels of nitrogen, and assess the *contrast* between this value and the mean for level 0.

As we have already seen, in Section 1.6, you can do this by using the general Analysis of Variance menu (Figure 3.5), instead of the One- and two-way Analysis of Variance menu.

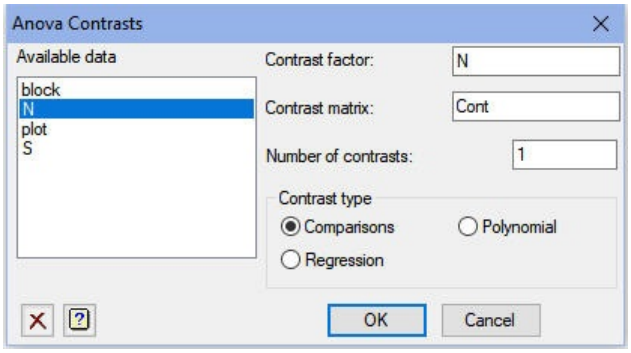


Figure 3.6

To define the contrasts, you click on the Contrasts button to open the ANOVA Contrasts menu. The Contrast factor and Contrast type fields in the menu shown in Figure 3.6, indicate that we want to assess *comparisons* between the levels of the factor N, and the Number of contrasts field indicates that we want to fit one contrast.

When we click on OK, a Genstat spreadsheet appears (Figure 3.7) containing the contrast matrix Cont whose name was specified in the Contrast matrix field; this name was selected automatically by the ANOVA Contrasts menu, but you can specify your own name if you prefer, or if you have already formed a suitable matrix. You use the spreadsheet to specify the coefficients that define the comparison. In Figure 3.7, the matrix defines the comparison:

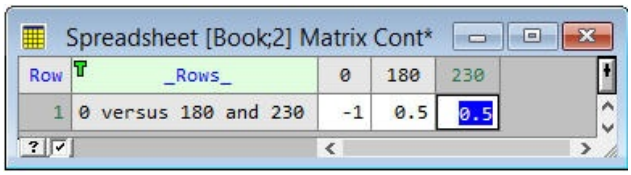


Figure 3.7

$$(N_{180} + N_{230}) / 2 \; - \; N_0$$

Notice that you can also define names for the contrasts, using the Rows column.

Back in the Analysis of Variance menu (Figure 3.8) you can see that the Treatment 1 field now contains a *function* of N, namely COMP(N;1;Cont). The syntax of these functions is described in Section 3.4.

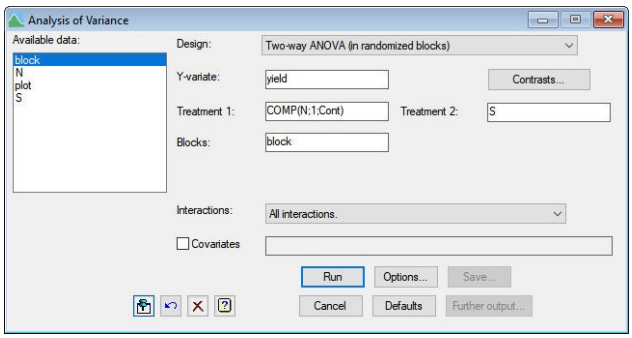


Figure 3.8

There is a box controlling the printing of contrasts in the Display section of the ANOVA options menu (obtained as usual by clicking on the Options button in the main Analysis of Variance menu). In Figure 3.9, we have checked this together with the AOV table and F-probabilities boxes. These request the output below.

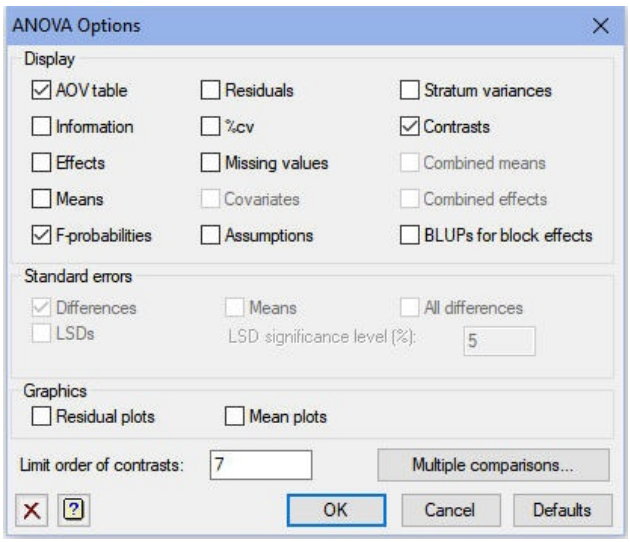


Figure 3.9

Analysis of variance

Variate: yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
block stratum	2	0.30850	0.15425	3.44	
block.*Units* stratum					
N	2	4.59223	2.29611	51.22	<.001
0 versus 180 and 230	1	4.54954	4.54954	101.48	<.001
S	3	0.97720	0.32573	7.27	0.001
N.S	6	0.64851	0.10808	2.41	0.061
0 versus 180 and 230.S	3	0.59907	0.19969	4.45	0.014
Residual	22	0.98625	0.04483		
Total	35	7.51269			

Tables of contrasts

Variate: yield

block.*Units* stratum

N contrasts

0 versus 180 and 230 0.754, s.e. 0.0749, ss.div. 8.00

N.S contrasts

0 versus 180 and 230.S, e.s.e. 0.150, ss.div. 2.00

S	0	10	20	40
	-0.35	-0.18	0.21	0.32

Notice that, in the analysis-of-variance table, the line for the main effect `N` is now accompanied by a line entitled "`0 versus 180 and 230`" giving the degrees of freedom, sum of squares and so on for that comparison. In addition the `N.S` interaction is accompanied by a line "`0 versus 180 and 230.S`" which represents the interaction between the comparison and the factor `S` (that is, it measures how the size of the comparison varies according to the level of `S`).

The section headed "`Tables of contrasts`" then shows the estimate of the contrast, 0.754, with standard error 0.0749. The "`ss. div`" value is analogous to the replication of a table of means or effects: it is the divisor used in calculating the estimated values of the contrasts. This is useful mainly where there is a range of e.s.e.'s for a table of contrasts: the contrasts with the smallest values of the ss. div. are those with the largest e.s.e., and vice versa. (The ss. div. of each estimated contrast is in fact the sum of squares of the values of the coefficients used to calculate it, weighted according to the replication.) The `N.S` contrasts table shows how the overall value of the contrast varies according to the level of `S`. So, at level 0 of `S`, the estimated contrast is 0.754−0.35.

When a factor like sulphur (or nitrogen) has quantitative levels, you might want to investigate whether the yield increases linearly with the amount of sulphur (or nitrogen); you could also include a quadratic term to check for curvature in the response.

Put the cursor into the Treatment 2 box of the Analysis of Variance menu, and click on the Contrasts button to produce the Anova Contrasts menu again. To fit *polynomial* contrasts of sulphur, we select Polynomial within the Contrast type box in the ANOVA Contrasts menu, set the Contrast factor to `S`, and (for a quadratic polynomial) set the Number of contrasts to 2; see Figure 3.10. After we click on OK, the Treatment 2 box of the Analysis of Variance menu will contain the function `POL(S;2)`. If we change the setting of the Treatment 1 box back to `N`, and then click on Run, we obtain the output below.

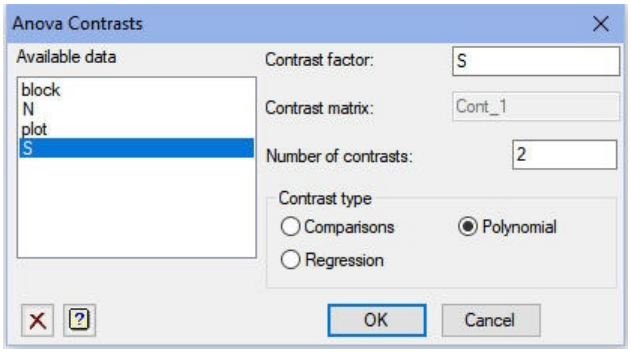


Figure 3.10

Analysis of variance

Variate: yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
block stratum	2	0.30850	0.15425	3.44	
block.*Units* stratum					
N	2	4.59223	2.29611	51.22	<.001
S	3	0.97720	0.32573	7.27	0.001
Lin	1	0.69741	0.69741	15.56	<.001
Quad	1	0.19577	0.19577	4.37	0.048
Deviations	1	0.08403	0.08403	1.87	0.185
N.S	6	0.64851	0.10808	2.41	0.061
N.Lin	2	0.52294	0.26147	5.83	0.009
N.Quad	2	0.07788	0.03894	0.87	0.433
Deviations	2	0.04769	0.02385	0.53	0.595
Residual	22	0.98625	0.04483		
Total	35	7.51269			

Tables of contrasts

Variate: yield

block.*Units* stratum

S contrasts

Lin 0.0094, s.e. 0.00239, ss.div. 7875.

Quad -0.00042, s.e. 0.000199, ss.div. 1131429.

Deviations, e.s.e. 0.0706, ss.div. 9.00

S	0	10	20	40
	-0.028	0.074	-0.055	0.009

N.S contrasts

N.Lin, e.s.e. 0.00413, ss.div. 2625.

N	0	180	230
	-0.0115	0.0058	0.0058

N.Quad, e.s.e. 0.000345, ss.div. 377143.

N	0	180	230
	0.00028	-0.00035	0.00007

Deviations, e.s.e. 0.122, ss.div. 3.00

N	S	0	10	20	40
0		-0.02	0.06	-0.05	0.01
180		0.03	-0.07	0.05	-0.01
230		0.00	0.01	-0.01	0.00

Equation of the polynomial for S

0.8561 + 0.0266 * S - 0.0004 * S**2

Equations of the polynomials for N.S

N	
0	0.6112 + 0.0035 * S - 0.0001 * S**2
180	0.8944 + 0.0469 * S - 0.0008 * S**2
230	1.0629 + 0.0295 * S - 0.0003 * S**2

In the analysis of variance, the sum of squares for sulphur is partitioned into the amount that can be explained by a linear relationship of the yields with sulphur (the line marked `Lin`), the extra amount that can be explained if the relationship is quadratic (the line `Quad`), and the amount represented by deviations from a quadratic polynomial. A cubic term would be labelled as `Cub`, and a quartic as `Quart`. You are not allowed to fit more than fourth-order polynomials. The interaction of nitrogen and sulphur is also partitioned: `N.Lin` lets you assess the effect of fitting three different linear relationships, one for each level of nitrogen; `N.Quad` assesses the effect of fitting a different quadratic contrast for each level of `N`; and the deviations line represents deviations from these quadratic polynomials. So, the analysis shows strong evidence for linear and quadratic effects of sulphur, and for interactions between these contrasts and nitrogen (as we would have expected from the plot in Figure 3.3). The tables of contrasts again provide estimates of the parameters of the contrasts. For example, the overall linear effect is 0.0094, and the effect for level 0 of nitrogen is 0.0094−0.0115

You can fit more than one set of contrasts at a time. If we had retained the nitrogen comparison, we would have obtained the output below.

Analysis of variance

Variate: yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
block stratum	2	0.30850	0.15425	3.44	
block.*Units* stratum					
N	2	4.59223	2.29611	51.22	<.001
0 versus 180 and 230	1	4.54954	4.54954	101.48	<.001
S	3	0.97720	0.32573	7.27	0.001
Lin	1	0.69741	0.69741	15.56	<.001
Quad	1	0.19577	0.19577	4.37	0.048
Deviations	1	0.08403	0.08403	1.87	0.185
N.S	6	0.64851	0.10808	2.41	0.061
0 versus 180 and 230.Lin	1	0.52294	0.52294	11.67	0.002
0 versus 180 and 230.Quad	1	0.04448	0.04448	0.99	0.330
Residual	22	0.98625	0.04483		
Total	35	7.51269			

Tables of contrasts

Variate: yield

block.*Units* stratum

N contrasts

0 versus 180 and 230 0.754, s.e. 0.0749, ss.div. 8.00

S contrasts

Lin 0.0094, s.e. 0.00239, ss.div. 7875.

Quad -0.00042, s.e. 0.000199, ss.div. 1131429.

Deviations, e.s.e. 0.0706, ss.div. 9.00

S	0	10	20	40
	-0.028	0.074	-0.055	0.009

N.S contrasts

0 versus 180 and 230.Lin 0.0173, s.e. 0.00506, ss.div. 1750.

0 versus 180 and 230.Quad -0.00042, s.e. 0.000422, ss.div. 251429.

The interaction between nitrogen and sulphur is now partitioned according to the nitrogen comparison. The line "`0 versus 180 and 230.Lin`" assesses the effect of fitting two different linear relationships, one for each level 0 of nitrogen, and one for levels 180 and 230 of nitrogen, instead of a single overall linear contrast. Similarly, the line "`0 versus 180 and 230.Quad`" represents the difference between the two quadratic contrasts. So you can define contrasts on any treatment factor, and Genstat will automatically estimate their interactions.

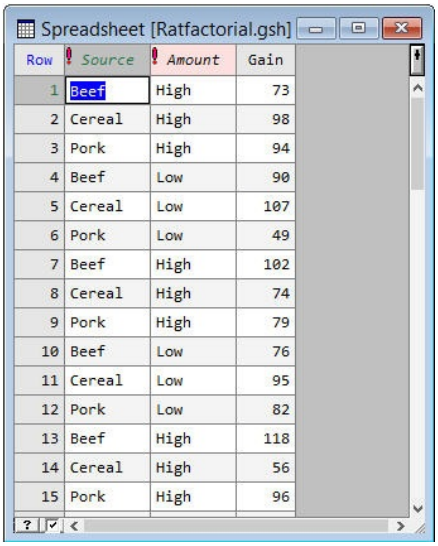
As explained in Section 1.6, to fit polynomial contrasts, Genstat calculates *orthogonal polynomials* and does a multiple regression of the effects of factor using the polynomials as x-variates. Regression contrasts are similar to polynomial contrasts, except that here you can supply your own matrix of x-variates. Genstat orthogonalizes the x-variates for you, so that each one represents the effect adding this x-variable to a model containing all the earlier ones.

3.3 Practical

Spreadsheet file Ratfactorial.gsh contains data from an experiment to study the effect of 6 different diets on the gain in weight of rats (data from Snedecor and Cochran, Statistical Methods p.305). Each diet was at either High or Low protein (factor Amount), and the protein was derived from either Beef, Cereal or Pork (factor Source).

Analyse the data as a 3×2 factorial, and assess whether there is evidence for an interaction between Amount and Source.

Fit two comparison contrasts between levels of the Source factor: Animal vs Vegetable, and Beef vs Pork.


Spreadsheet [Ratfactorial.gsh]

Row	Source	Amount	Gain
1	Beef	High	73
2	Cereal	High	98
3	Pork	High	94
4	Beef	Low	90
5	Cereal	Low	107
6	Pork	Low	49
7	Beef	High	102
8	Cereal	High	74
9	Pork	High	79
10	Beef	Low	76
11	Cereal	Low	95
12	Pork	Low	82
13	Beef	High	118
14	Cereal	High	56
15	Pork	High	96

Figure 3.11

3.4 Syntax of model formulae

The structure of the design and the treatment terms to be fitted in a Genstat analysis of variance are specified by *model formulae*. In the simpler menus, like those we have used earlier in this chapter, the formulae are constructed automatically behind the scenes. However, for the more advanced menus and analyses you will need to specify your own formulae.

Several of the menus allow you to specify any number of treatment factors, interactions and so on. So, for example, the General analysis of variance, the General treatment structure (no blocking) and the General treatment structure (in randomized blocks) menus all have a box entitled Treatment structure into which a formula (known as the *treatment formula*) needs to be entered.

The general Analysis of Variance menu also allows you to define any *underlying structure* for the design (for example completely randomized, randomized-block, split-plot, split-split-plot, and so on). This is specified by a model formula (the *block formula*) which is entered into the Block structure box; this can be left blank with unstructured (completely randomized) designs. This formula defines the strata and thus the error terms for the analysis.

In its simplest form, a model formula is a list of *model terms*, linked by the operator "+". For example,

```
A + B
```

is a formula containing two terms, A and B, representing the main effects of factors A and B respectively. *Higher-order terms* (like interactions) are specified as series of factors separated by dots, but their precise meaning depends on which other terms the formula contains, as we explain below. The other operators provide ways of specifying a formula more succinctly, and of representing its structure more clearly.

The *crossing operator* * is used to specify factorial structures. The formula

```
N * S
```

was used by Genstat to specify the two-way analysis of variance introduced in Section 3.1. This is expanded to become the formula

```
N + S + N.S
```

which has three terms: `N` for the nitrogen main effect, `S` for the main effect of sulphur, and `N.S` for the nitrogen by sulphur interaction. Higher-order terms like `N.S` represent all the joint effects of the factors `N` and `S` that have not been removed by earlier terms in the formula. Thus here it represents the interaction between nitrogen and sulphur as both main effects have been removed.

The other most-commonly used operator is the *nesting operator* (`/`). This occurs most often in block formulae. For example, the formula

```
block /  plot
```

is expanded to become the formula

```
block + block.plot
```

This specification assumes that there is no special similarity between the plot numbered 1, for example, in block 1 and plot 1 in any other block. So the formula contains no "main effect" for `plot`, and the term `block.plot` thus represents *plot-within-block* effects (that is the differences between individual plots after removing any overall similarity between plots that belong to the same block). This is similar to the block model for the randomized design in Section 2.2 except that we have the factor `plot` instead of `*Units*`.

Treatments can be nested too. For example, in a study of potential energy crops, we may want to study two varieties of Miscanthus (M_1 ... M_2) and three of Reed Canary Grass (R_1 ... R_3). We will certainly be interested in assessing overall differences between Miscanthus and Reed Canary Grass. We may also be interested in how much variation there is between Mp_1 and Mp_2, and amongst {R_1, R_2 and R_3}; that is whether there is variability of the varieties beyond the variability of the individual plants of each variety. The model of interest (assuming that there is no blocking) would then be

$$y_{ijk} = \mu + s_i + sv_{ij} + \varepsilon_{ijk}$$

where parameters

s_i represent the effects of the species (i = 1, 2), and

sv_{ij} represent the variety *within* species effects (j = 1,2 for i=1, j = 1...3 for i=2).

Notice that we do not have any term for a variety main effect – the actual number allocated to each variety does imply any special similarity for example between the strain numbered 2 for Miscanthus and the strain numbered 2 for Reed Canary Grass.

A formula can contain more than one of these operators. The three-factor factorial model

```
A * B * C
```

becomes

```
A + B + C + A.B + A.C + B.C + A.B.C
```

The interaction `A.B.C` then assesses whether the joint effects of factors `A` and `B` differ according to the level of `C` (or, equivalently, whether the joint effects of `A` and `C` differ

according to the level of B, and so on).

The nested structure

```
block / wplot / subplot
```

which occurs as the block model of a split-plot design (Section 5.1) becomes

```
block + block.wplot + block.wplot.subplot
```

The crossing and nesting operators can also be mixed in the same formula. For example, the factorial-plus-added-control study in Section 3.5 has treatment structure

```
Control / (Drug * Dose)
```

which expands to

```
Control + Control.Drug + Control.Dose + Control.Drug.Dose
```

In general, if *l* and *m* are two model formulae:

```
l * m  =  l + m + l.m
l / m  =  l + fac(l).m
```

(where *l*`.`*m* is the sum of all pairwise dot products of a term in *l* and a term in *m*, and `fac(`*l*`)` is the dot product of all factors in *l*). For example:

```
(A + B) * (C + D) = (A + B) + (C + D) + (A + B).(C + D)
                  = A + B + C + D + A.C + A.D + B.C + B.D
(A + B)/C = A + B + fac(A + B).C  =  A + B + A.B.C
```

Terms in the treatment formula can be partitioned into contrasts by specifying a function of the factor.

`COMPARISON(`*factor*`;` *scalar*`;` *matrix*`)` partitions the *factor* into the comparisons specified by the *matrix*. There is a row of the matrix for each comparison, and the *scalar* specifies how many of them are to be fitted.

`POL(`*factor*`;` *scalar*`;` *variate*`)` partitions the *factor* into polynomial contrasts (linear, quadratic and so on). The *scalar* gives the maximum order of contrast (1 for linear only, 2 for linear and quadratic, and so on) and the *variate* gives a numerical value for each level of the factor. If the variate is omitted, the levels defined when the factor was declared will be used.

`REG(`*factor*`;` *scalar*`;` *matrix*`)` partitions the *factor* into the (user-defined) regression contrasts specified by the coefficients in each row of the *matrix*. The *scalar* defines the number of contrasts to be fitted.

3.5 Factorial plus added control

One important model that includes crossing and nesting is the *factorial plus added control* structure. For example, suppose we have four different fumigants used to control nematodes (*CN*, *CS*, *CM* and *CK*), which we wish to try at two levels (*single* and *double*), and that we also want to include a control treatment (*none* = no fumigant at any dose). The control represents a "zero" level for both factors, and the factorial structure of `Type` × `Amount` operates only when some sort of fumigant has been applied. The table below indicates which combinations of `Type` and `Amount` are feasible, and also shows the extra factor `Fumigant` that is necessary to define the model.

Fumigant	Amount	Type *none*	*CN*	*CS*	*CM*	*CK*
not fumigated	*none*	✔	✘	✘	✘	✘
fumigated	*single*	✘	✔	✔	✔	✔
fumigated	*double*	✘	✔	✔	✔	✔

In Genstat terms, we need a model

```
Fumigant / ( Amount * Type )
```

in which the factorial structure `Amount * Type` is nested within the factor `Fumigant` (in fact `Amount` and `Type` have their factorial structure only within the *fumigated* level of `Fumigant`). The model expands to

```
Fumigant + Fumigant.Amount + Fumigant.Type +
Fumigant.Amount.Type
```

in which

`Fumigant` represents the overall effect of any fumigant at any (non-zero) dose,

`Fumigant.Amount` represents the comparison between *single* and *double* doses (averaged over the different types),

`Fumigant.Type` represents overall differences between types (averaged over single and double doses), and

`Fumigant.Amount.Type` represents the interaction between `Amount` and `Type` (given that some sort of fumigant has been applied).

Results of the experiment, a classic study carried out at Rothamsted in 1935, are available in spreadsheet file `Nematode.gsh` (also see Cochran & Cox 1957, *Experimental Designs*, page 46). As it is thought that effects will proportionate the Calculate menu (Figure 3.12) is used to transform the counts to logarithms. Transformations are discussed further in Chapter 4.

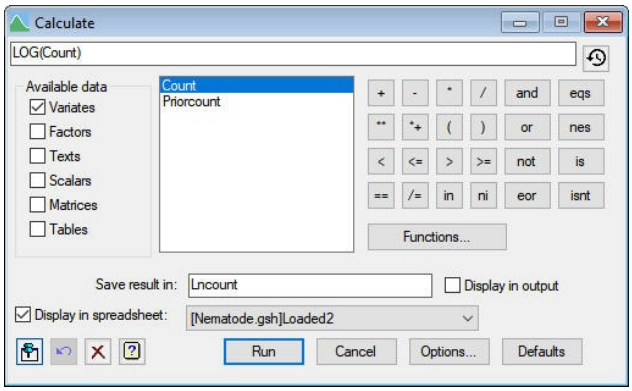


Figure 3.12

The analysis can be done by selecting the General treatment structure (in randomized blocks) setting of the Design drop-down list box in the general Analysis of Variance menu (Figure 3.13). There is now a Treatment structure box, in which we can define any treatment model, using the syntax explained in Section 3.4).

The resulting output is shown below.

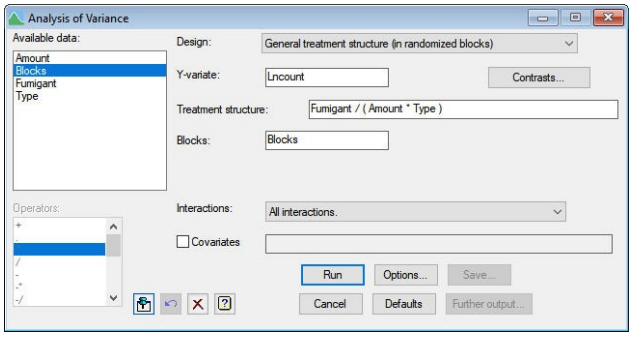


Figure 3.13

Analysis of variance

Variate: Lncount

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Blocks stratum	3	5.5727	1.8576	7.80	
Blocks.*Units* stratum					
Fumigant	1	1.0186	1.0186	4.28	0.046
Fumigant.Amount	1	0.0028	0.0028	0.01	0.915
Fumigant.Type	3	1.5153	0.5051	2.12	0.114
Fumigant.Amount.Type	3	0.2471	0.0824	0.35	0.792
Residual	36	8.5688	0.2380		
Total	47	16.9253			

Tables of means

Variate: Lncount

Grand mean 5.582

Fumigant	Not fumigated	Fumigated
	5.788	5.479
rep.	16	32

Fumigant	Amount	None	Single	Double
Not fumigated		5.788		
Fumigated			5.488	5.469

Fumigant	Type	None	CN	CS	CM	CK
Not fumigated		5.788				
	rep.	16				
Fumigated			5.529	5.153	5.763	5.470
	rep.		8	8	8	8

Fumigant	Amount	Type	None	CN	CS	CM	CK
Not fumigated	None		5.788				
		rep.	16				

Fumigated	Single		5.483	5.280	5.818	5.371
		rep.	4	4	4	4
	Double		5.575	5.026	5.707	5.570
		rep.	4	4	4	4

Standard errors of differences of means

Table	Fumigant	Fumigant Amount	Fumigant Type	Fumigant Amount Type	
rep.	unequal	16	unequal	unequal	
d.f.	36	36	36	36	
s.e.d.			0.2439	0.3450	min.rep
	0.1494	0.1725	0.2113	0.2727	max-min
			0.1725X	0.1725X	max.rep

(No comparisons in categories where s.e.d. marked with an X)

Notice that, when tables of means have unequal replication, the general Analysis of Variance menu provides three standard errors of difference for each table:

- to compare a pair of means each with the minimum replication of those in the table,
- to compare a mean with minimum replication with one with maximum replication,
- and to compare a pair of means that both have the maximum replication.

The "X" beside the standard errors of difference for maximum replication indicates that there is actually only one mean in the table with the maximum replication. So this is an unavailable comparison.

3.6 Covariates

Covariates incorporate additional quantitative information into an analysis. Sometimes you may have measurements made on the units before the experiment was carried out. This can be used to allocate the units to blocks but, even after this grouping, they may contain additional useful information. Analysis of covariance incorporates quantitative information of this sort into the analysis – providing a further way of decreasing variability.

In the example in Section 3.5, nematode counts were done prior to the experiment as well as afterwards. Analysis of covariance includes the (transformed) initial counts as a linear term in the model, rather like a regression analysis except that here we have the factors for blocks and treatments as well.

$$y_{ijkl} = \mu + \beta_i + f_j + ft_{jk} + fl_{jl} + ftl_{jkl} + b \times (x_{ijkl} - \bar{x}) + \varepsilon_{ijkl}$$

where y_{ijkl} and x_{ijkl} are the logarithms of the counts.

To do an analysis of covariance, you simply need to check the Covariates box in the Analysis of Variance menu, and enter the covariate in the box immediately to the right, as shown in Figure 3.14. If you have several covariates, you can enter them as a list (separated by commas). You can even enter a model formula: for example, you could put `Lnpriorcount.Blocks` to fit a different regression coefficient in each block.

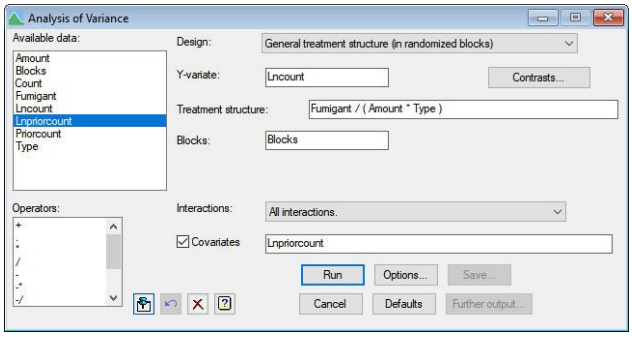


Figure 3.14

Clicking on Run in Figure 3.14 produces an analysis-of-variance table that contains extra lines to assess how much the final (log) counts depend on the initial counts, after removing the effects of treatments. The treatment effects (and s.s.) are also adjusted to take account of the fact that the plots with the various treatments had different numbers of nematodes before the experiment. This adjustment causes some loss of efficiency in the treatment estimation. The remaining efficiency is measured by the *covariance efficiency factor*, shown for each treatment term in the "`cov. ef.`" column of the analysis-of-variance table. The values are in the range zero to one. A value of zero indicates that the treatment contrasts are completely correlated with the covariates: after the covariates have been fitted there is no information left about the treatments. A value of one indicates that the covariates and the treatment term are orthogonal. Usually the values will be around 0.8 to 0.9. A low value should be taken as a warning: either the measurements used as covariates have been affected by the treatments, which can occur when the measurements on covariates are taken after instead of before the experiment; or the random allocation of treatments has been unfortunate in that some treatments are on units with generally low values of the covariates while others are on generally high ones.

For a residual line in the analysis of variance, the value in the "`cov. ef.`" column measures how much the covariates have improved the precision of the experiment. This is calculated by dividing the residual mean square in the unadjusted analysis (which excludes the covariates) by its value in the adjusted analysis.

To assess the full effect of the covariate on the estimation of a treatment term, you should multiply its covariance efficiency factor by the covariance efficiency factor of the residual with which it is to be compared. For `Fumigant.Amount` in the example, the calculation would be 0.99×2.48. So fitting the covariate has improved the precision with which `Fumigant.Amount` is estimated. You can see this in its sed (0.1097), which is equal to the earlier sed (0.1725), divided by $\sqrt{(0.99 \times 2.48)}$.

Analysis of variance (adjusted for covariate)

Variate: Lncount
Covariate: Lnpriorcount

Source of variation	d.f.	s.s.	m.s.	v.r.	cov.ef.	F pr.

Blocks stratum						
Covariate	1	4.76145	4.76145	11.74		0.076
Residual	2	0.81127	0.40563	4.23	4.58	
Blocks.*Units* stratum						
Fumigant	1	1.16420	1.16420	12.13	1.00	0.001
Fumigant.Amount	1	0.03514	0.03514	0.37	0.99	0.549
Fumigant.Type	3	2.09342	0.69781	7.27	0.92	<.001
Fumigant.Amount.Type	3	0.31977	0.10659	1.11	1.00	0.358
Covariate	1	5.21084	5.21084	54.31		<.001
Residual	35	3.35793	0.09594		2.48	
Total	47	16.92526				

Covariate regressions

Variate: Lncount

Covariate	coefficient	s.e.
Blocks stratum		
Lnpriorcount	0.54	0.157
Blocks.*Units* stratum		
Lnpriorcount	0.585	0.0794
Combined estimates		
Lnpriorcount	0.573	0.0684

Tables of means (adjusted for covariate)

Variate: Lncount
Covariate: Lnpriorcount

Grand mean 5.582

Fumigant	Not fumigated	Fumigated
	5.805	5.470
rep.	16	32

Fumigant	Amount	None	Single	Double
Not fumigated		5.805		
Fumigated			5.508	5.432

Fumigant	Type	None	CN	CS	CM	CK
Not fumigated		5.805				
	rep.	16				
Fumigated			5.798	5.220	5.667	5.195
	rep.		8	8	8	8

Fumigant	Amount	Type	None	CN	CS	CM	CK
Not fumigated	None		5.805				
		rep.	16				
Fumigated	Single			5.713	5.399	5.745	5.174
		rep.		4	4	4	4
	Double			5.882	5.041	5.589	5.216
		rep.		4	4	4	4

Standard errors of differences of means

Table	Fumigant	Fumigant Amount	Fumigant Type	Fumigant Amount Type	
rep.	unequal	16	unequal	unequal	
d.f.	35	35	35	35	
s.e.d.			0.1596	0.2226	min.rep
	0.0949	0.1097	0.1382	0.1760	max-min
			0.1129X	0.1113X	max.rep

(No comparisons in categories where s.e.d. marked with an X)

You can find more information about analysis of covariance in Genstat in the Guide to the Genstat Command Language, Part 2, Section 4.3.

3.7 Practical

Spreadsheet file `Ratmuscles.gsh` contains data from an experiment to study the effect of electrical stimulation in preventing the wasting away of denervated muscles, using rats as the subjects (Solandt, DeLury & Hunter, 1943, *Archives of Neurology & Psychiatry*, **49**, 802-807; also see Cochran & Cox, 1957, *Experimental Designs 2nd Edition*, page 176). There were three treatment factors: length of each treatment, number of treatment periods per day and the type of current. The experiment used a complete randomized block design with two blocks. The denervated muscles were the gastrocnemius muscles on one side of the rat. To improve precision, the normal muscle on the other side of each rat was also measured, for use as a covariate in the analysis.

Spreadsheet [Ratmuscles.gsh]

Row	Block	Length	Number	Type	Normal	Denervated
1	1	1	1	Galvanic	152	72
2	1	1	3	Galvanic	131	74
3	1	1	6	Galvanic	131	69
4	1	1	1	Faradic	130	61
5	1	1	3	Faradic	129	61
6	1	1	6	Faradic	126	65
7	1	1	1	60 cycle	141	62
8	1	1	3	60 cycle	112	65
9	1	1	6	60 cycle	111	70
10	1	1	1	25 cycle	147	85
11	1	1	3	25 cycle	125	76
12	1	1	6	25 cycle	130	61
13	1	2	1	Galvanic	136	67
14	1	2	3	Galvanic	110	52
15	1	2	6	Galvanic	122	62
16	1	2	1	Faradic	111	60
17	1	2	3	Faradic	180	55
18	1	2	6	Faradic	122	59

Figure 3.15

Analyse the experiment. Has the covariate improved the precision of the estimates? Which tables of means would you present in the report?

3.8 Summaries of results

When you have a complicated experiment, it may be difficult to decide what to report. The Summary of results box in the ANOVA Further Output menu provides a summary of the analysis, containing information useful for a report. It prints the name of the y-variate, the block and treatment models and any covariates. It lists the significant terms, and then it prints the relevant tables of means. These tables are those that contain significant treatment effects. Also, the tables are formed so that each one contains *all* the significant effects involving any of its factors.

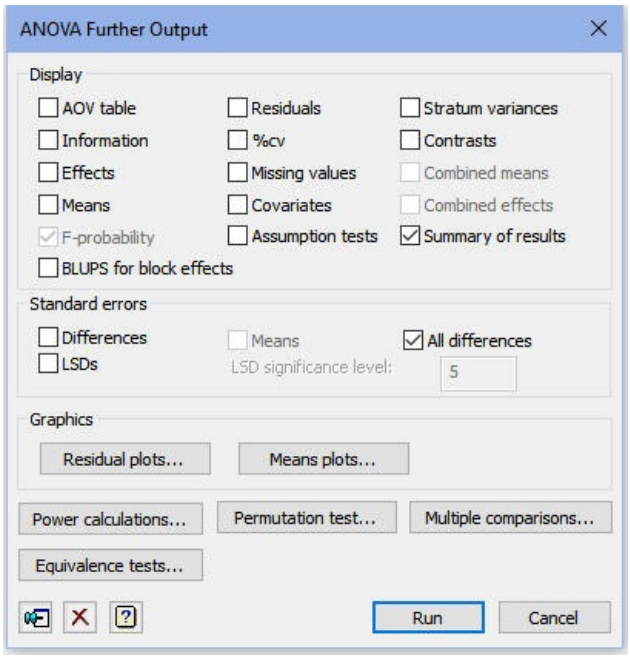


Figure 3.16

In the example in Section 3.6, `Fumigant` and `Fumigant.Type` are significant. `Fumigant` is included in the two-way classified by `Fumigant` and `Type`, and so Genstat does not print the one-way table for `Fumigant`. (As the effect of `Fumigant` depends on the `Type`, it does not make sense to consider `Fumigant` on its own.)

The standard errors for differences between means in a table are not all the same. Genstat then prints them all in a triangular array, which may be easier to use than the summary usually provided with the tables of means.

Results from analysis of variance

Variate: Lncount
Treatment structure: Fumigant/Amount*Type
Block structure: Blocks
Covariates: Lnpriorcount, Priorcount
Factorial: 3

Significant treatment terms

Fumigant	1%	(pr. 0.001)
Fumigant.Type	<0.1%	(pr. <.001)

Predicted means for Fumigant.Type

Type	None	CN	CS	CM	CK
Fumigant					
Not fumigated	5.806	*	*	*	*
Fumigated	*	5.794	5.219	5.670	5.194

Standard errors of differences between means

Not fumigated, None	1	*			
Not fumigated, CN	2	*	*		
Not fumigated, CS	3	*	*	*	
Not fumigated, CM	4	*	*	*	*
Not fumigated, CK	5	*	*	*	*
Fumigated, None	6	*	*	*	*
Fumigated, CN	7	0.1408	*	*	*
Fumigated, CS	8	0.1408	*	*	*
Fumigated, CM	9	0.1408	*	*	*
Fumigated, CK	10	0.1408	*	*	*
		1	2	3	4

Not fumigated, CK	5	*			
Fumigated, None	6	*	*		
Fumigated, CN	7	*	*	*	
Fumigated, CS	8	*	*	0.1641	*
Fumigated, CM	9	*	*	0.1641	0.1641
Fumigated, CK	10	*	*	0.1641	0.1641
		5	6	7	8

Fumigated, CM	9	*	
Fumigated, CK	10	0.1641	*
		9	10

Rows and columns are labelled by the labels/levels of the factors: Fumigant and Type.

3.9 Practical

Produce a summary of the results from the analysis in Practical 3.7.

4 Checking the assumptions

In this chapter you will learn

- what assumptions are needed to ensure the validity of an analysis of variance
- why the variance must be homogeneous (for example the variability of the residuals should be the same at high values of the response variable as at low values)
- how to assess whether the variance is homogeneous
- that the residuals should come from identical and independent Normal distributions
- how to assess the Normality of the residuals
- why the model must be additive (that is, differences between treatment effects must remain the same however large or small the underlying size of the variable measured)
- how to identify outliers
- how transforming the response variate may correct for failures in the assumptions ★
- how to print back-transformed tables of means ★
- how to do a permutation or exact test ★

Note: the topics marked ★ are optional.

4.1 Homogeneity of variance

It is assumed that the variance is homogeneous, that is, the size of the random variation is similar over all the units. Homogeneity of variance can easily be assessed by plotting the residuals (estimates of the random error) against the fitted values: if the variance is homogeneous, the residuals should lie within a uniform band as in Figure 4.1 below.

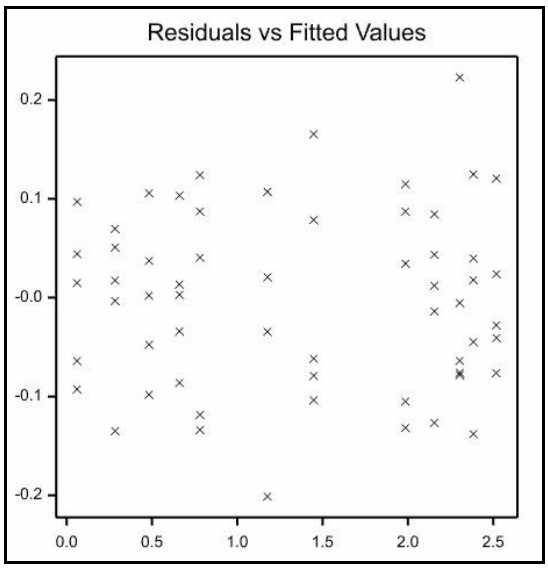


Figure 4.1

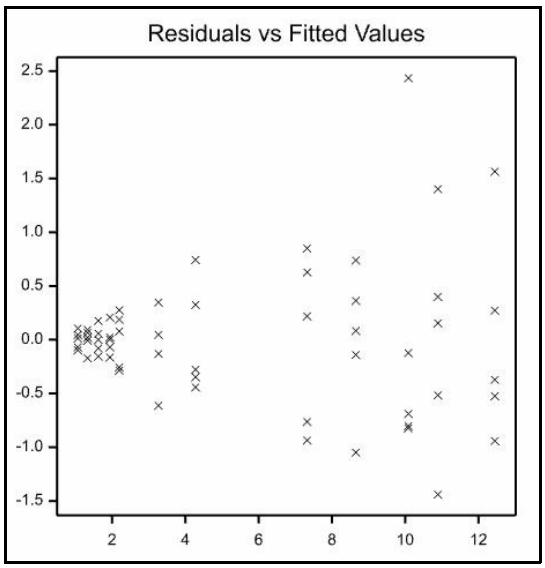


Figure 4.2

It is quite common, especially with count data, to find that the variation of the residuals increases as the value of the response increases, as in Figure 4.2. In this case, the standard errors of differences between treatments will be over-estimated for differences between treatments with low means, and under-estimated for differences between larger means, causing incorrect conclusions to be drawn. If a plot of residuals against fitted values indicates non-homogeneity of variances, a transformation of the data should be considered, as we show in Section 4.5.

One situation where unequal variances can occur, but where a transformation may not help, is when analyses are performed on data collected in different years or at different locations. It is then important to check that the variances within the years (or at each location) are homogeneous. Otherwise a weighted analysis will be required, with the data from each year being weighted by the reciprocal of the variance at that year. (This can be done automatically by using the Multiple Experiments / Meta Analysis (REML) menu, although we do not cover that here.)

4.2 Normality and independence of the residuals

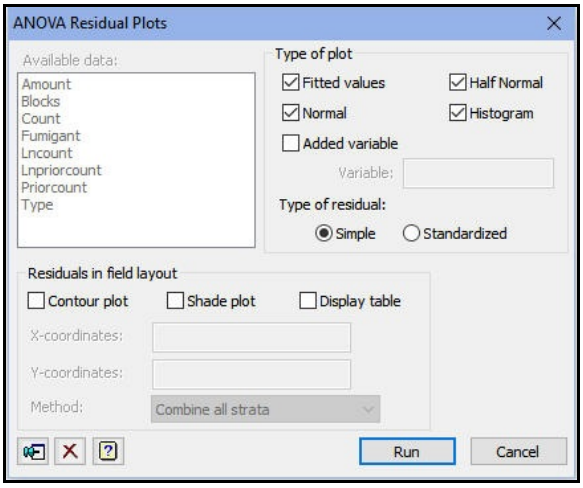


Figure 4.3

Analysis of variance assumes that the data contains random error (estimated by the residuals) that is independent and Normally distributed for each data value. Non-Normality of the residuals is usually also associated with non-homogeneity of variances and can be examined graphically in several ways. First the residuals can be plotted as a histogram – this should look approximately like a normal distribution, a non-skew bell-shaped distribution. Alternatively a *Normal* plot (or *half-Normal* plot) can be used. This plots the ordered residuals (or their absolute values) against the quantiles of a Normal distribution. If the residuals have a Normal distribution, these graphs should be straight lines.

These graphs, together with the plot of residuals against fitted values, can be produced by the ANOVA Residual Plots menu . This is obtained by clicking the Further output button on the Analysis of Variance menu, and then the Residual plots button on the ANOVA Further Output menu. The menu allows you to select the plots that you would like to see. The plots in Figures 4.4 and 4.5 were produced by the default settings, shown in Figure 4.3. Added variable plots can be used to plot the residuals against a potential covariate, to assess whether its relationship with the response variate is linear, and whether it may be worth including in an analysis of covariance (Section 3.6).

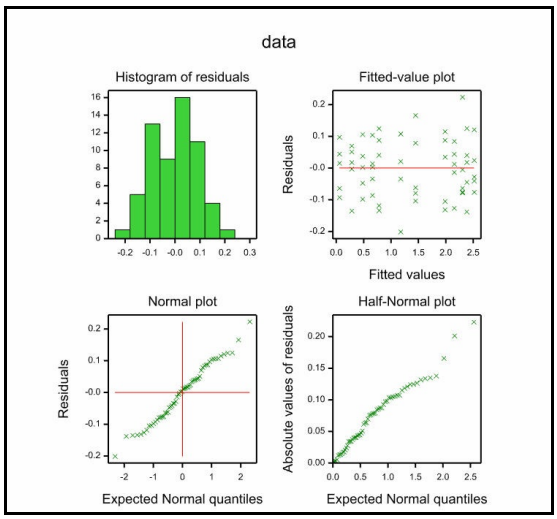


Figure 4.4

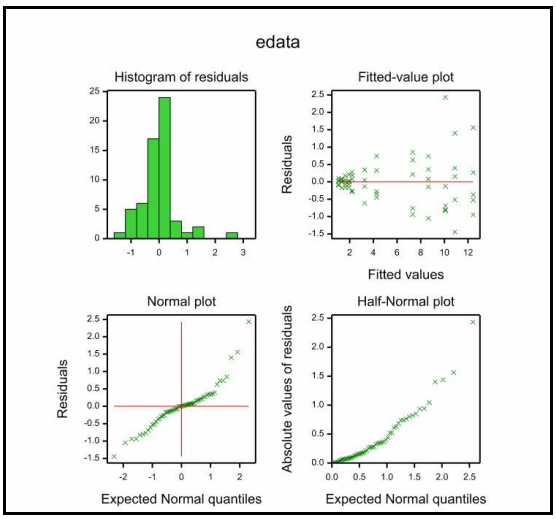


Figure 4.5

The plots in Figures 4.4 and 4.5 are from analyses of artificial data. The data on the left (Figure 4.4) was generated from a Normal distribution, the data on the right (Figure 4.5) is from a non-Normal distribution where the variance increases with the size of the response variable. Note that the histogram of residuals in Figure 4.5 is slightly skew, but there is a relatively small difference between the Normal and half-Normal plots. The difference between the two data sets is clearest in the plot of residuals against fitted values.

4.3 Additivity of the model

If you fit an *additive* model to your data, you are assuming that differences between treatment effects remain the same however large or small the underlying size of the variable measured. For example, in a randomized-block design, the assumption is that the theoretical value of the difference between two treatments remains the same within a block where the recorded values are generally low, as in one where the values are generally high. An example of non-additivity occurs where treatments give a proportionate increase or decrease to data values. In an additive model, the effect of a treatment is a constant increase or decrease.

If you fit an additive model where non-additivity is present this will often lead to the detection of interactions in the analysis. Of course, genuine interactions between treatment terms may also occur, for example associated with one treatment modifying the mode of action of another. However, the additive model assumes that interactions between blocks and treatments do not occur and so examining these interactions is a good way to look for evidence of non-additivity. You will usually find that data which shows signs of non-additivity also violates other assumptions.

4.4 Outliers

An *outlier* is an extreme observation, which leads to a unit with a very large residual. Genstat `ANOVA` will produce warnings if any units have large residuals compared to the standard error of the units. You can also use the diagnostic plots produced by the ANOVA Residual Plots menu to detect outliers in your data. Outliers will appear as extreme observations in the graph of residuals against fitted values, or in a histogram of residuals. They will also appear as single values away from the line in a normal or half-normal plot.

Outliers may arise from an error in recording or punching data, if the wrong treatment has been applied to a unit, or where something else has gone wrong in the experimental procedure. When outliers are present, they can distort treatment means as well as inflating the error variance so that the precision of estimates is decreased. If any observation appears to be an outlier, you should investigate the observation to try and find out if an error has occurred. If you can uncover an error and use the correct data value, then you should do so. If you find an error but cannot recover the correct data value, then you should replace the incorrect value by a missing value. If you cannot track down any possible source of error, you should consider whether the outlier might be a true data value, and whether your model for the data is wrong!

4.5 Transformations

Failures of the assumptions can often be corrected by transforming the data, using the Calculate menu. Different transformations are appropriate for different types of data. The most common types of data requiring transformations are counts, percentages and proportions. Some transformations are used only to stabilize the variance (i.e. to make it homogeneous), but it is equally important to consider the additivity of the model. In some situations a transformation can be chosen both to provide additivity and to stabilise the variance. If this proves to be impossible, you should consider using a generalized linear model; see the *Guide to the Genstat Command Language*, Part 2, Section 3.5.

Count data occur where an experiment counts the occurrences of some event with no preset upper limit, for example, the number of accidents occurring on a section of road, numbers of hits on a web site, numbers of weed plants in a plot, and so on. Conventional wisdom is to stabilize the variance, using a square-root transformation. However, this will usually not provide an additive model – the treatments generally take the effect of a proportionate increase (or decrease). A logarithmic transformation would then give an additive scale for the treatments, and will often be found also to give adequate stability for the variance. To guard against zero counts it is usual to add a small constant to the response `y` before taking the logarithms: for example to use `LOG10(y+1)` or `LOG(y+0.5)`.

Proportion or percentage data can arise in several ways. Sometimes, the data value is a natural continuous percentage measure, for example, the percentage area of a plot that has been infected by a disease. Treatment effects are often then found to be approximately proportional to the amount infected for low percentages, while for percentages near to 100% they tend to be proportional to the amount uninfected. If the percentages are obtained by visual assessment of areas such as infected parts of leaves, the same pattern is found: for low percentages the eye tends to examine the amount infected, while nearer to 100% it is the amount uninfected that is assessed. In this situation, a logit transformation, $\log(p/(100-p))$, would both stabilize the variance and give an additive model.

Alternatively, the data may count the number of occurrences (r) of some event in a population of fixed size n (binomial data), for example, the number of children to have been vaccinated out of a class of 30, or the number of infected plants out of a sample of 40. Binomial data can be converted to percentages ($p=100\times r/n$) for analysis. Conventional wisdom is to stabilize the variance of binomial data by taking an angular transformation, $\arcsin(\sqrt{(p/100)})$. However, this will generally not give an additive model, so it may be worth considering a logit transformation instead. To guard against 0 or 100% values, you can then calculate the percentage as $p=100\times(r+0.5)/(n+1)$.

Finally, where data values span a very large range, for example, where the range of the data is more than two or three times the mean value, the treatment effects and the variance are often both found to be proportional to the size of response. It would then be appropriate to take a logarithmic transformation.

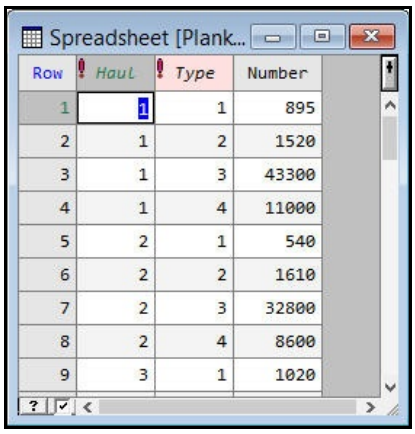


Figure 4.6

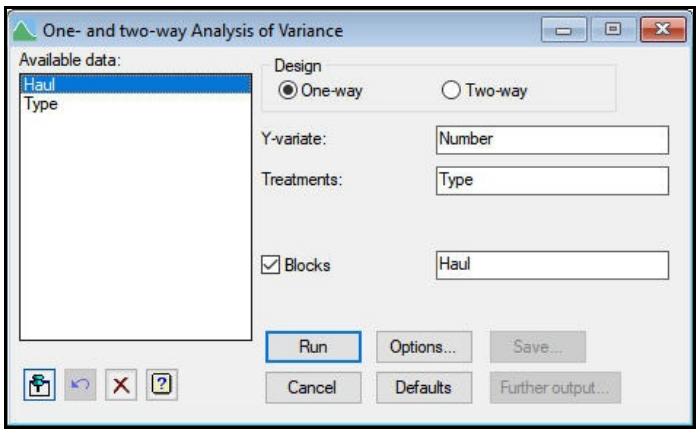


Figure 4.7

Spreadsheet `Plankton.gsh` contains data from a study of plankton numbers (Snedecor & Cochran 1967, *Statistical Methods, 6th Edition*, page 329). Four types of plankton were sampled in 12 hauls. In the analysis, hauls are treated as blocks, and types of plankton as treatments (Figure 4.7). The first analysis is of the untransformed counts.

Analysis of variance

Variate: Number

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Haul stratum	11	2.153E+08	1.957E+07	1.91	
Haul.*Units* stratum					
Type	3	7.035E+09	2.345E+09	228.71	<.001
Residual	33	3.384E+08	1.025E+07		
Total	47	7.589E+09			

Tables of means

Variate: Number

Grand mean 10636.

Type	1	2	3	4
	671.	1701.	30775.	9396.

Standard errors of differences of means

Table	Type
rep.	12
d.f.	33
s.e.d.	1307.2

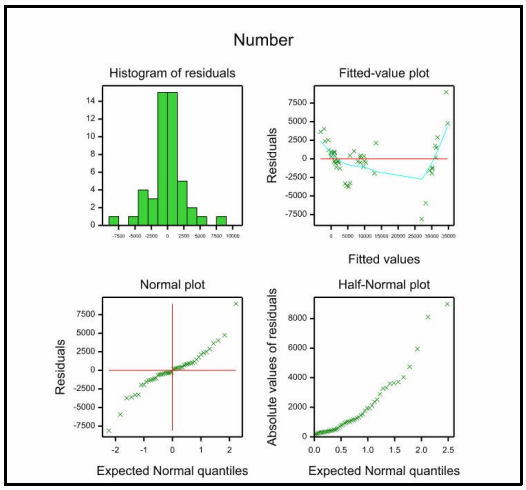


Figure 4.8

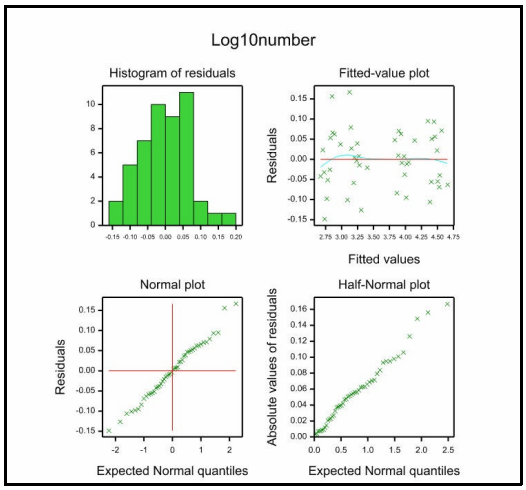


Figure 4.9

Figure 4.8 shows the residual plot from the untransformed analysis, and Figure 4.9 shows the residual plot from the analysis of the log-transformed numbers. The output from the transformed is shown below. The untransformed fitted-value plot shows clear evidence that the variance is increasing with the size of the number – which is corrected in the transformed analysis.

Analysis of variance

Variate: Log10number

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Haul stratum	11	0.337442	0.030677	4.41	
Haul.*Units* stratum					
Type	3	20.169765	6.723255	965.74	<.001
Residual	33	0.229737	0.006962		
Total	47	20.736944			

Tables of means

Variate: Log10number

Grand mean 3.616

Type	1	2	3	4
	2.803	3.221	4.478	3.962

Standard errors of differences of means

Table	Type
rep.	12

d.f.	33
s.e.d.	0.0341

If you are analysing transformed data, it is important to remember that the statistical properties of the analysis apply only on the transformed scale. So, for example, comparisons between means must be assessed on the transformed scale (i.e. using the tables of means and s.e.d.'s, or l.s.d.'s, from the analysis of the transformed data). For interpretation, though, it is often helpful also to present the tables of means back-transformed to the original scale. These values are often given in brackets under the transformed values. To save the means, you click on the Save button on the Analysis of Variance menu, to open the ANOVA Save Options menu. Check the Means box, and then fill in an identifier for the table (here `Meanlogplankton`) to store the means.

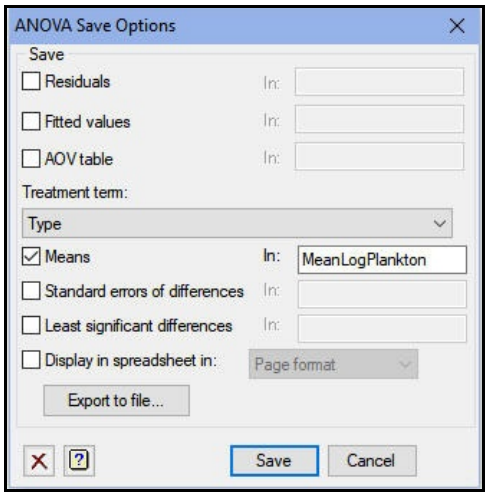


Figure 4.10

You can calculate the back-transform the means by using the Calculate menu (accessible from the Data menu on the menu bar); see Figure 4.11.

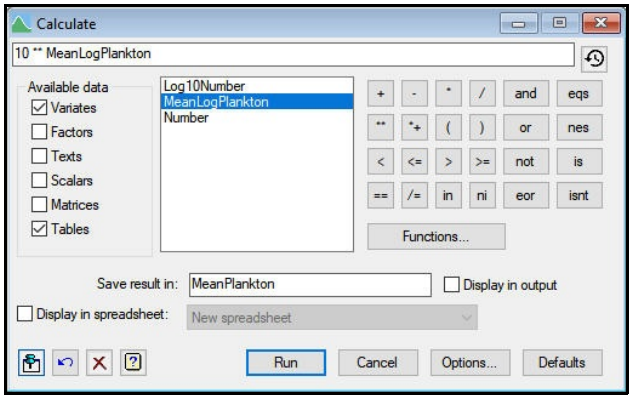


Figure 4.11

To display the tables click on the Display Data in Output option of the Data menu on the menu bar. In the resulting Display Data in Output menu (Figure 4.12), use the arrow to put the two tables into the right-hand box. Highlight each

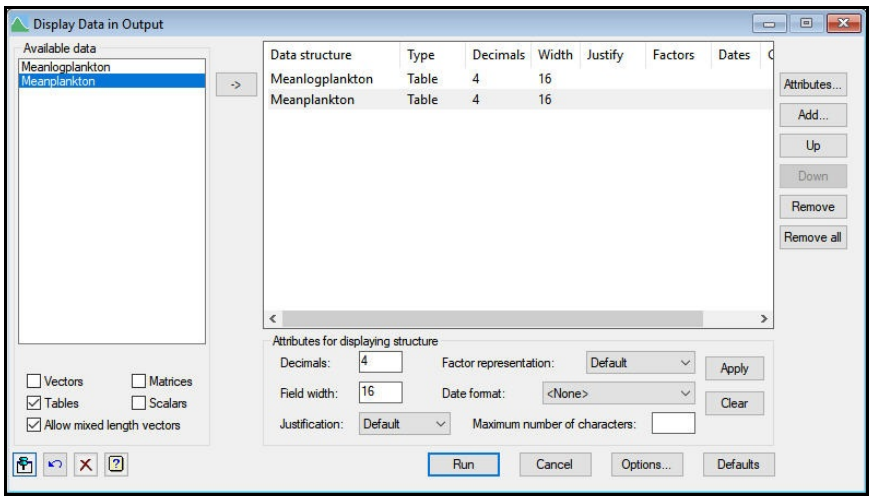


Figure 4.12

table in that box, enter the number of Decimals and Field width and click on Apply. Then click on Run to produce the output below.

Type	Meanlogplankton	Meanplankton
1	2.8026	634.80
2	3.2213	1664.39
3	4.4783	30084.69
4	3.9621	9164.54

4.6 Automatic testing of the assumptions

In addition to the visual checks of the assumptions, described earlier in this chapter, you can also make automatic checks when using the general Analysis of Variance menu. We can illustrate these using the plankton data, analysed above..

First we set up the menu to specify the analysis, as shown in Figure 4.13.

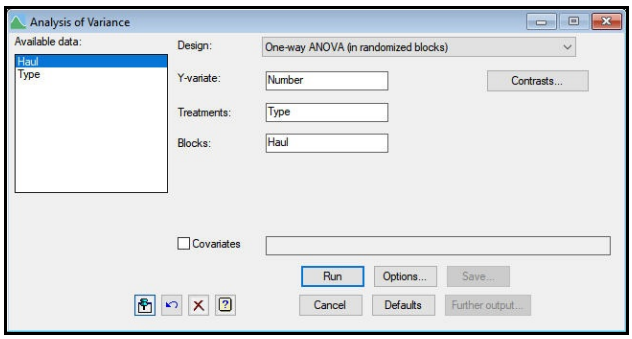


Figure 4.13

Then we open the ANOVA Options menu, and check the Assumptions box, as shown on Figure 4.14. To avoid duplication, we will not print any other output this time.

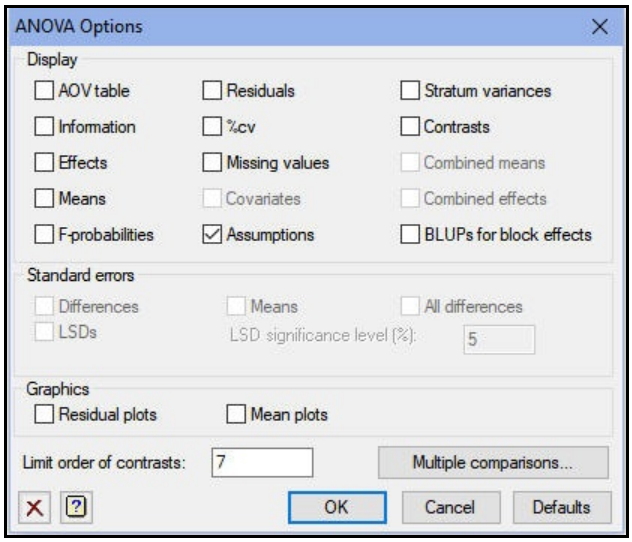


Figure 4.14

Genstat now performs three types of check. Firstly, it performs Levene tests to check whether the residual variance seems to be affected by any of the terms in the analysis (here `Type` and `Haul`). Then it performs a Shapiro-Wilk test to check for evidence that the residuals do not come from a Normal distribution. Finally, it performs two Levine tests to check whether the residual variance differs according to the size of the response. The data are divided into three groups (small, intermediate and large) according to the sizes of their fitted values. The tests compare the variance of the residuals in the first (small) group with those in the third (large) group, and the variance of the second (intermediate) group with the variance of other two groups combined. Warning messages are given if any of the tests generates a test probability less than or equal to 0.025. This is the same as the value used for the similar messages that may occur with the summary of analysis in regression. It is important to realise that the estimated residuals (from either regression or analysis of variance) will be correlated. The Levene

and Shapiro-Wilk tests assume that the residuals are independent Normally-distributed observations. Their test probabilities may therefore be too low – and generate too many significant results. So the use of a smaller critical probability value provides some protection against spurious messages.

As expected, Genstat reports evidence of both non-homogeneity of the residual variance, and of non-Normality.

Message: evidence of non-homogeneity of residual variance for Type and Haul.
Message: the Shapiro-Wilk test shows evidence of non-Normality.

The ANOVA Options menu does not print the tests themselves, but these are given if you use the Assumption tests box ANOVA Further Output menu (Figure 4.15). The setting in the options menu is intended to allow unobtrusive background testing, while that in the further output menu gives further output – as requested.

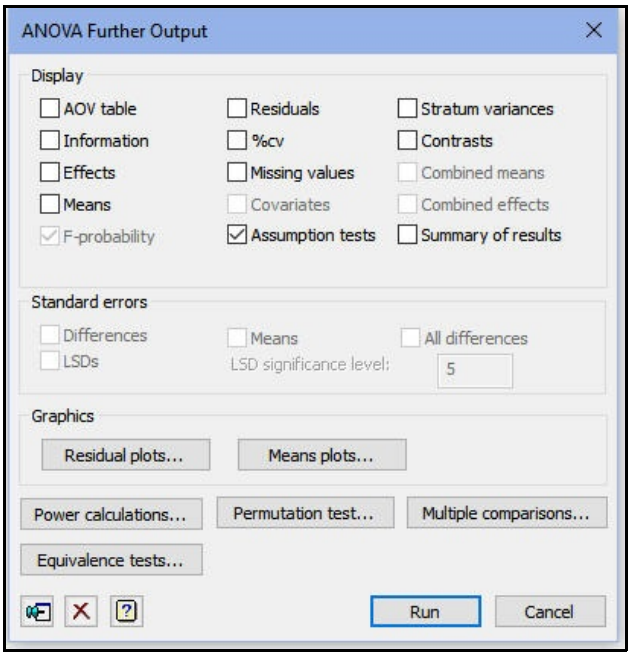


Figure 4.15

Tests of assumptions for ANOVA

Variate: Number

Levene tests for homogeneity of variance

Analysis of variance

Variate: Absolute residuals

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Haul stratum	11	13.8440	1.2585	7.17	

Haul.*Units* stratum					
Type	3	6.0333	2.0111	11.45	<.001
Residual	33	5.7949	0.1756		
Total	47	25.6721			

Tables of means

Variate: Absolute residuals

Grand mean 0.682

Type	1	2	3	4
	0.584	0.594	1.259	0.291

Standard errors of differences of means

Table	Type
rep.	12
d.f.	33
s.e.d.	0.1711

Levene tests for stability of variance

Test	t-statistic	d.f.	pr.
Small vs. large responses	2.285	12.703	0.040
Intermediate v.s. small & large responses	1.906	16.762	0.074

Shapiro-Wilk test for Normality

Data variate:	Residuals
Test statistic W:	0.9351
Probability:	0.011

Message: evidence of non-homogeneity of residual variance for Type and Haul.
Message: the Shapiro-Wilk test shows evidence of non-Normality.

The output shows that the type-3 plankton numbers are more variable than the other types. (This is not surprising as many more of this type of plankton have been recorded in the experiment than the other types.)

If we repeat the analysis with the log-transformed numbers, there is no evidence that the assumptions are broken, and no warnings are given.

4.7 Practical

An experiment was conducted to assess the percentage of alcohol by volume of five types of wine labelled A to E. Three bottles of each type were tested in the laboratory in a random order, as listed below and stored in file `Wine.gsh`.

```
E  4.931
D  7.263
A  4.857
C  3.361
B  6.871
E  4.141
C  3.164
B  3.012
A  5.668
D 12.185
B  4.223
E  3.323
A  4.668
C  2.686
D  7.776
```

Analyse the experiment and plot a graph of the residuals against the fitted values.

Transform the data using a logit transformation, re-analyse the data and plot another graph of residuals against fitted values.

4.8 Permutation and exact tests

If the distributional assumptions for the analysis of variance are not satisfied, you might use a permutation test an alternative way to assess the significance of the terms in the analysis. You still need the model to be additive for the results to be meaningful, but there is no longer any need for the residuals to follow Normal distributions with equal variances.

Clicking on the Permutation test button in the ANOVA Further Output menu (Figure 1.10) produces the menu in Figure 4.16. This asks Genstat to make 4999 random permutations of the values of the response variate (see the Number of permutations box), and repeat the analysis with each one. The Seed box specifies the seed to use for the random-number generator that is used to construct the permutations. The value 0 initializes the seed automatically (and prints the value in the output) if this is the first use of the generator in this run of Genstat; otherwise the seed is chosen to continue the existing sequence.

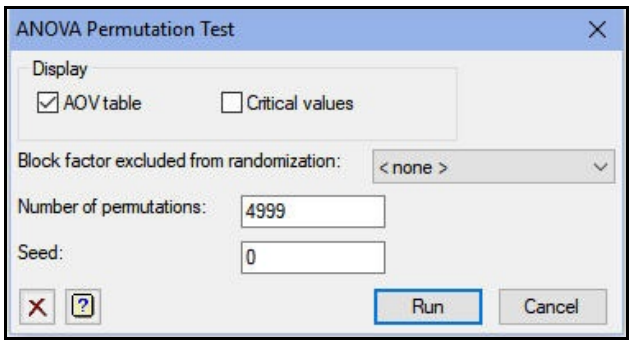


Figure 4.16

The probability for each treatment term is now determined from its distribution over the randomly permuted data sets. The output below prints a probability value <.001 for `Type`, which means that the observed data set was one of the 5 sets with the largest variance ratios out of the 5000 sets that have been examined (1 observed data set + 4999 randomly permuted data sets).

Message: Default seed for random number generator used with value 582564

Analysis of variance

Variate: Log10number
Probabilities determined from 4999 random permutations

Source of variation	d.f.	s.s.	m.s.	v.r.	prob.
Haul stratum	11	0.33744	0.03068	4.41	
Haul.*Units* stratum					
Type	3	20.16976	6.72325	965.74	<.001
Residual	33	0.22974	0.00696		

If you ask for more permutations than the number that are possible for your data, Genstat will instead do an *exact test*, which uses each permutation once.

4.9 Practical

Extend the analysis of the logit-transformed percentage of alcohol from Practical 4.6 by performing a permutation test, and checking whether the assumptions are still broken..

5 Designs with several error terms

The randomized-block design is undoubtedly the most popular of the designs in common use, but sometimes more sophisticated arrangements may be required involving units of different sizes. For example, there are sometimes treatments, like plant varieties or irrigation, that cannot conveniently be applied to the small plots that are feasible for treatments like levels of fertiliser or types of fungicide. In this chapter you will learn

- how a split-plot design is constructed
- how to analyse a split-plot design, and interpret the output
- why the analysis of variance table for a split-plot design has more than one *stratum* (or error term)
- how to define the block structure for other *stratified* designs ★
- what happens when the response variate contains missing values ★

Note: the topics marked ★ are optional.

5.1 Split-plot design

V3 N3	V3 N2	V3 N2	V3 N3
V3 N1	V3 N0	V3 N0	V3 N1
V1 N0	V1 N1	V2 N0	V2 N2
V1 N3	V1 N2	V2 N3	V2 N1
V2 N0	V2 N1	V1 N1	V1 N2
V2 N2	V2 N3	V1 N3	V1 N0
V3 N2	V3 N0	V2 N3	V2 N0
V3 N1	V3 N3	V2 N2	V2 N1
V1 N3	V1 N0	V1 N2	V1 N3
V1 N1	V1 N2	V1 N0	V1 N1
V2 N1	V2 N0	V3 N2	V3 N3
V2 N2	V2 N3	V3 N1	V3 N0
V2 N1	V2 N2	V1 N2	V1 N0
V2 N3	V2 N0	V1 N3	V1 N1
V3 N3	V3 N1	V2 N3	V2 N2
V3 N2	V3 N0	V2 N0	V2 N1
V1 N0	V1 N3	V3 N0	V3 N1
V1 N1	V1 N2	V3 N2	V3 N3

In the *split-plot* design shown here, the treatments are three varieties of oats (*Victory*, *Golden rain* and *Marvellous*) and four levels of nitrogen (0, 0.2, 0.4 and 0.6 cwt). As it is feasible to work with smaller plots for fertiliser than for varieties, the six blocks were initially split into three whole-plots and then each whole-plot was split into four subplots. The varieties were allocated (at random) to the whole-plots within each block, and the nitrogen levels (at random) to the subplots within each whole-plot. In a randomized-block design, we have a hierarchical structure with blocks and then plots within blocks.

Results from the experiment are in spreadsheet file `Oats.gsh` in the `Data` folder.

The split-plot is another design with a customized setting in the general Analysis of Variance menu, as shown in Figure 5.1. The treatment structure is a factorial with two factors, and is specified by a model formula as described in Chapter 3. The block structure is set up automatically by Genstat from the factors specified in the Blocks, Whole plots and Sub-plots fields.

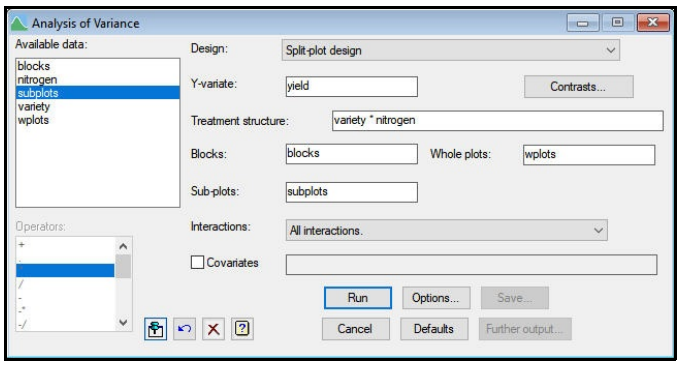


Figure 5.1

The analysis-of-variance table shows that we now have three *strata* in the hierarchy: blocks, whole-plots within blocks, and subplots within whole plots (within blocks). Moreover, the analysis has more than one residual: in the split-plot design we need to consider the random variability of the whole-plots as well as the variability of the

subplots. The sum of squares for `Variety` (which was applied to complete whole-plots) can correctly be compared with a residual which represents the random variability of the whole-plots. Conversely, `Nitrogen` (which was applied to subplots) and the `Variety.Nitrogen` interaction are compared with the residual for subplots within whole-plots.

Analysis of variance

Variate: yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
blocks stratum	5	15875.3	3175.1	5.28	
blocks.wplots stratum					
variety	2	1786.4	893.2	1.49	0.272
Residual	10	6013.3	601.3	3.40	
blocks.wplots.subplots stratum					
nitrogen	3	20020.5	6673.5	37.69	<.001
nitrogen.variety	6	321.8	53.6	0.30	0.932
Residual	45	7968.8	177.1		
Total	71	51985.9			

Tables of means

Variate: yield

Grand mean 104.0

nitrogen	0 cwt	0.2 cwt	0.4 cwt	0.6 cwt
	79.4	98.9	114.2	123.4

variety	Victory	Golden rain	Marvellous
	97.6	104.5	109.8

nitrogen	variety	Victory	Golden rain	Marvellous
0 cwt		71.5	80.0	86.7
0.2 cwt		89.7	98.5	108.5
0.4 cwt		110.8	114.7	117.2
0.6 cwt		118.5	124.8	126.8

Standard errors of differences of means

Table	nitrogen	variety	nitrogen variety
rep.	18	24	6
s.e.d.	4.44	7.08	9.72
d.f.	45	10	30.23

Except when comparing means with the same level(s) of variety	7.68
d.f.	45

The standard errors accompanying the tables of means also take account of the stratum where each treatment term was estimated. The `Variety` s.e.d. of

$$7.08 = \sqrt{(2 \times 601.3 / 24)}$$

is based on the residual mean square for `Blocks.Wplots`, while that for `Nitrogen`

$$4.44 = \sqrt{(2 \times 177.1 / 18)}$$

is based on that for `Blocks.Wplots.Subplots`. The `Variety` × `Nitrogen` table is more interesting. There are two s.e.d.'s according to whether the two means to be compared are for the same variety. If they are, then the subplots from which the means are calculated will all involve the same set of whole-plots, so any whole-plot variability will cancel out, giving a smaller s.e.d. than for a pair of means involving different varieties.

Split-plot designs do not only occur in field experiments, but they can occur in animal trials (where, for example, the same diet may need to be fed to all the animals in a pen but other treatments may be applied to individual animals), or in industrial experiments (where different processes may require different sized batches of material), or even in cookery experiments (see, for example, Cochran & Cox 1957, page 299). There can also be more than one treatment factor applied to the units of any stratum; to analyse the results in Genstat, you simply need to specify the blocking factors, as above, and then whatever treatment structure is appropriate.

Genstat specifies the structure of the design, and thus the different sources of variability (or strata) in the model, using the `BLOCKSTRUCTURE` directive (see Chapter 9). For Figure 5.1, this was

```
BLOCKSTRUCTURE  Blocks / Wplots / Subplots
```

where the operator / indicates that a factor is nested within another factor. So we have `Subplots` nested within `Wplots` (whole-plots) nested within `Blocks`, as required. The model formula expands to the list of model terms

```
Blocks  +  Blocks.Wplots  +  Blocks.Wplots.Subplots
```

which defines the strata to represent the variation between the blocks, between whole-plots within blocks, and between subplots within whole plots (within blocks) shown in the analysis-of-variance table.

The next section shows how you can define your own block structure in the menu, and specify any stratified design.

5.2 Practical

In an experiment to study the effect of two meat-tenderizing chemicals, the two (back) legs were taken from four carcasses of beef and one leg was treated with chemical 1 and the other with chemical 2. Three sections were then cut from each leg and allocated (at random) to three cooking temperatures, all 24 sections (4 carcasses × 2 legs × 3 sections) being cooked in separate ovens. The table below shows the force required to break a strip of meat taken from each of the cooked sections (the data are also in the file `Meat.gsh`). Analyse the experiment.

Leg		1			2		
Carcass	Section	Chemical	Temp	Force	Chemical	Temp	Force
1	1	1	2	5.5	2	3	6.3
	2	1	3	6.5	2	1	3.5
	3	1	1	4.3	2	2	4.8
2	1	2	1	3.2	1	3	6.2
	2	2	3	6.0	1	2	5.0
	3	2	2	4.7	1	1	4.0
3	1	2	1	2.6	1	2	4.6
	2	2	2	4.3	1	1	3.8
	3	2	3	5.6	1	3	5.8
4	1	1	3	5.7	2	2	4.1
	2	1	1	3.7	2	3	5.9
	3	1	2	4.9	2	1	2.9

On the assumption that the temperature levels are equally spaced and increasing, use the polynomial contrast menu to see whether the force increases linearly with temperature.

5.3 Other stratified designs

The ideas behind the split-plot design can easily be extended to allow for further subdivisions. For example, in a split-split-plot design if we would split the subplots into sub-subplots with a further factor, `Subsubplot`, to obtain a block structure of

```
Blocks / Wplots / Subplots / Subsubplot
```

leading to a further term (and thus stratum)

```
Blocks.Wplots.Subplots.Subsubplot
```

Designs like this can be specified using the General analysis of variance design setting of the Analysis of Variance menu. Provided the necessary factors are correctly defined, Genstat will determine automatically the stratum where each treatment term is estimated, and calculate appropriate s.e.d's for each table of means.

D3 N1	D2 N2	D1 N2	D4 N2
D3 N2	D2 N1	D1 N1	D4 N1
D1 N1	D4 N1	D3 N1	D2 N2
D1 N2	D4 N2	D3 N2	D2 N1
D4 N1	D1 N1	D2 N2	D3 N1
D4 N2	D1 N2	D2 N1	D3 N2
D2 N2	D3 N1	D4 N2	D1 N1
D2 N1	D3 N2	D4 N1	D1 N2

You can also have designs involving both crossing and nesting. The plan above shows an experiment set up to study the effects of cutting date and a nitrogen treatment on the

yield of a forage crop. The main-plot treatment is `Cutdate` (D1-4 on the plan), and the individual plots of the square have been split into pairs to allow for the two `Nitrogen` treatments (0 and 0.3). The subplot factor is nested below the usual block formula for a Latin square

```
(Rows * Columns) / Subplots
=  Rows  + Columns  + Rows.Columns  + Rows.Columns.Subplots
```

to give an extra stratum `Rows.Columns.Subplots` to represent the variation of the subplots within the plots of the Latin square.

The data are in spreadsheet file `Forage.gsh`, and the variate to be analysed is the yield of forage.

Again, the two-way table of means has two s.e.d's depending on the level of the factor that was applied to the plots of the design.

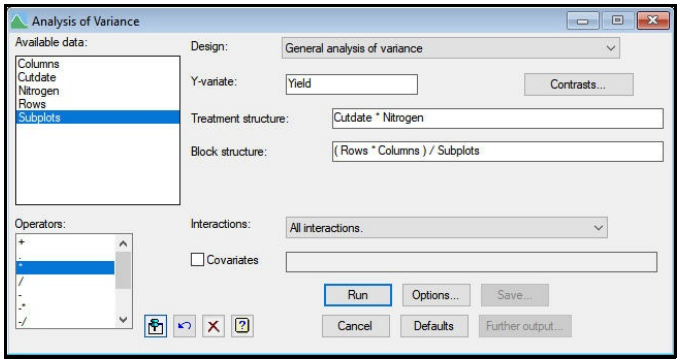


Figure 5.2

Analysis of variance

Variate: Yield

Source of variation	d.f.	(m.v.)	s.s.	m.s.	v.r.	F pr.
Rows stratum	3		87.603	29.201	0.81	
Columns stratum	3		110.181	36.727	1.02	
Rows.Columns stratum						
Cutdate	3		23019.485	7673.162	212.53	<.001
Residual	5	(1)	180.515	36.103	17.11	
Rows.Columns.Subplots stratum						
Nitrogen	1		232.890	232.890	110.37	<.001
Cutdate.Nitrogen	3		27.004	9.001	4.27	0.035
Residual	10	(2)	21.102	2.110		
Total	28	(3)	21265.627			

Tables of means

Variate: Yield

Grand mean 62.64

Cutdate	Jun11	Jul01	Jul22	Aug12
	20.60	58.95	80.48	90.53

Nitrogen	0.0	0.3
	59.94	65.34

Cutdate	Nitrogen	0.0	0.3
Jun11		18.73	22.48
Jul01		56.40	61.50
Jul22		76.25	84.72
Aug12		88.40	92.67

Standard errors of differences of means

Table	Cutdate	Nitrogen	Cutdate Nitrogen
rep.	8	16	4
s.e.d.	3.004	0.514	3.091
d.f.	5	10	5.59
Except when comparing means with the same level(s) of			
Cutdate			1.027
d.f.			10

(Not adjusted for missing values)

This example also shows how the analysis can cope with missing values as may occur if a unit is damaged or, for some reason, fails to be measured. Here we have lost one complete plot and half another one. The residual degrees of freedom are adjusted (as shown in brackets) and the missing values are estimated as part of the analysis. The analysis involves approximations but, provided only a few units are missing, these should be acceptable. (See the *Guide to the Genstat, Part 2: Statistics,* Section 4.4 for more details.)

5.4 Practical

Spreadsheet file `Rice.gsh` contains data from an experiment that studied the effect of three levels of nitrogen fertilizer on the yields of six varieties of rice (Gomez & Gomez, 1984, *Statistical Procedures for Agricultural Research*, page 110).

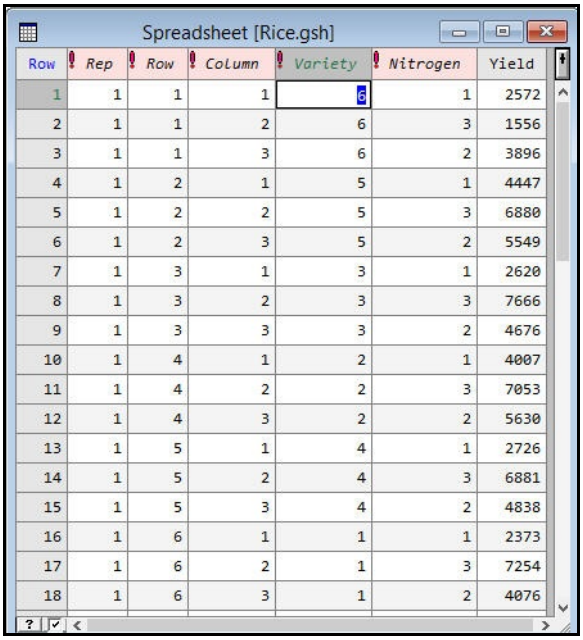
The experiment used a strip-plot design. This is a replicated row and column design. Each replicate had three columns and six rows. Within each replicate, the nitrogen levels were randomized onto the columns, and the varieties were randomized onto the rows. So the block structure is

```
Rep / (Row * Column)
```

and the treatment structure is

```
Variety * Nitrogen
```

Analyse the yields.



Row	Rep	Row	Column	Variety	Nitrogen	Yield
1	1	1	1	6	1	2572
2	1	1	2	6	3	1556
3	1	1	3	6	2	3896
4	1	2	1	5	1	4447
5	1	2	2	5	3	6880
6	1	2	3	5	2	5549
7	1	3	1	3	1	2620
8	1	3	2	3	3	7666
9	1	3	3	3	2	4676
10	1	4	1	2	1	4007
11	1	4	2	2	3	7053
12	1	4	3	2	2	5630
13	1	5	1	4	1	2726
14	1	5	2	4	3	6881
15	1	5	3	4	2	4838
16	1	6	1	1	1	2373
17	1	6	2	1	3	7254
18	1	6	3	1	2	4076

Figure 5.3

6 Design and sample size

In this chapter you will learn

- how to use the Generate a Standard Design menu
- how to decide how many replicates you need, using the Replications Required menu
- how to assess the *power* of the design i.e. the probability that it will be able to detect the treatment effects that you expect
- how to include additional control treatments ★

Note: the topics marked ★ are optional.

6.1 Designing an experiment

The Generate a Standard Design menu enables you to generate many standard experimental designs. It is obtained by clicking Stats on the menu bar and selecting Design, followed by Standard Design. The type of design is selected using the Design list box. The categories parallel those in the Analysis of Variance menu – again each with its appropriate boxes and buttons.

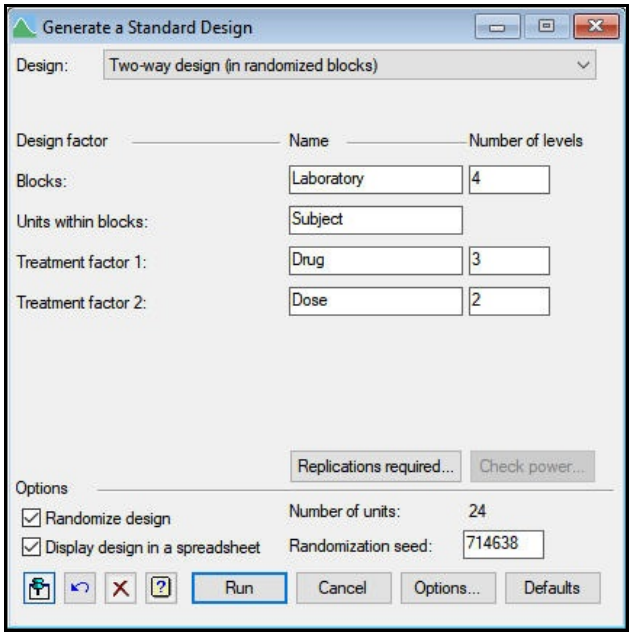


Figure 6.1

The menu in Figure 6.1 generates a randomized-block design with four blocks (corresponding to four different laboratories) to study two treatment factors: `Drug` with three levels, and `Dose` with two levels. Checking the Randomize design box asks Genstat to randomize the design. Genstat automatically determines the appropriate type of randomization from the inter-relationships of the blocking factors of the design. For a randomized-block design, this amounts to randomizing the allocation of the treatments independently within each block; see Section 6.3. (However, if you want to do your own randomization, you can use the Randomize menu, obtained by clicking Stats on the menu bar and selecting Design, followed by Randomize.) The Randomization seed box supplies a seed used to generate the random numbers for the randomization. Genstat suggests a seed automatically (at random), in the same way that it suggests defaults for the other fields in the menu. However, you can supply your own seed if you prefer, and keeping the same seed will generate the same randomization if you want to reproduce the exact design in future.

The Generate a Standard Design Options menu (Figure 6.2) provides further contols. In Figure 6.2, the Generate plot / unit labels box is checked to form labels to identify the units of the design. It is often more convenient to use a single numerical code to identify observations from an experiment, rather than having to use the levels of all the blocking factors (here subjects within laboratories). The labels will be integer numbers 1, 2 and so on. These will be saved in the variate `Subjcode`, specified in the Column name for labels window. The Design box is checked to print the design, and the Dummy ANOVA table box is checked to generate a *skeleton* analysis-of-variance. We now click on OK to return to the main menu.

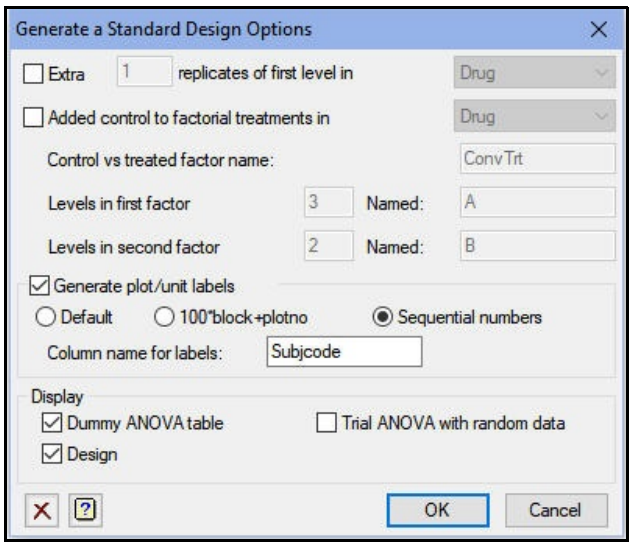


Figure 6.2

Back in the Generate a Standard Design menu (Figure 6.1), clicking on the Replications required button produces a menu that allows you to determine the replication (Figure 6.3). For a randomized-block design, the replication depends on the number of blocks (here laboratories). To make the calculation, Genstat needs to know which treatment term you are concerned about (here `Drug.Dose`) and the size of the smallest difference that you need to detect (here 1.5). You also need to indicate how large you expect the within-block variance to be (here we are assuming 0.5). The variance is best obtained from an earlier analysis of similar data, and is provided by the residual mean square in the "block.plot" (in this case, `Laboratory.Subject`) stratum. Other boxes allow you to set the significance level that you plan to use to detect the difference (i.e. *alpha*) and the probability of detection (i.e. the *power* required for the test).

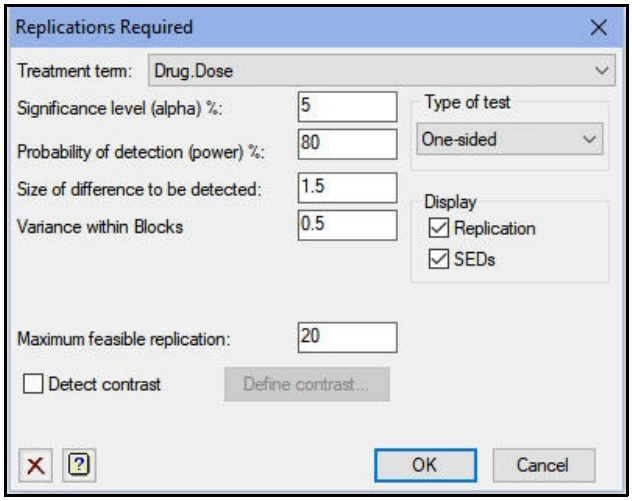


Figure 6.3

Clicking OK in Figure 6.3, pops ups the menu shown in Figure 6.4, which indicates the required number of replicates. You can then either click Apply to enter that number automatically into the design menu (Figure 6.1), click Cancel to close the menu with no actions, or click Change to return to the Replications Required menu (Figure 6.4). The result here, of 4, matches what we had hoped to find (and the value that we had already entered into the main menu!). So we can simple click Cancel.

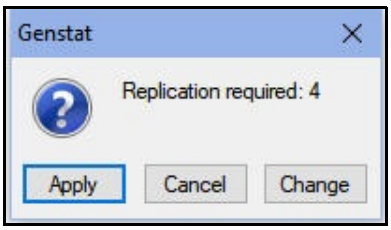


Figure 6.4

The Replication and SEDs boxes were checked in Figure 6.3, so Genstat prints a table giving the power (and the standard errors of differences) for up to 20 replicates, and a report of the required replication.

Power

Number of replicates	Residual d.f.	Residual m.s.	s.e.d.	RESPONSE / s.e.d.	t-value	Power
2	5	0.5000	0.7071	2.121	2.015	0.572
3	10	0.5000	0.5774	2.598	1.812	0.780
4	15	0.5000	0.5000	3.000	1.753	0.888
5	20	0.5000	0.4472	3.354	1.725	0.944
6	25	0.5000	0.4082	3.674	1.708	0.973
7	30	0.5000	0.3780	3.969	1.697	0.987
8	35	0.5000	0.3536	4.243	1.690	0.994
9	40	0.5000	0.3333	4.500	1.684	0.997
10	45	0.5000	0.3162	4.743	1.679	0.999
11	50	0.5000	0.3015	4.975	1.676	0.999
12	55	0.5000	0.2887	5.196	1.673	1.000
13	60	0.5000	0.2774	5.408	1.671	1.000
14	65	0.5000	0.2673	5.612	1.669	1.000
15	70	0.5000	0.2582	5.809	1.667	1.000
16	75	0.5000	0.2500	6.000	1.665	1.000
17	80	0.5000	0.2425	6.185	1.664	1.000
18	85	0.5000	0.2357	6.364	1.663	1.000
19	90	0.5000	0.2294	6.538	1.662	1.000
20	95	0.5000	0.2236	6.708	1.661	1.000

Replication

To detect a treatment difference of 1.500, at a significance level of 0.050,
with a power of 0.800, using a one-sided test, requires a replication of 4.

The Replications required button is available for any design where the replication can be modified simply by altering the number of levels of one of the factors (for example split-plot designs, split-split-plot designs, criss-cross designs and so on), but not e.g. for Latin squares where the replication cannot be changed without changing the number of levels of the treatment factor.

The Generate a Standard Design menu (Figure 6.1) will now be back as the active window. We have set our options and checked that the replication will be sufficient. So we now click on Run to generate the design, and the output below.

Treatment combinations on each unit of the design

Laboratory	1	2	3	4
Subject				
1	1 1	3 1	3 2	1 1
2	3 2	2 1	1 1	1 2
3	2 1	3 2	1 2	3 1
4	1 2	1 1	2 2	3 2
5	2 2	1 2	3 1	2 1
6	3 1	2 2	2 1	2 2

Treatment factors are listed in the order: Drug, Dose.

Analysis of variance

Source of variation	d.f.
Laboratory stratum	3
Laboratory.Subject stratum	
Drug	2
Dose	1
Drug.Dose	2
Residual	15
Total	23

The Display design in spreadsheet box was checked in the Generate a Standard Design menu in Figure 6.1. So the design factors are loaded into a new spreadsheet as shown in Figure 6.5. Genstat's spreadsheet facilities can now be used to redefine the factor levels or to specify labels. To do this, you click Spread on the menu bar, followed by Factor and then either Edit Levels or Edit Labels as required.

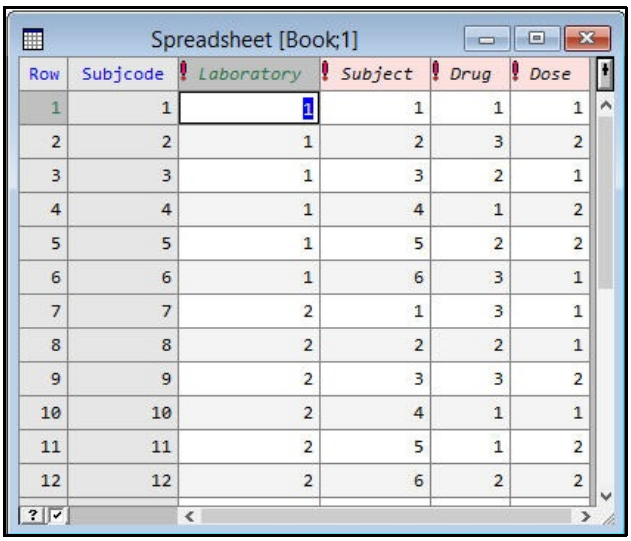


Figure 6.5

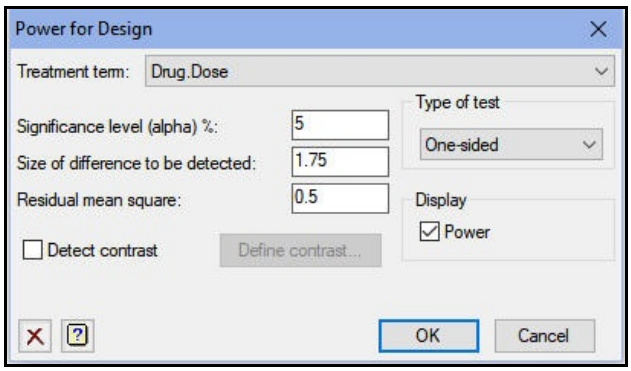


Figure 6.6

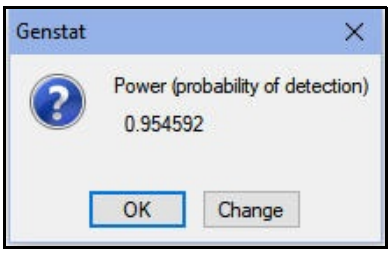


Figure 6.7

The Generate a Standard Design menu has a Check power button, which you can press once you have generated the design. This pops up the Power for Design menu, which allows you to calculate the power, or probability with which various sizes of treatment responses will be detected. In Figure 6.6 we have set the treatment term to be `Drug.Dose`, and the size of difference to be 1.75. When we click OK Genstat pops up the menu shown in Figure 6.7, telling us that the power would be 0.95.

6.2 Practical

Construct a randomized block design for three factors `Additive`, `Timing` and `Amount` with three, two and two levels, respectively. (Hint: select the design setting General Treatment structure (in randomized blocks) in the Generate a Standard Design menu. Set the number of replicates so that the design has a 90% chance (or power) to detect a difference of 1.5 in the effects of the 3-way interaction, assuming a variance within blocks (residual mean square) of 0.5 and using the F ratio with a significance level of 5%.

Your client now tells you that he cannot manage more than five replicates. What will the power now be for the detection of the interaction?

6.3 Control treatments

We now look at some of the other possibilities in the Standard Design Options menu. The Extra check box enables you to add extra replicates to the first level of any of the treatment factors. This could be useful if the first level is a control treatment against which the other levels are to be compared. When you check Extra box, the two other boxes in the top line of the menu become accessible, for you to select the factor of interest (in the right-hand box), and specify the number of extra replications. In Figure 6.8

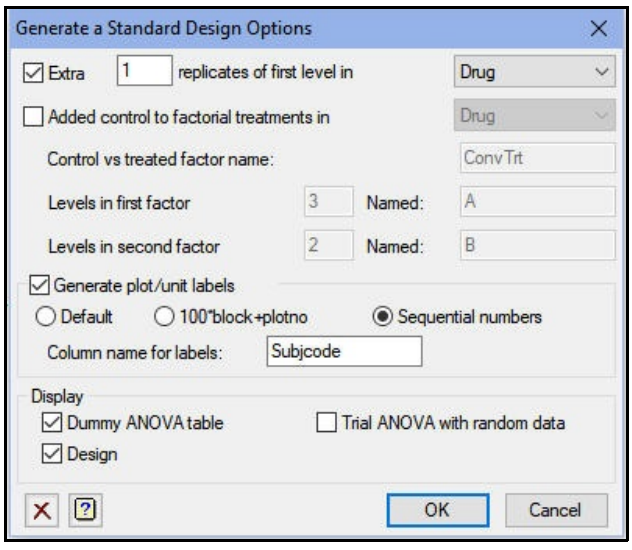


Figure 6.8

we have asked for one extra replicate for the first drug (making two replicates altogether).

The Added control to factorial treatments in box is relevant if you want to add a control treatment that is relevant to more than one treatment factor. Suppose we want to include a placebo drug in the example above. We shall now have seven treatment combinations: the six existing treatments (three drugs at two doses), and the additional placebo treatment (no drug at any dose). To set up the design, we need to revise the main menu as in Figure 6.9, to show One-way design (in randomized blocks) in the Design box, and to give a name (here `Treat`) for the factor representing the full set of treatment combinations. You do not need to set the number of levels for `Treat`, as this will be determined automatically by the options menu.

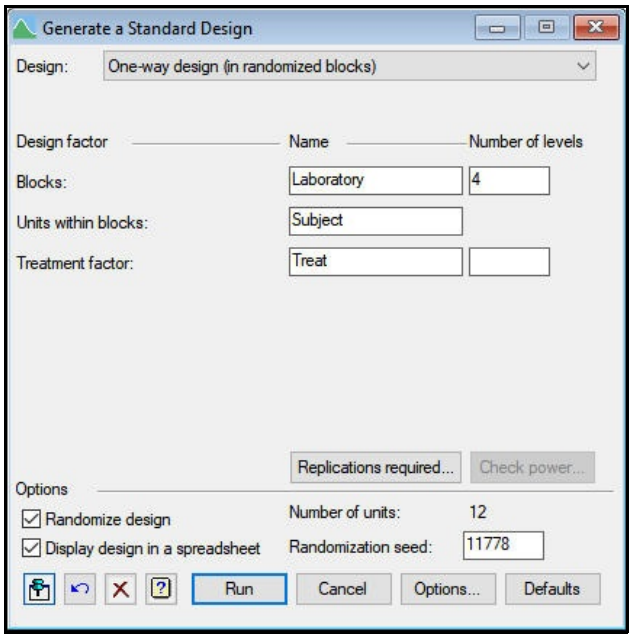


Figure 6.9

Then, in the Standard Design Options menu (Figure 6.10), we need to check the box Added control to factorial treatments in, select the factor to be subdivided into the added control plus factorial structure (here `Treat`), and specify names for the factors to represent the substructure within `Treat`. The factor `Control` represents the comparison between the placebo and any sort of drug or dose; `Drug` represents the three drugs as before, and `Dose` the doses.

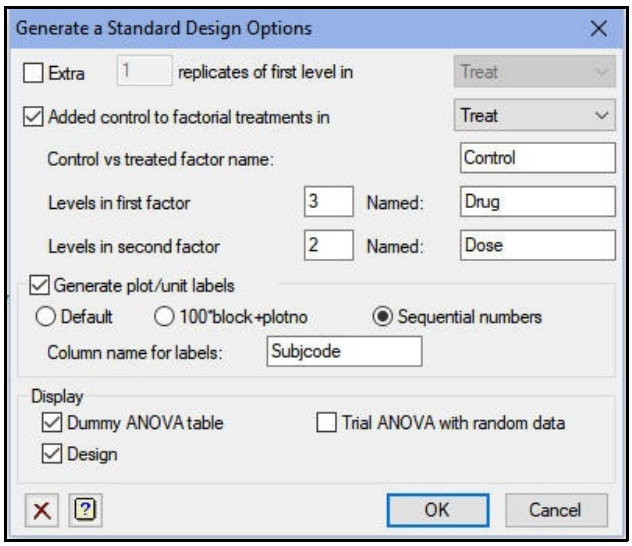
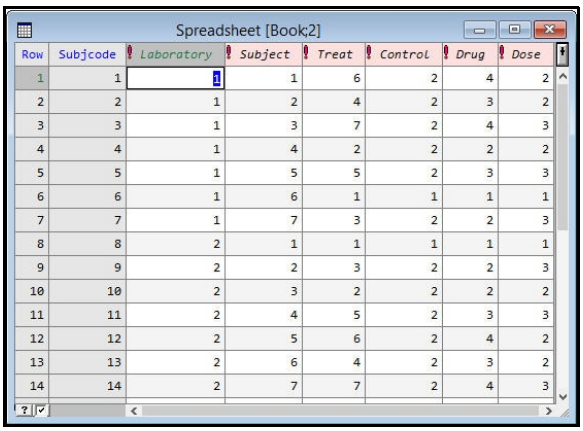


Figure 6.10

Figure 6.11 shows the spreadsheet containing the design factors, and the skeleton analysis-of-variance table is shown below. The `Control` line in the analysis of variance represents the overall effect of any drug at any (non-zero) dose, `Control.Drug` represents overall differences between the drugs (averaged over the two doses), `Control.Dose` represents the comparison between the two doses (averaged over the different drugs), and `Control.Drug.Dose` represents the interaction between `Drug` and `Dose` (assuming that some sort of drug has been taken).



Spreadsheet [Book;2]

Row	Subjcode	Laboratory	Subject	Treat	Control	Drug	Dose
1	1	1	1	6	2	4	2
2	2	1	2	4	2	3	2
3	3	1	3	7	2	4	3
4	4	1	4	2	2	2	2
5	5	1	5	5	2	3	3
6	6	1	6	1	1	1	1
7	7	1	7	3	2	2	3
8	8	2	1	1	1	1	1
9	9	2	2	3	2	2	3
10	10	2	3	2	2	2	2
11	11	2	4	5	2	3	3
12	12	2	5	6	2	4	2
13	13	2	6	4	2	3	2
14	14	2	7	7	2	4	3

Figure 6.11

Analysis of variance

Source of variation	d.f.
Laboratory stratum	3
Laboratory.Subject stratum	
Control	1
Control.Drug	2
Control.Dose	1
Control.Drug.Dose	2
Residual	18
Total	27

The "factorial plus added control" treatment structure is not one of the constructs covered directly by the Analysis of Variance menu, although the necessary *model formula* can be typed explicitly into the Treatment structure box that appears when General analysis of variance or any of the General treatment structure settings are selected in the Design box (see Section 3.5). However, the spreadsheet also contains commands to analyse the design, which can be used as an alternative to the Analysis of Variance menus, when the data values have been collected and entered as extra columns in the spreadsheet. The menu to run these commands is obtained by clicking Spread on the menu bar and selecting Sheet, followed by Analysis.

Genstat provides several more-specialized types of design. These are obtained by selecting Design from the Stats menu and then clicking on Select Design.

6.4 Practical

Modify the design that you set up in Practical 6.2 so that the first additive has twice as many replicates as the second and third additives.

7 Balance and non-orthogonality

In this chapter you will learn

- how treatment terms can be *confounded* with block terms ★
- the meaning of the *efficiency factor*, which measures how much information on a treatment term is contained in each stratum ★
- how means are formed when treatments are estimated in several strata ★
- the conditions for a design to be balanced, and analysable by the Genstat ANOVA directive ★
- how to analyse unbalanced designs with two treatment factors, using the One- and two-way Analysis of Variance menu
- how to analyse unbalanced designs with several treatment factors, using the Unbalanced ANOVA menu

Note: the topics marked ★ are optional.

7.1 Confounding and efficiency factors

In the split-plot design it is the main effect of one of the treatment factors that is estimated in the higher stratum. Statistically, we would say that this main effect is *confounded* with whole plots within blocks. For the factor `Variety` in Section 5.1, this is completely acceptable; the main interest in the trial was to look at the `Nitrogen` factor and the interaction between `Nitrogen` and `Variety`. However, on other occasions, we may want all the main effects to be estimated with the extra precision that should be available in the bottom stratum, and so we may want the interactions to be estimated in the higher strata instead.

n 0	0 *k*	*n* 0	0 *k*	0 0	0 0	*n k*	*n k*

The plan above shows a design in which the interaction between the factors `N` and `K` is confounded with blocks. The definition of the `N` × `K` interaction is that it is the difference between the effect of `N` estimated at the different levels of `K`. Here we have factors at two levels 0 and *n* for `N`, and 0 and *k* for `K`. For the 0 level of `K`, the effect of adding `N` is given by the mean of the plots with the combination (*n*, 0) minus the mean of the plots with (0, 0); while for `K` at level *k*, it is given by the mean of the plots with (*n*, *k*) minus the mean of the plots with (0, *k*). So the difference between the two estimates (which gives the interaction contrast) is

{ mean of plots with (*n*, 0) + mean of plots with (0, *k*) }
− { mean of plots with (0, 0) + mean of plots (*n*, *k*) }

The left-hand block above contains only combinations (*n*, 0) and (0, *k*), while the right-hand block contains only combinations (0, 0) and (*n*, *k*). Consequently the difference between the means of the plots in the two blocks also estimates the interaction: that is, the `N` × `K` interaction is *confounded* with blocks.

Usually, in a situation like this, you would have more than two blocks. In fact, the two blocks above are part of a design with eight blocks, each with four plots, that was used to study factors `N`, `K` and `D` (see Yates, 1937, *Design and Analysis of Factorial Experiments*, page 21; also John, 1972, *Statistical Design and Analysis of Experiments*, page 135). The left-hand block in the plan is block 3 of the design, and the right-hand block is block 4. If we analyse just those two blocks with treatment model `N*K`, the analysis of variance table below confirms that the interaction is estimated in the `Blocks` stratum (and, as we have analysed only these two blocks, there are no degrees of freedom left over for the residual).

Spreadsheet [Potatoes.gsh]

Row	Blocks	Plots	N	K	D	Yield of potatoes in tons/acre
1	1	1	0	0	0	2.71
2	1	2	N	K	0	7.79
3	1	3	N	0	D	9.99
4	1	4	0	K	D	10.66
5	2	1	N	0	0	2.84
6	2	2	0	K	0	7.10
7	2	3	0	0	D	8.36
8	2	4	N	K	D	12.05
9	3	1	N	0	0	2.38
10	3	2	0	K	0	7.29
11	3	3	N	0	D	9.05
12	3	4	0	K	D	10.90
13	4	1	0	0	0	2.84
14	4	2	0	0	D	8.68
15	4	3	N	K	0	8.20
16	4	4	N	K	D	12.03

Figure 7.1

Analysis of variance

Variate: Yield of potatoes in tons/acre

Source of variation	d.f.	s.s.	m.s.	v.r.
Blocks stratum				
N.K	1	0.56	0.56	
Blocks.*Units* stratum				
N	1	0.48	0.48	0.04
K	1	29.86	29.86	2.25
Residual	4	53.17	13.29	
Total	7	84.06		

0 0 0	*n k* 0	*n* 0 *d*	0 *k d*	*n* 0 0	0 *k* 0	0 0 *d*	*n k d*
n 0 0	0 *k* 0	*n* 0 *d*	0 *k d*	0 0 0	0 0 *d*	*n k* 0	*n k d*
n 0 0	0 0 *d*	*n k* 0	0 *k d*	0 0 0	0 *k* 0	*n* 0 *d*	*n k d*
0 *k* 0	0 0 *d*	*n k* 0	*n* 0 *d*	0 0 0	*n* 0 0	0 *k d*	*n k d*

The plan for the whole design, above, illustrates some further sophistication. It is set up so that `N.K.D` is confounded in blocks 1 and 2, `N.K` in blocks 3 and 4, `N.D` in blocks 5 and 6, and `K.D` in blocks 7 and 8. Thus, for example, `N.K` is estimated *between* blocks 3 and 4, and *within* blocks 1, 2, 5, 6, 7 and 8. So 6/8 of the information about `N.K` is in the `Blocks.Plots` stratum, and 2/8 is in the `Blocks.Plots` stratum. The main effects of `N`, `K` and `D` can be estimated in every block: they are *orthogonal* to blocks and all their information is in the `Blocks.Plots` stratum.

The amount of information available about a term in a particular stratum is known as its *efficiency factor*. The efficiency factors of non-orthogonal terms (i.e. those whose efficiency is less than one) are listed in the Information Summary, which can be obtained by checking the Information box in the ANOVA Options menu.

The whole design can be analysed using the general Analysis of Variance menu, with the Design drop-down list box set to General analysis of variance, the Block structure set to `Blocks/Plots`, and the Treatment structure set to `N*K*D`. The analysis is shown below.

Analysis of variance

Variate: Yield of potatoes in tons/acre

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Blocks stratum					
N.K	1	0.5597	0.5597	3.02	0.180
N.D	1	0.1981	0.1981	1.07	0.377
K.D	1	1.8340	1.8340	9.91	0.051
N.K.D	1	0.0807	0.0807	0.44	0.556
Residual	3	0.5554	0.1851	0.81	
Blocks.Plots stratum					
N	1	2.4863	2.4863	10.86	0.004
K	1	115.6375	115.6375	505.21	<.001
D	1	200.0482	200.0482	873.99	<.001
N.K	1	0.0202	0.0202	0.09	0.770
N.D	1	1.2934	1.2934	5.65	0.029
K.D	1	8.2713	8.2713	36.14	<.001
N.K.D	1	0.0326	0.0326	0.14	0.711
Residual	17	3.8911	0.2289		
Total	31	334.9085			

Tables of means

Variate: Yield of potatoes in tons/acre

Grand mean 7.81

N	O	N
	7.53	8.09

K	O	K
	5.91	9.71

D	O	D
	5.31	10.31

N	K	O	K
O		5.66	9.40
N		6.16	10.02

N	D	O	D
O		5.26	9.80
N		5.36	10.82

K	D	O	D
O		2.82	9.00
K		7.80	11.62

	K	O		K	
N	D	O	D	O	D
O		2.84	8.48	7.69	11.12
N		2.80	9.52	7.91	12.13

Standard errors of differences of means

Table	N	K	D	N K
rep.	16	16	16	8
d.f.	17	17	17	17
s.e.d.	0.169	0.169	0.169	0.239
Except when comparing means with the same level(s) of				
N				0.258
K				0.258

Table	N D	K D	N K D
rep.	8	8	4
d.f.	17	17	17
s.e.d.	0.239	0.239	0.352
Except when comparing means with the same level(s) of			
N	0.258		0.365
K		0.258	0.365
D	0.258	0.258	0.365
N.K			0.378
N.D			0.378
K.D			0.378

As in Practical 2.2, the y-variate (`Yield`) has a description "of potatoes in tons/acre" associated with it. (You can see how to define one of these, by putting the cursor into the `Wear` column of the spreadsheet, and clicking on Spread on the menu bar, followed by Column and then Rename.) Notice how the description is appended to the variate name in the output, to provide additional annotation.

The means produced by `ANOVA` take the effects of each term only from the lowest stratum where it is estimated. Thus the effects for `N.K` are taken from the `Blocks.Plots` stratum. The different efficiency factors for the component terms of the two-way and three-way tables of means in the example lead to different standard errors for some comparisons. For example, the s.e.d. for the `N.K.D` table is 13.15 when comparing means with different levels of all three factors, it is 13.64 if the level of one of the factors is identical for both means, and it is 14.12 if two of the factors are at identical levels.

The effects from the lowest stratum are usually those that are estimated most precisely; the lower strata generally have smaller mean squares and, in most designs, terms will have higher efficiency factors in the lower strata. Moreover, under the usual assumptions of Normality of residuals, differences between the means can be tested by the usual t-statistics. Nevertheless, for prediction you will often want to present means and effects that combine the information about each term from all the strata where it is estimated. Provided the design is a *generally-balanced design*, these can be requested using the ANOVA Options menu or the ANOVA Further Output menu (Figure 7.2). Payne & Tobias (1992, *Scandinavian Journal of Statistics*, **19**, 3-23) give a full definition of the method and of the design properties. However, you do not need to know the details – Genstat checks the design automatically and will let you know if it is not generally balanced.

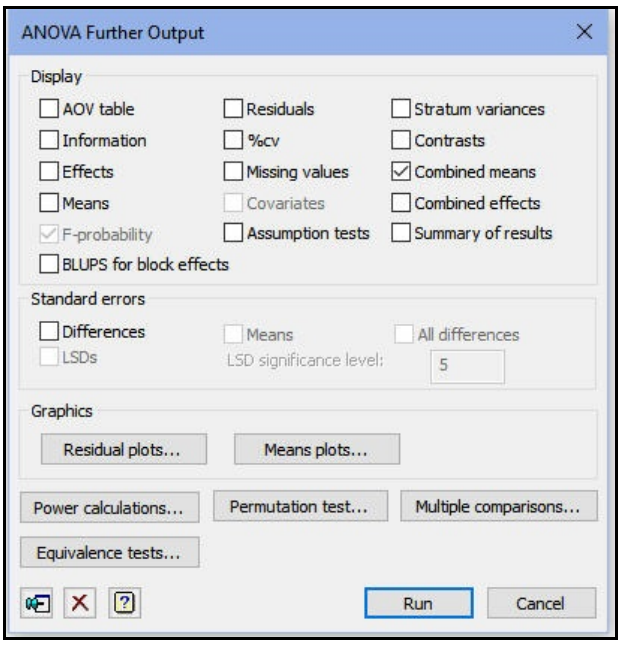


Figure 7.2

The combined means for the potato example are shown below.

Tables of combined means

Variate: Yield of potatoes in tons/acre

N	O	N
	7.53	8.09

K	O	K
	5.91	9.71

D	O	D
	5.31	10.31

N	K	O	K
O		5.71	9.35
N		6.11	10.07

N	D	O	D
O		5.18	9.89
N		5.44	10.73

K	D	O	D
O		2.85	8.97
K		7.77	11.65

N	K D	O O	D	K O	D
O		2.85	8.58	7.51	11.20
N		2.85	9.36	8.04	12.10

Standard errors of differences of combined means

Table	N	K	D	N K
rep.	16	16	16	8
s.e.d.	0.170	0.170	0.170	0.241
effective d.f.	17.90	17.90	17.90	17.90
Except when comparing means with the same level(s) of				
N				0.243
effective d.f.				21.76
K				0.243
effective d.f.				21.76

Table	N D	K D	N K D
rep.	8	8	4
s.e.d.	0.241	0.241	0.342
effective d.f.	17.90	17.90	19.94
Except when comparing means with the same level(s) of			
N	0.243		0.344
effective d.f.	21.76		21.76
K		0.243	0.344
effective d.f.		21.76	21.76
D	0.243	0.243	0.344
effective d.f.	21.76	21.76	21.76
N.K			0.345
effective d.f.			23.14
N.D			0.345
effective d.f.			23.14
K.D			0.345
effective d.f.			23.14

The effective d.f. are calculated by an algorithm based on Satterthwaite's method (Payne 2004, *COMPSTAT 2004 Proceedings in Computational Statistics*, 1629-1636), and can be used for approximate t-tests for differences between means. For further information, see the *Guide to the Genstat Command Language*, Part 2, Section 4.7.1.

7.2 Balance

The designs that are analysable by the ANOVA directive must have the property of *first-order balance*. Essentially this requires the contrasts of each term to all have a single efficiency factor, wherever the term is estimated. In the example in Section 7.1, all the terms have only one degree of freedom, and so represent only one contrast. So it is clear that the design is balanced.

Suppose instead that the treatment combinations were represented by a single factor T with eight levels:

```
FACTOR [LABELS=!T('OOO','OOD','OKO','OKD',\
                  'NOO','NOD','NKO','NKD')] T
```

The main effect of T would not be balanced: the comparison of levels

{'OOO' 'OOD' 'OKO' 'OKD'}

with {'NOO' 'NOD' 'NKO' 'NKD'}

has efficiency factor one in the Blocks.Plots stratum and zero in the Blocks stratum (this contrast is equivalent to the main effect of N in the original specification); but the comparison of levels

{'NOO' 'NOD' 'OKO' 'OKD'}

with {'OOO' 'OOD' 'NKO' 'NKD'}

has efficiency 0.25 in the Blocks stratum and 0.75 in the Blocks.Plots stratum (this is equivalent to N.K in the original specification). Thus the main effect of T is not balanced, since in the Block.Plots stratum some of its contrasts have efficiency factor one, while others have efficiency factor 0.75. Genstat can detect unbalanced designs like this, and will give you an error diagnostic.

Fault 23, code AN 1, statement 1 on line 78

Command: ANOVA
Design unbalanced - cannot be analysed by ANOVA.
Model term T (non-orthogonal to term Blocks) is unbalanced, in the Blocks.Plots stratum.

It is still possible to analyse this particular design by ANOVA, by defining pseudo-factors (see *Guide to the Genstat Command Language*, Part 2, Section 4.7.3). However, this requires extra skill for the specification, and it may not be feasible in many cases. So, if you have a single error term, you can use the Unbalanced ANOVA menu (Section 7.4). Alternatively, if you have several error terms you can use the REML menus (Chapter 8).

7.3 Practical

Factorial designs with interactions confounded with blocks can be constructed using the Generate Factorial Designs in Blocks menu, which can be opened by clicking on the Generate a Factorial Design in Blocks sub-option of the Design option of the Stats menu (Figure 7.3).

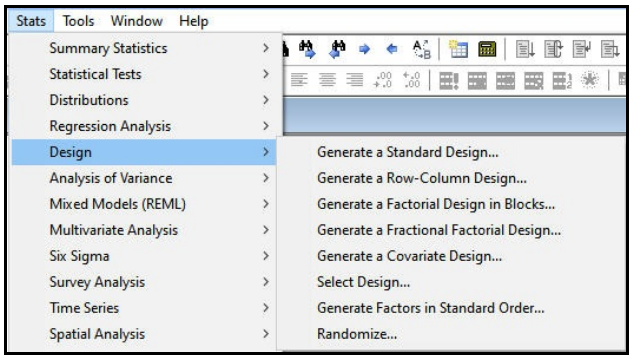


Figure 7.3

Use the menu, as shown in Figure 7.4, to construct a design for a single replicate of a 2×2×2×2 design in blocks of size 8.

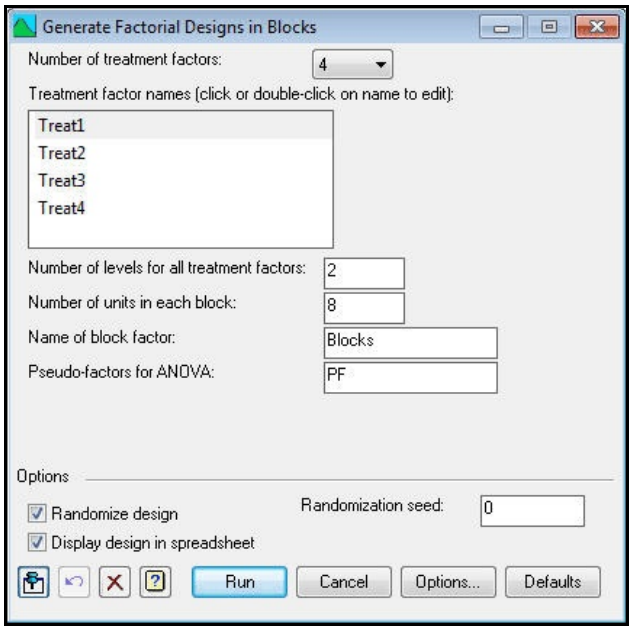
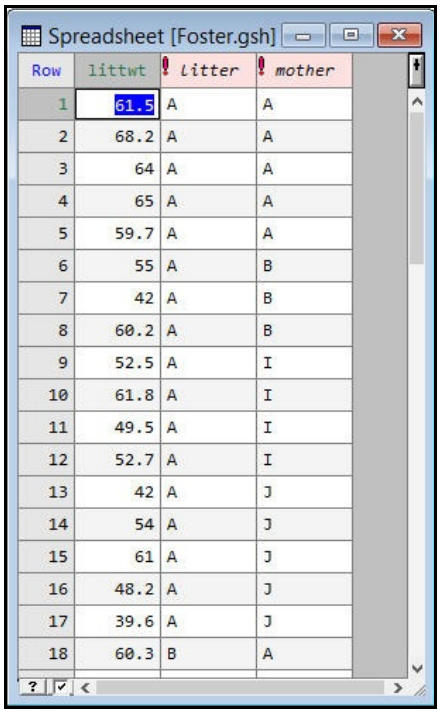


Figure 7.4

7.4 Unbalanced designs with two treatment factors

Most of the designs covered by the Analysis of Variance menus are *balanced* and, in fact, all of those discussed so far in the earlier chapters have been *orthogonal*. Essentially this means that the order in which the treatment terms are fitted is unimportant (other than that each main effect must be fitted before any of its interactions). So we could have specified sulphur as the first treatment factor and nitrogen as the second treatment factor in the menus in Figures 3.2 and 3.5, and still have obtained the same sums of squares and effects. This contrasts with the situation in multiple linear regression (see e.g. Section 5.2 of the *Introduction to Genstat for Windows*), where the x-variates are usually correlated (i.e. non-orthogonal), and so different regression coefficients are obtained for each x-variate according to which other x-variates had been fitted beforehand.

Genstat spreadsheet file `Foster.gsh` (Figure 7.5) contains the results of an experiment to study the effect of foster feeding of rats (Scheffe, 1959, *The Analysis of Variance*; also see McConway, Jones & Taylor, 1999, *Statistical Modelling using GENSTAT*, Example 7.6). The rats were from four different genotypes (A, B, I or J), the experimental unit was a litter of four rats, and the response variate was the weight of the litter after a period of feeding. The interest was in whether the genotype of a foster mother would affect the weight. So there are two treatment factors, each with four levels, the



Row	littwt	litter	mother
1	61.5	A	A
2	68.2	A	A
3	64	A	A
4	65	A	A
5	59.7	A	A
6	55	A	B
7	42	A	B
8	60.2	A	B
9	52.5	A	I
10	61.8	A	I
11	49.5	A	I
12	52.7	A	I
13	42	A	J
14	54	A	J
15	61	A	J
16	48.2	A	J
17	39.6	A	J
18	60.3	B	A

Figure 7.5

genotype of the mother and the genotype of the foster mother. It was impossible to balance the numbers of litters over the two factors, and so the design is unbalanced.

The One- and two-way Analysis of Variance menu (Figure 7.6) automatically detects that a design is unbalanced, and calculates the analysis instead by using the Genstat regression commands.

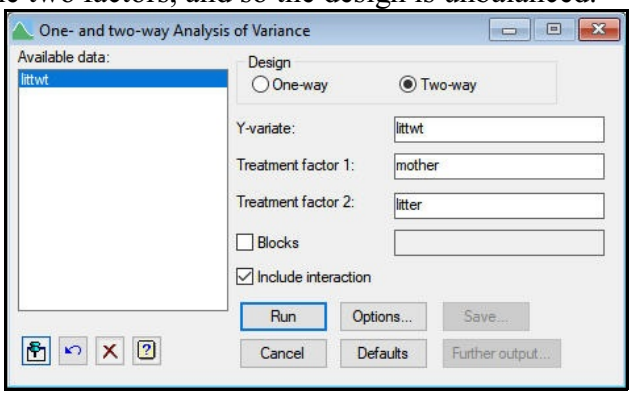


Figure 7.6

The analysis-of-variance table is modified so that it shows the effect of fitting each of the factors either before or after the other one. So the line "`mother ignoring litter`" fits the effect of `mother` first. The alternative line "`mother elimining litter`" fits the effect of `mother` after fitting the `litter` effect. So it looks to see if there are any effects of the foster mother that cannot be explained by the genotype of the litter itself. (Remember, though, that interactions are always fitted after their main effects.)

Notice that the means are now predicted means (from the Genstat `PREDICT` directive). These are accompanied by a summary of the standard errors of difference over the pair of means within the table. You can print s.e.d.'s for every possible comparison of pairs of means within the table, by using the Unbalanced ANOVA menu, as shown in Section 7.6.

Analysis of variance

Source	d.f.	s.s.	m.s.	v.r.	F pr.
mother ignoring litter	3	771.61	257.20	4.74	0.006
mother eliminating litter	3	775.08	258.36	4.76	0.006
litter ignoring mother	3	60.16	20.05	0.37	0.775
litter eliminating mother	3	63.63	21.21	0.39	0.760
mother.litter	9	824.07	91.56	1.69	0.120
Residual	45	2440.82	54.24		
Total	60	4100.13	68.34		

Grand mean

53.97

Predictions from regression model

Response variate: littwt

	Prediction
mother	
A	54.79
B	58.08
I	53.60
J	48.34

Minimum standard error of difference	2.641
Average standard error of difference	2.753
Maximum standard error of difference	2.863

Predictions from regression model

Response variate: littwt

litter	Prediction
A	54.97
B	53.07
I	52.82
J	53.50

Minimum standard error of difference	2.659
Average standard error of difference	2.755
Maximum standard error of difference	2.848

Predictions from regression model

Response variate: littwt

	Prediction			
litter	A	B	I	J
mother				
A	63.68	52.3	47.10	54.35
B	52.40	60.64	64.37	56.10
I	54.13	53.93	51.60	54.53
J	48.96	45.90	49.43	49.06

Minimum standard error of difference	4.658
Average standard error of difference	5.499
Maximum standard error of difference	6.723

7.5 Practical

Spreadsheet file `Unbalanced2way.gsh` (Figure 7.7) contains results from an experiment with two factors `A` and `B`. Analyse the response variate `Y` using the One- and two-way Analysis of Variance menu.

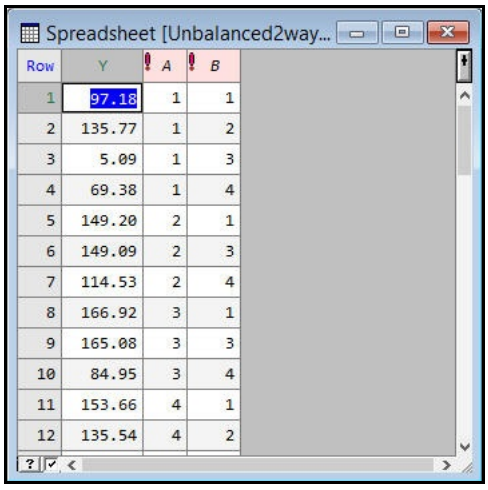
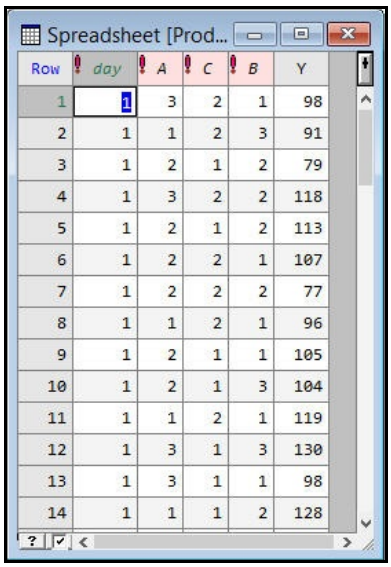


Figure 7.7

7.6 Unbalanced designs with several treatment factors



Row	day	A	C	B	Y
1	1	3	2	1	98
2	1	1	2	3	91
3	1	2	1	2	79
4	1	3	2	2	118
5	1	2	1	2	113
6	1	2	2	1	107
7	1	2	2	2	77
8	1	1	2	1	96
9	1	2	1	1	105
10	1	2	1	3	104
11	1	1	2	1	119
12	1	3	1	3	130
13	1	3	1	1	98
14	1	1	1	2	128

Figure 7.8

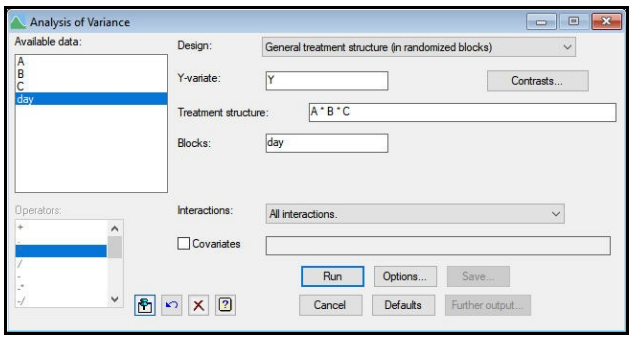


Figure 7.9

Genstat spreadsheet file `Product.gsh`, displayed in Figure 7.8 contains the results of an experiment to study the effects of factors `A`, `B` and `C` on the yield `Y` of a production process. The intention was originally to run the experiment in two separate days, and to have two observations of each treatment combination on each day. However, due to time constraints, there were several combinations (chosen at random) in each of the days that could only be performed once.

If the design had been constructed with equal replication, as planned, it could have been analysed using the General treatment structure (in randomized blocks) design setting. The block factor would be `day`, and the treatment structure would be a factorial with three factors: `A*B*C`, as shown in Figure 7.9. However, this generates a fault message (below) reporting that the design is unbalanced.

Fault 27, code AN 1, statement 1 on line 37

Command: ANOVA [PRINT=aovtable,information,means; FACT=32; CONTRASTS=7; PCONTRAS
Design unbalanced - cannot be analysed by ANOVA.
Model term A.B (non-orthogonal to term day) is unbalanced, in the day.*Units* stratum.

Instead we need to use the Unbalanced ANOVA menu, setting, obtained by clicking on the Unbalanced Designs line in the Analysis of Variance section of the Stats menu (see Figure 1.7). The menu, in Figure 7.10, is not customized for any particular design, but merely has two boxes to define the model to be fitted. The Blocking (nuisance terms) box contains the main effect of days as we are not interested in testing for day effects, we simply want to remove any day differences before assessing

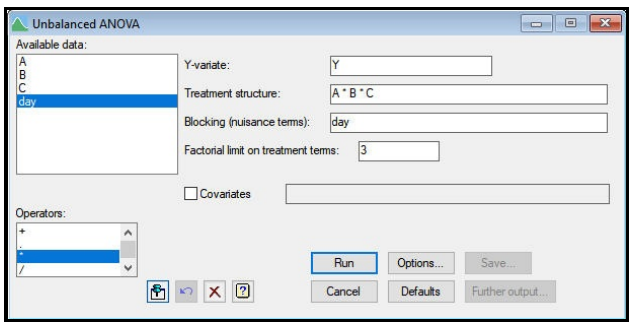


Figure 7.10

the treatments. The Treatment structure box contains a factorial model with treatment factors A, B and C.

The commands that are generated by this setting of the menu use the Genstat regression facilities (via procedure AUNBALANCED) rather than the analysis-of-variance facilities. So Genstat produces an accumulated analysis-of-variance, indicating the order in which the terms were fitted. The term day is fitted first because this is a *nuisance* term, reflecting random variability which we want to eliminate before we assess the treatments. The +A line then gives the (main) effect of A after eliminating day. The +B line gives the main effect of B, eliminating day and A, and so on. Each line in the table presents the effect of a particular term, eliminating the terms in the lines above, but ignoring the terms in the lines below. This is technically true also in the examples presented in earlier chapters but there the designs were orthogonal and so the ordering of the treatment terms was unimportant. Here if we had specified C*A*B, the sums of squares for A, B and C would have been 1699.1, 429.4 and 1063.0 respectively, and there would also have been changes to the sums of squares for the interactions. The results would have led to the same conclusions to those from the earlier order (namely that there are main effects of A and C, and an A by C interaction), but in a design with a greater degree of non-orthogonality you would be well advised to investigate several orderings.

Alternatively, the Options menu for the designs with Unbalanced Treatment Structure (Figure 7.11) contains a check box to allow you to request *screening tests*.

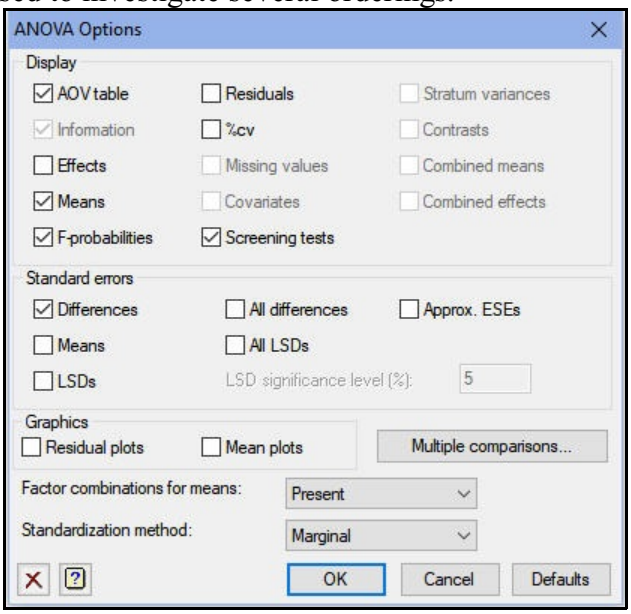


Figure 7.11

In the *marginal test* (the column headed "mtest" below) the term is added to the simplest possible model. So A.B would be added to a model containing only the main effects A and B. This assesses the effect of the term ignoring as many other terms as possible, and so it checks to see if there is any evidence for the term having an effect.

In the *conditional test* (the column headed "ctest" below) the term is added to the most complex possible model. So, A would be added to a model containing B, C and B.C. This checks to see if the term has any effect that cannot be explained by any other terms.

Ideally (as here) the tests will both lead to the same conclusion. If not, the conclusion is that there is more than one plausible model for the data, but the design is too unbalanced to allow you to choose between them.

Screening of terms in an unbalanced design

Variate: Y

Marginal and conditional test statistics, degrees of freedom and number of observations used

degrees of freedom for denominator (full model): 48

term	mtest	mdf	ctest	cdf
A	3.42	2	3.47	2
B	0.76	2	0.84	2
C	4.27	1	4.78	1

term	mtest	mdf	ctest	cdf
A.B	1.04	4	1.00	4
A.C	5.25	2	4.81	2
B.C	0.71	2	0.57	2

term	mtest	mdf	ctest	cdf
A.B.C	1.40	4	1.40	4

P-values of marginal and conditional tests

term	mprob	cprob
A	0.041	0.039
B	0.474	0.439
C	0.044	0.034

term	mprob	cprob
A.B	0.395	0.415
A.C	0.009	0.013
B.C	0.498	0.569

term	mprob	cprob
A.B.C	0.248	0.248

Analysis of an unbalanced design using Genstat regression

Variate: Y

Accumulated analysis of variance

Change	d.f.	s.s.	m.s.	v.r.	F pr.
+ day	1	914.0	914.0	3.67	0.061
+ A	2	1706.8	853.4	3.42	0.041
+ B	2	418.8	209.4	0.84	0.438
+ C	1	1065.9	1065.9	4.28	0.044
+ A.B	4	1166.0	291.5	1.17	0.336
+ A.C	2	2456.7	1228.3	4.93	0.011
+ B.C	2	284.4	142.2	0.57	0.569
+ A.B.C	4	1397.4	349.4	1.40	0.248
Residual	48	11960.4	249.2		

Total	66	21370.4	323.8

Grand mean

106.6

Predictions from regression model

Response variate: Y

A	Prediction
1	113.2
2	101.2
3	105.3

Minimum standard error of difference	4.679
Average standard error of difference	4.795
Maximum standard error of difference	4.909

Predictions from regression model

Response variate: Y

B	Prediction
1	103.2
2	108.1
3	108.3

Minimum standard error of difference	4.724
Average standard error of difference	4.788
Maximum standard error of difference	4.896

Predictions from regression model

Response variate: Y

C	Prediction
1	110.6
2	102.4

Standard error of differences between predicted means 3.903

Predictions from regression model

Response variate: Y

Prediction

B	1	2	3
A			
1	115.2	112.3	111.8
2	97.9	99.9	106.4
3	96.7	113.2	106.8

Minimum standard error of difference	7.894
Average standard error of difference	8.313
Maximum standard error of difference	9.393

Predictions from regression model

Response variate: Y

	Prediction	
C	1	2
A		
1	125.9	100.9
2	101.7	100.7
3	104.6	105.9

Minimum standard error of difference	6.454
Average standard error of difference	6.778
Maximum standard error of difference	7.103

Predictions from regression model

Response variate: Y

	Prediction	
C	1	2
B		
1	110.2	96.5
2	111.9	104.5
3	109.7	106.9

Minimum standard error of difference	6.454
Average standard error of difference	6.770
Maximum standard error of difference	7.215

Predictions from regression model

Response variate: Y

		Prediction	
	C	1	2
A	B		
1	1	136.1	95.1
	2	124.1	100.8
	3	116.2	107.6
2	1	102.1	93.8
	2	101.8	98.1
	3	101.3	111.3
3	1	92.3	101.1

2	110.6	115.8
3	112.6	101.2

Minimum standard error of difference	11.16
Average standard error of difference	11.74
Maximum standard error of difference	14.42

Like the One- and two-way Analysis of Variance menu, the Unbalanced ANOVA menu uses the PREDICT directive to form the predicted means, but it gives more control over the way in which they are formed. The first step (A) of the calculation forms the full table of predictions, classified by every factor in the model. The second step (B) averages the full table over the factors that do not occur in the table of means. The Factor combination for means box specifies which cells of the full table are to be formed in Step A. The default setting, Estimable, fills in all the cells other than those that involve parameters that cannot be estimated, for example because of aliasing. Alternatively, the setting Present excludes the cells for factor combinations that do not occur in the data. The Standardization method box then defines how the averaging is done in Step B. The default setting, Marginal, forms a table of marginal weights for each factor, containing the proportion of observations with each of its levels; the full table of weights is then formed from the product of the marginal tables. The setting Equal weights all the combinations equally. Finally, the setting Observed uses the WEIGHTS option of PREDICT to weight each factor combination according to its own individual replication in the data. The One- and two-way Analysis of Variance menu, always uses the default settings.

In an unbalanced design, there will usually be a different standard error for differences between each pair of means. Here we have simply printed a summary giving the minimum, average and maximum standard errors for differences between pairs of means. The Options menu (Figure 7.11) allows you to print a symmetric matrix giving the standard errors for differences between every possible pair of means, but this is omitted here to save space. In the earlier designs in this chapter, the treatment combinations were all equally replicated, and so the standard errors were the same for every pair of means.

7.7 Practical

Reanalyse the data in the Spreadsheet file Unbalanced2way.gsh, first analysed in Section 7.5, using the Unbalanced ANOVA menu. Print the standard errors of differences for all pairs of means. (Note, you do not have any Blocking or Nuisance terms.)

8 REML analysis of unbalanced designs

The Analysis of Variance menus, described in the earlier chapters, deal mainly with balanced designs. This ideal situation, however, is not always achievable. The randomized-block design in Section 2.2 is balanced because every block contained one of each treatment combination. However, there may sometimes be so many treatments that the blocks would become unrealistically large. Designs where each block contains less than the full set of treatments include cyclic designs and Alpha designs (both of which can be generated within Genstat by clicking Stats on the menu bar, selecting Design and then Select Design), neither of which tend to be balanced. In experiments on animals, some subjects may fail to complete the experiment for reasons unconnected with the treatments. So even an initially balanced experiment may not yield a balanced set of data for analysis. The Mixed Models (REML) menus, which use the Genstat `REML` directive, are designed to handle these situations. They also allow you to fit models to the complex correlation structures that occur in repeated measurements or in spatially-correlated data from field experiments.

In this chapter you will learn

- how to use the Linear Mixed Models menu
- what output is given by a Genstat `REML` analysis, and how it compares to Genstat `ANOVA`
- how to assess fixed terms using Wald and F statistics
- how effects and means can be produced by Genstat `ANOVA`, combining all the available information when treatment terms that are estimated in several strata ★

Note: the topics marked ★ are optional.

8.1 Linear mixed models: split-plot design

We start by reanalysing the split-plot data (Oats.gsh) in Section 5.1, to highlight the differences and similarities between REML and ANOVA.

Figure 8.1 shows the Linear Mixed Models menu, obtained by clicking Stats on the menu bar and selecting Mixed Models (REML), followed by Linear Mixed Models. The Fixed model box corresponds to the Treatment structure box in the split-plot menu, and specifies the terms defining the *fixed* effects in the model to be fitted. The Linear Mixed Models menu provides general facilities covering any type of design, and so the *random* effects are defined explicitly by the contents of the Random model box, instead of being defined automatically as in the split-plot menu. The model is the same though, namely

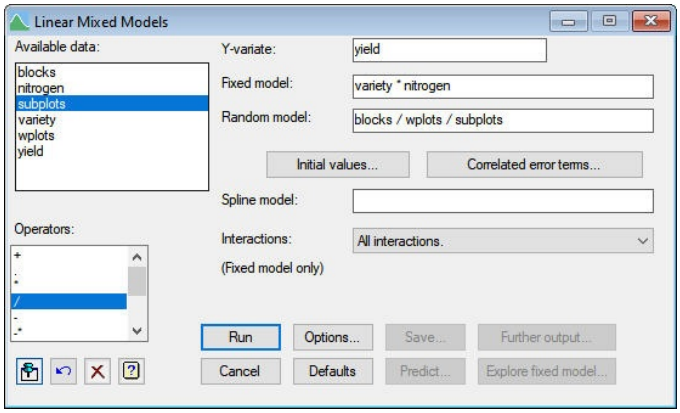


Figure 8.1

```
blocks/wplots/subplots
```

which expands to give the three (random) terms; see Section 3.4.

```
block + block.wplot + block.wplot.subplot
```

Similarly, the fixed model

```
variety * nitrogen
```

expands as before to

```
variety + nitrogen + nitrogen.variety
```

to request that Genstat fits the main effects of nitrogen and variety, and their interaction. (The Interactions box, which operates just like the one in the Analysis of Variance menu, has requested all interactions in the fixed model to be included.)

The Options button produces the Linear Mixed Model Options menu, shown in Figure 8.2. The standard model options (as shown in the figure) are fine for this design, so we need only select the output to display (and then click OK).

Returning to the main menu (Figure 8.1): initial values are seldom required for simple REML analyses like this, and the Spline model box is not relevant (this is mainly useful with repeated measurements), so we can click on Run and generate the output shown below.

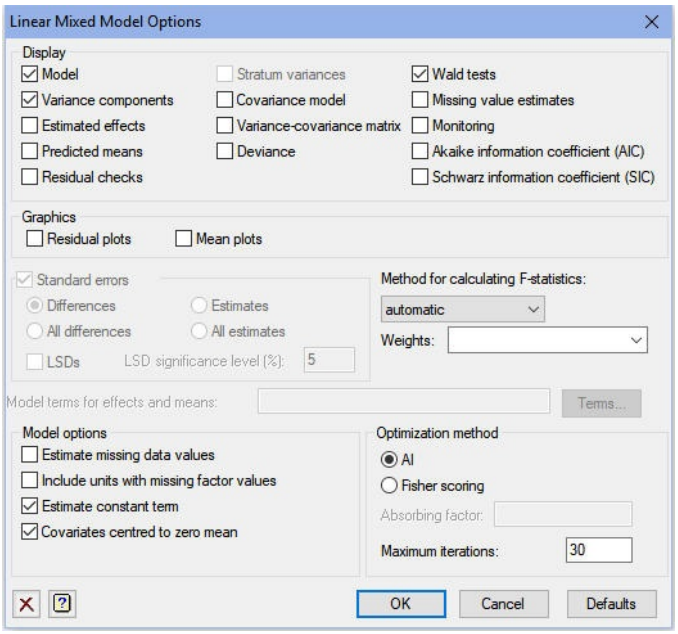


Figure 8.2

REML variance components analysis

Response variate:	yield
Fixed model:	Constant + variety + nitrogen + variety.nitrogen
Random model:	blocks + blocks.wplots + blocks.wplots.subplots
Number of units:	72

blocks.wplots.subplots used as residual term

Sparse algorithm with AI optimisation

Estimated variance components

Random term	component	s.e.
blocks	214.5	168.8
blocks.wplots	106.1	67.9

Residual variance model

Term	Model(order)	Parameter	Estimate	s.e.
blocks.wplots.subplots	Identity	Sigma2	177.1	37.3

Tests for fixed effects

Sequentially adding terms to fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
variety	2.97	2	1.49	10.0	0.272
nitrogen	113.06	3	37.69	45.0	<0.001
variety.nitrogen	1.82	6	0.30	45.0	0.932

Dropping individual terms from full fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
variety.nitrogen	1.82	6	0.30	45.0	0.932

Message: denominator degrees of freedom for approximate F-tests are calculated using algebraic derivatives ignoring fixed/boundary/singular variance parameters.

The output first lists the terms in the fixed and random model, and indicates the residual term. The residual term is a random term with a parameter for every unit in the design. Here we have specified a suitable term, `blocks.wplots.subplots`, explicitly. However, if we had specified only `blocks` and `blocks.wplots` as the Random Model (for example by putting `blocks/wplots`), Genstat would have added an extra term `*units*` to act as residual. (`*units*` would be a private factor with a level for every unit in the design.)

Genstat estimates a *variance component* for every term in the random model, apart from the residual. The variance component measures the inherent variability of the term, over and above the variability of the sub-units of which it is composed. Generally, this is positive, indicating that the units become more variable the larger they become. So here the whole-plots are more variable than the subplots, and the blocks are more variable than the whole-plots within the blocks. (This is the same conclusion that you would draw from the analysis-of-variance table in Section 5.1 and, in fact, you can also produce the variance components as part of the stratum variances output from the Analysis of Variance menu.) However, the variance component can sometimes be *negative*, indicating that the larger units are *less* variable than you would expect from the contributions of the sub-units of which they are composed. This could happen if the sub-units were negatively correlated.

The section of output summarising the residual variance model indicates that we have not fitted any specialized correlation model on this term (see the column headed `Model`), and gives an estimate of the residual variance; this is the same figure as is given by the mean square in the residual line in the `blocks.wplots.subplots` stratum in the split-plot analysis-of-variance table.

The next section, however, illustrates a major difference between the two analyses. When the design is balanced, Genstat is able to partition the variation into *strata* with an appropriate random error term (or residual) for each treatment term (see Section 5.1). No such partitioning is feasible for the unbalanced situations that `REML` is designed to handle. Instead Genstat produces a *Wald statistic* to assess each fixed term.

If the design is orthogonal, the Wald statistic is equal to the treatment sum of squares divided by the stratum residual mean square. So under the usual assumption that the residuals come from Normal distributions, the Wald statistic divided by its degrees of freedom will have an F distribution, $F_{m,n}$, where m is the number of degrees of freedom of the fixed term, and n is the number of residual degrees of freedom for the fixed term.

By default, unless the design is large or complicated, Genstat estimates *n*, and prints it in the column headed "d.d.f." (i.e. *denominator degrees of freedom*); *m* is shown the column headed "n.d.f." (i.e. *numerator degrees of freedom*). For orthogonal designs, the F statistics and probabilities are identical to those produced by the Analysis of Variance menus, and can be used in exactly the same way. In other situations, the printed F statistics have approximate F distributions. So you need to be careful if the value is close to a critical value.

The Linear Mixed Model Options menu (Figure 8.2) has a list box Method for calculating F statistics to control how and whether to calculate the F statistics. With the default setting, automatic, Genstat itself decides whether the statistics can be calculated quickly enough to be useful, and the best method to use. The other settings allow you to select to use either algebraic or numerical derivatives, or to print just Wald statistics (none).

The Wald statistics themselves would have exact χ^2 distributions if the variance parameters were known but, as they must be estimated, they are only asymptotically distributed as χ^2. In practical terms, the χ^2 values will be reliable if the residual degrees of freedom for a fixed term is large compared to its own degrees of freedom. Otherwise they tend to give significant results rather too frequently. The F statistics, if available, are more reliable than the Wald statistics. If they are not calculated, Genstat produces probabilities for the Wald statistics instead, which should again be used with care especially when the value is close to a critical value.

In the example, the treatment terms are *orthogonal* so it makes no difference whether nitrogen or variety is fitted first. In a non-orthogonal design, however, the ordering of fitting is important, and you should be aware that each test in the "Sequentially adding terms to fixed model" section represents the effect of adding the term concerned to a model containing all the terms in the preceding lines. The next section, headed "Dropping individual terms from full fixed model" looks at the effect of removing terms from the complete fixed model: so the lines here allow you to assess the effects of a term after eliminating all the other fixed terms. This is particularly useful for seeing how the model might be simplified. Notice that the only relevant term here is the variety by nitrogen interaction. We cannot remove a main effect (such as nitrogen or variety) from a model that contains an interaction involving that factor.

The Further output button generates the Linear Mixed Models Further Output menu. In Figure 8.3, we have checked the boxes to produce tables of predicted means and standard errors of differences between means. The Model terms for effects and means box enables you to specify the terms for which you want tables of means (and, if you had checked the Estimated effects box, tables of effects). The default, which is fine here, is to produce a table for each term in the fixed model. Clicking Run then generates the tables shown below. Because the fixed terms are orthogonal, the means are identical to those produced by the Analysis of Variance menu (Section 5.1).

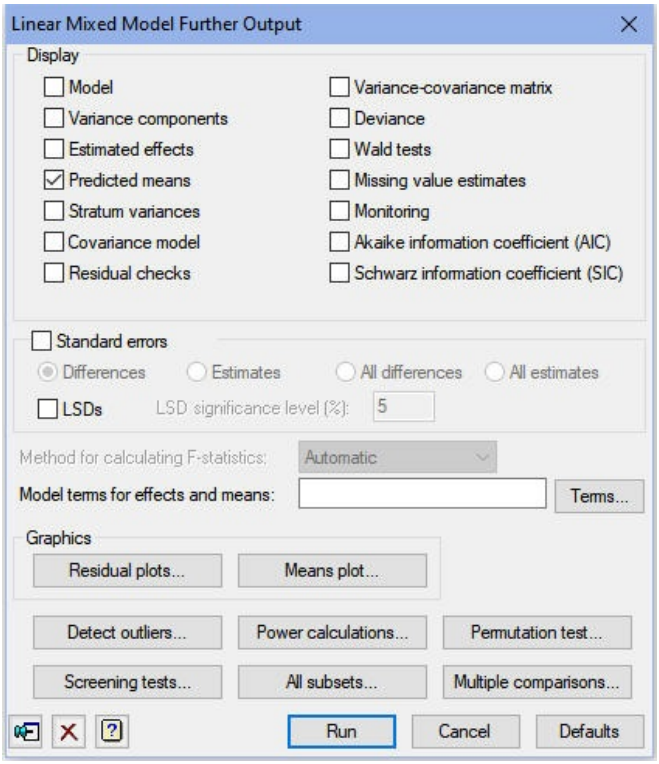


Figure 8.3

Table of predicted means for Constant

104.0 Standard error: 6.64

Table of predicted means for variety

variety	Victory	Golden rain	Marvellous
	97.6	104.5	109.8

Standard error of differences: 7.079

Table of predicted means for nitrogen

nitrogen	0 cwt	0.2 cwt	0.4 cwt	0.6 cwt
	79.4	98.9	114.2	123.4

Standard error of differences: 4.436

Table of predicted means for variety.nitrogen

nitrogen variety	0 cwt	0.2 cwt	0.4 cwt	0.6 cwt
Victory	71.5	89.7	110.8	118.5
Golden rain	80.0	98.5	114.7	124.8
Marvellous	86.7	108.5	117.2	126.8

Standard errors of differences

Average:	9.161
Maximum:	9.715
Minimum:	7.683

Average variance of differences: 84.74

Standard error of differences for same level of factor:

	variety	nitrogen
Average:	7.683	9.715
Maximum:	7.683	9.715
Minimum:	7.683	9.715

The REML facilities thus produce the same information as that given by the Analysis of Variance menu where this is possible in their more general context, but they are not able to match its more specialized output. The advantage of the REML menus, however, lies in the fact that they can also analyse unbalanced designs.

8.2 Practical

Use the Linear Mixed Models menu to reanalyse the experiment on meat-tenderizing chemicals (spreadsheet file Meat.gsh), but without fitting the polynomials to temperature. Compare the analysis with the split-plot analysis, originally performed in Section 5.2, using the Analysis of Variance menu.

8.3 Linear mixed models: a non-orthogonal design

We now consider the analysis of a rather more complicated field experiment (at Slate Hall Farm in 1976), previously analysed by Gilmour *et al*. (1995). The design was set up to study 25 varieties of wheat, and contained six replicates (each with one plot for every variety) laid out in a two by three array. The variety grown on each plot is shown in the plan below.

Each replicate has a block structure of rows crossed with columns, so the random model is

```
replicates / (rows * columns)
```

(rows crossed with columns, nested within replicates), which expands to give

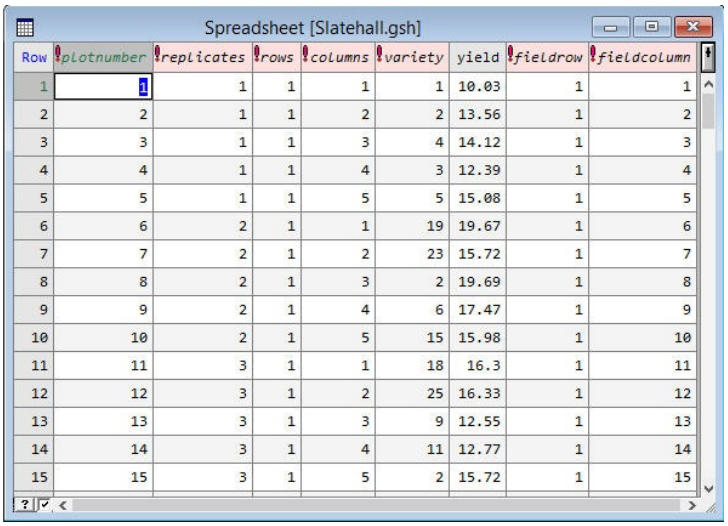
```
replicates + replicates.rows + replicates.columns +
```

`replicates.rows.columns`

So we have random terms for replicates, rows within replicates, columns within replicates and, finally, `replicates.rows.columns` represents the residual variation. The fixed model contains just the main effect of the factor `variety`.

1	2	4	3	5	19	23	2	6	15	18	25	9	11	2
6	7	9	8	10	8	12	16	25	4	5	7	16	23	14
21	22	24	23	25	11	20	24	3	7	6	13	22	4	20
11	12	14	13	15	22	1	10	14	18	24	1	15	17	8
16	17	19	18	20	5	9	13	17	21	12	19	3	10	21
3	18	8	13	23	16	24	10	13	2	10	4	17	11	23
1	16	6	11	21	12	20	1	9	23	12	6	24	18	5
5	20	10	15	25	4	7	18	21	15	19	13	1	25	7
2	17	7	12	22	25	3	14	17	6	21	20	8	2	14
4	19	9	14	24	8	11	22	5	19	3	22	15	9	16

Figure 8.4 shows a Genstat spreadsheet file, stored as `Slatehall.gsh`, containing the data. As well as the factors already mentioned, the sheet also contains factors `fieldrow` and `fieldcolumn` (defining the row and column positions within the whole field, rather than within each replicate). Chapter 3 of the *Guide to REML in Genstat for Windows* shows how these can be used to define spatial correlation structures.



Row	plotnumber	replicates	rows	columns	variety	yield	fieldrow	fieldcolumn
1	1	1	1	1	1	10.03	1	1
2	2	1	1	2	2	13.56	1	2
3	3	1	1	3	4	14.12	1	3
4	4	1	1	4	3	12.39	1	4
5	5	1	1	5	5	15.08	1	5
6	6	2	1	1	19	19.67	1	6
7	7	2	1	2	23	15.72	1	7
8	8	2	1	3	2	19.69	1	8
9	9	2	1	4	6	17.47	1	9
10	10	2	1	5	15	15.98	1	10
11	11	3	1	1	18	16.3	1	11
12	12	3	1	2	25	16.33	1	12
13	13	3	1	3	9	12.55	1	13
14	14	3	1	4	11	12.77	1	14
15	15	3	1	5	2	15.72	1	15

Figure 8.4

Figure 8.5 shows the Linear Mixed Models menu with the necessary boxes filled in. If we use the Linear Mixed Model Options menu (Figure 8.2) to request predicted means and standard errors of differences of means (in addition to the existing Display options), and then click on Run in the Linear Mixed Models menu itself, the following output is produced.

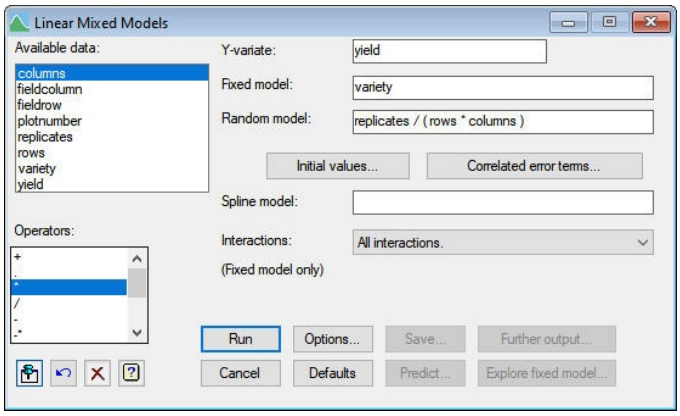


Figure 8.5

REML variance components analysis

Response variate: yield
Fixed model: Constant + variety
Random model: replicates + replicates.rows + replicates.columns + replicates.rows.columns
Number of units: 150

replicates.rows.columns used as residual term

Sparse algorithm with AI optimisation

Estimated variance components

Random term	component	s.e.
replicates	0.4262	0.6890
replicates.rows	1.5595	0.5091
replicates.columns	1.4812	0.4865

Residual variance model

Term	Model(order)	Parameter	Estimate	s.e.
replicates.rows.columns	Identity	Sigma2	0.806	0.1340

Tests for fixed effects

Sequentially adding terms to fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
variety	212.26	24	8.84	79.3	<0.001

Dropping individual terms from full fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
variety	212.26	24	8.84	79.3	<0.001

Message: denominator degrees of freedom for approximate F-tests are calculated using algebraic derivatives ignoring fixed/boundary/singular variance parameters.

Table of predicted means for Constant

14.70 Standard error: 0.422

Table of predicted means for variety

variety	1	2	3	4	5	6	7	8
	12.84	15.49	14.21	14.52	15.33	15.27	14.01	14.57
variety	9	10	11	12	13	14	15	16
	12.99	11.93	13.27	14.84	16.19	13.27	14.98	13.46
variety	17	18	19	20	21	22	23	24
	14.98	15.92	16.70	16.40	14.93	16.44	13.29	15.46
variety	25							
	16.31							

Standard error of differences: 0.6202

Unusually for a large variety trial, this particular design is balanced (in fact it is a lattice square), and we can gain additional insights into the REML analysis by looking at the output that we could have obtained from the Analysis of Variance menu. The menu is not customized for the design, but we can use the General analysis of variance setting in the Design box, and specify the Treatment structure and Block structure as shown in Figure 8.6. The standard analysis of variance output (analysis-of-variance table, information summary, means and standard errors of differences) is shown below.

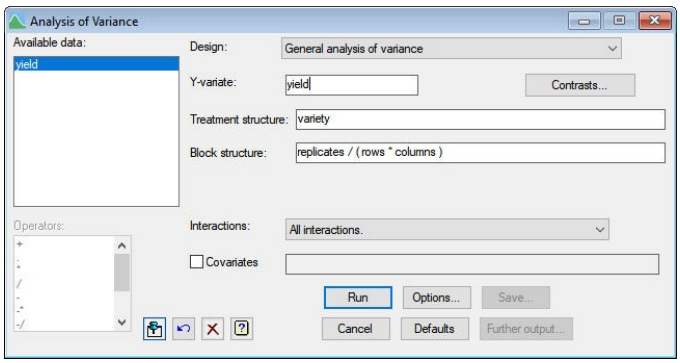


Figure 8.6

Analysis of variance

Variate: yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
replicates stratum	5	133.3273	26.6655		
replicates.rows stratum					
variety	24	215.9053	8.9961		
replicates.columns stratum					
variety	24	229.8094	9.5754		
replicates.rows.columns stratum					
variety	24	166.7675	6.9486	8.58	<.001
Residual	72	58.3011	0.8097		
Total	149	804.1105			

Information summary

Model term	e.f.	non-orthogonal terms
replicates.rows stratum		
variety	0.167	
replicates.columns stratum		
variety	0.167	replicates.rows
replicates.rows.columns stratum		
variety	0.667	replicates.rows
		replicates.columns

Message: the following units have large residuals.

replicates 6	-1.895	approx. s.e.	0.943
replicates 1 rows 4 columns 3	-1.665	approx. s.e.	0.623
replicates 1 rows 5 columns 2	1.710	approx. s.e.	0.623

Tables of means

Variate: yield

Grand mean 14.704

variety	1	2	3	4	5	6	7
	12.962	15.561	14.152	14.560	15.481	15.358	14.008
variety	8	9	10	11	12	13	14
	14.428	12.968	11.928	13.222	14.835	16.176	13.187
variety	15	16	17	18	19	20	21
	15.067	13.287	14.968	15.881	16.742	16.277	15.048
variety	22	23	24	25			
	16.430	13.283	15.464	16.344			

Standard errors of differences of means

Table	variety
rep.	6
d.f.	72
s.e.d.	0.6363

Notice that the analysis-of-variance table has *three* lines for `variety`. As each row contains a different set of varieties, the differences between the rows in each replicate enable us to obtain estimates of the variety effects (which appear in the `replicates.rows` stratum). The same is true of the columns. The design is balanced because the various comparisons between varieties are all estimated with the same efficiency in the `replicates.rows` stratum; the Information Summary indicates the efficiency is in fact 0.167. Similarly, they all have efficiency 0.167 in the `replicates.columns` stratum, and efficiency 0.667 in the `replicates.rows.columns` stratum. So, the possible information on variety is split (1/6 : 1/6 : 2/3) between the three strata.

We can see the estimates obtained in each stratum by checking the Effects box in the ANOVA Further Output menu (Figure 8.7) and then clicking Run, and you can verify that the standard table of means produced by `ANOVA`, above, is calculated using the estimated effects from the lowest stratum (`replicates.rows.columns`): the mean 12.962 for variety 1 is the grand mean 14.704 plus the effect of variety 1 in the `replicates.rows.columns` table, namely −1.742.

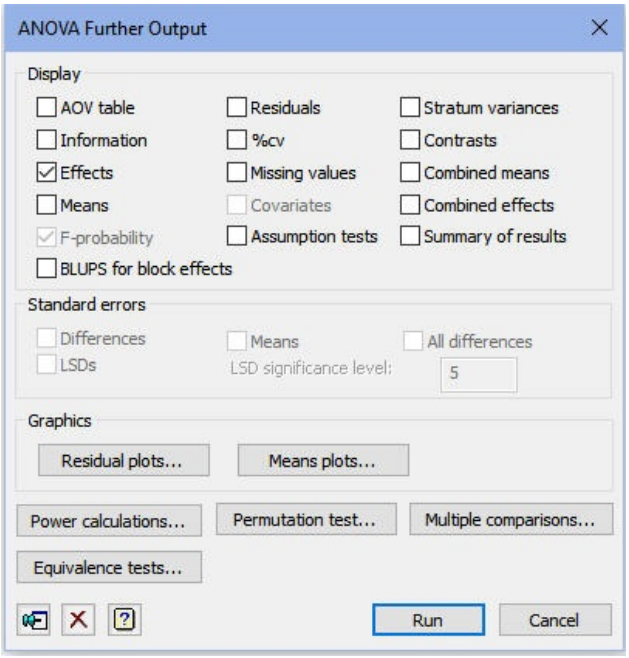


Figure 8.7

Tables of effects

Variate: yield

replicates.rows stratum

variety effects, e.s.e. *, rep. 6

variety	1	2	3	4	5	6	7
	-5.614	1.296	0.604	-1.468	-3.522	2.790	-3.458
variety	8	9	10	11	12	13	14
	1.718	0.520	-3.814	-2.718	-2.544	1.020	1.236
variety	15	16	17	18	19	20	21
	0.582	5.598	3.786	3.480	3.902	3.530	-1.294
variety	22	23	24	25			
	-0.028	1.360	-3.058	-3.894			

replicates.columns stratum

variety effects, e.s.e. *, rep. 6

variety	1	2	3	4	5	6	7
	-3.432	-2.588	0.812	-0.650	-1.450	-4.948	1.930
variety	8	9	10	11	12	13	14
	4.064	-3.010	-1.584	1.852	2.828	2.540	-0.752
variety	15	16	17	18	19	20	21
	-3.536	-0.642	-2.494	0.740	-1.706	4.934	-2.9240
variety	22	23	24	25			
	3.990	-3.730	4.434	5.332			

replicates.rows.columns stratum

variety effects, e.s.e. 0.4499, rep. 6

variety	1	2	3	4	5	6	7
	-1.742	0.857	-0.553	-0.144	0.777	0.653	-0.697
variety	8	9	10	11	12	13	14
	-0.277	-1.736	-2.777	-1.482	0.130	1.471	-1.517
variety	15	16	17	18	19	20	21
	0.362	-1.418	0.263	1.176	2.037	1.573	0.343
variety	22	23	24	25			
	1.726	-1.421	0.760	1.639			

In contrast, the REML analysis has produced a single set of estimates, and these automatically combine (with an appropriate weighting) all the separate estimates. In fact the REML estimates correspond to the *combined* effects and means in the ANOVA Further Output menu. Below, we show these tables, together with the output generated by checking the Stratum variances box which contains the variance components. The combined means have a smaller standard error of difference than the standard means, but the complicated structure of their estimation means that we can no longer assume that differences between them follow t-distributions with a known number of degrees of freedom. (However, the *effective* numbers of degrees of freedom printed by ANOVA are

generally reasonably reliable.)

Tables of combined effects

Variate: yield

variety effects, e.s.e. 0.4385, rep. 6, effective d.f. 79.99

variety	1	2	3	4	5	6	7
	-1.869	0.786	-0.495	-0.186	0.628	0.570	-0.697
variety	8	9	10	11	12	13	14
	-0.131	-1.716	-2.772	-1.432	0.133	1.486	-1.438
variety	15	16	17	18	19	20	21
	0.276	-1.243	0.277	1.217	1.991	1.695	0.230
variety	22	23	24	25			
	1.739	-1.413	0.760	1.602			

Tables of combined means

Variate: yield

variety	1	2	3	4	5	6	7
	12.836	15.490	14.209	14.519	15.333	15.274	14.007
variety	8	9	10	11	12	13	14
	14.574	12.989	11.932	13.272	14.838	16.190	13.266
variety	15	16	17	18	19	20	21
	14.980	13.461	14.982	15.922	16.696	16.399	14.934
variety	22	23	24	25			
	16.444	13.291	15.465	16.306			

Standard errors of differences of combined means

Table	variety
rep.	6
s.e.d.	0.6202
effective d.f.	79.99

Estimated stratum variances

Variate: yield

Stratum	variance	effective d.f.	variance component
replicates	26.6655	5.000	0.4262
replicates.rows	8.6037	23.464	1.5595
replicates.columns	8.2120	23.438	1.4812
replicates.rows.columns	0.8062	73.099	0.8062

The example reinforces the point that the REML output is the same as that given by ANOVA when both are feasible, but that the generality of the REML method leaves aspects that it cannot duplicate. More importantly, though, it shows that the REML method makes use of all the available information about each fixed effect. These aspects indicate the efficiency and appropriateness of the methodology, and the exercises at the end of the chapter will illustrate its ability to handle designs that cannot be analysed by ANOVA. Another important advantage is that REML can fit models to spatial correlation structures. Details are given in the *Guide to the Genstat Command Language*, Part 2, Section 5.4, and the *Guide to REML in Genstat*, Chapters 3 and 4.

8.4 Practical

Genstat spreadsheet file Vartrial1.gsh contains data from a trial of 35 varieties of wheat. The design has two replicates each laid out in a five by seven plot array. Assuming that the same block structure is appropriate as in Section 8.3 (rows crossed with columns within replicates), analyse the data as a linear mixed model.

8.5 Analysis of variance by ANOVA, regression or REML

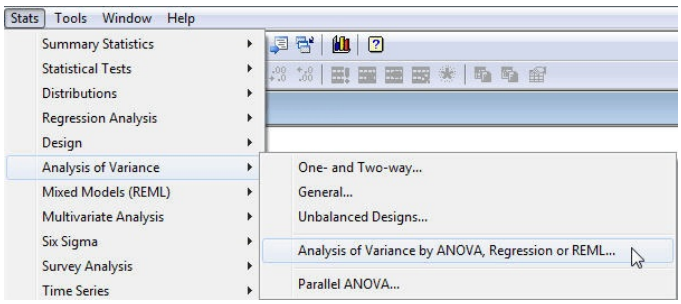


Figure 8.8

In the earlier chapters of this Guide, you have seen that, if your design is balanced you can produce an analysis if variance using the Analysis of Variance menu (Figure 1.8), or you may be able to use the One- and Two-way Analysis of Variance menu (Figure 3.2) if you have no more than two treatment factors. Genstat will tell you if the design is unbalanced. Then, if it has only one error term you can use the Unbalanced ANOVA menu (Figure 7.9), or if it has several you can use the Linear Mixed Models menu (Figure 8.1). A small complication is that you might want to use the Unbalanced ANOVA menu rather than the Linear Mixed Models menu, even when there several error terms, if the additional error terms contain very little information about the treatments (and this was why we did not use the Linear Mixed Models menu in Section 7.6).

So you could define a set of rules to decide how to analyse a complicated design. However, you might prefer Genstat to do this for you – and, in fact, it will do so if you use the menu for Analysis of Variance by ANOVA, Regression or REML. Figure 8.9 shows the use of the menu to analyse the production data from Section 7.6.

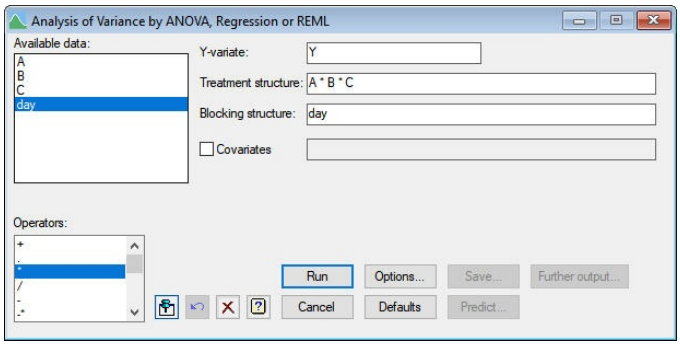


Figure 8.9

The Options menu (Figure 8.10) allows you to select only the types of output that are available from all the possible methods of analysis. You can also say how much information (i.e. efficiency) you are prepared to lose on any treatment term when deciding to use whether to use the Unbalanced ANOVA menu (which uses regression) rather than the Linear Mixed Models menu (which uses REML). The Information section will contain details of the recommended method, and the amount of information that may have been lost.

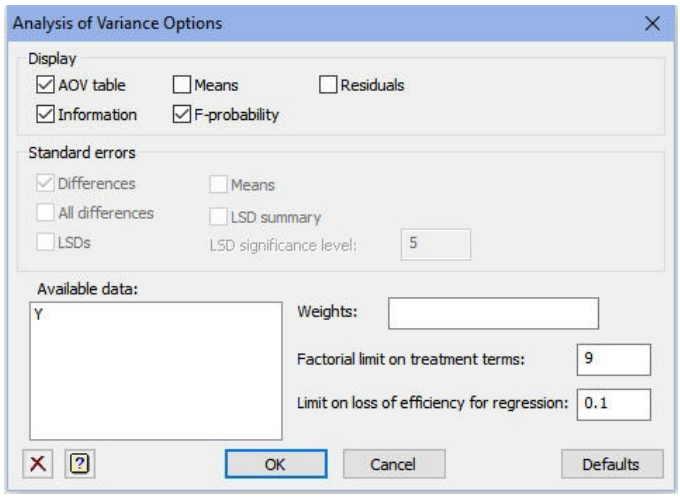


Figure 8.10

The output, below, confirms that it was acceptable to use Unbalanced ANOVA in Section 7.6: less than 1% of the information has been lost.

Analysis of variance by ANOVA, REML or regression

Information summary

Design unbalanced with weights or more than 2 treatment factors, and no more than 0.801% of information on any contrast estimated between block terms; analyse by AUNBALANCED.

Accumulated analysis of variance

Change	d.f.	s.s.	m.s.	v.r.	F pr.
+ day	1	914.0	914.0	3.67	0.061
+ A	2	1706.8	853.4	3.42	0.041
+ B	2	418.8	209.4	0.84	0.438
+ C	1	1065.9	1065.9	4.28	0.044
+ A.B	4	1166.0	291.5	1.17	0.336
+ A.C	2	2456.7	1228.3	4.93	0.011

+ B.C	2	284.4	142.2	0.57	0.569
+ A.B.C	4	1397.4	349.4	1.40	0.248
Residual	48	11960.4	249.2		
Total	66	21370.4	323.8		

8.6 Practical

Re-analyse the data in `Vartrial1.gsh` using the menu for Analysis of Variance by ANOVA, Regression or REML.

9 Commands for analysis of variance

This optional (★) chapter introduces the main commands that are used for analysis of variance in Genstat. The full descriptions, however, are in the *Genstat Reference Manual* (Part 2 for directives, or Part 3 for Procedures) or in the *Guide to the Genstat Command Language*. These can both be accessed on line, from the Help menu on the Genstat menu bar.

Most of the menus described in this course use the ANOVA directive, which analyses *generally balanced* designs. These include most of the commonly occurring experimental designs such as randomized blocks, Latin squares, split plots and other orthogonal designs, as well as designs with balanced confounding, like balanced lattices and balanced incomplete blocks. Many partially balanced designs can also be handled, using pseudo factors, so a very wide range of designs can be analysed.

Before using ANOVA we first need to define the model that is to be fitted in the analysis. Potentially this has three parts. The BLOCKSTRUCTURE directive defines the "underlying structure" of the design or, equivalently, the *error* terms for the analysis; in the simple cases where there is only a single error term this can be omitted. The TREATMENTSTRUCTURE directive specifies the treatment (or *systematic*, or *fixed*) terms for the analysis. The other directive, COVARIATE, lists the covariates if an analysis of covariance is required. Alternatively, the AFCOVARIATES procedure can define covariates from a model formula, for example to fit a different regression coefficient for every level of a factor like blocks; it calculates the variates required to represent the covariates and then specifies them as covariates for the analysis using the COVARIATE directive.

At the start of a job all these model-definition directives have null settings. However, once any one of them has been used, the defined setting remains in force for all subsequent analyses in the same job until it is redefined.

For example, the statements below were generated by the One-way ANOVA (no Blocking) menu to analyse the example in Section 1.5.

```
"One-way ANOVA (no Blocking)."
BLOCK "No Blocking"
TREATMENTS diet
COVARIATE "No Covariate"
ANOVA [PRINT=aovtable,information,mean; FPROB=yes] weight
```

The BLOCK (or, in full, BLOCKSTRUCTURE) directive is given a null setting to cancel any existing setting; so this indicates that the design is unstructured and has a single error term. Similarly, the COVARIATE statement cancels any covariates that may have been set in an earlier menu. The TREATMENTS (or, in full, TREATMENTSTRUCTURE) directive is used to specify that we have a single term in the analysis, the main effect of diet.

The first parameter of the ANOVA directive specifies the y-variate to be analysed. The PRINT option is set to a list of strings to select the output to be printed. These are similar to the check boxes of the Further Output menu. The most commonly used settings are:

aovtable	analysis-of-variance table,
information	details of large residuals, non-orthogonality and any aliasing in the model,
covariates	estimated coefficients and standard errors of any

	covariates,
effects	tables of effects,
residuals	tables of residuals,
contrasts	estimated coefficients of polynomial or other contrasts,
means	tables of means,
%cv	coefficient of variation, and
missingvalues	estimated missing values.

By default PRINT=aovtable,information,covariates,means,missing.

Probabilities are not printed by default for the variance ratios in the analysis-of-variance table, but these can be requested by setting the FPROBABILITY option to yes. ANOVA has a PSE option to control the standard errors printed for tables of means. The default setting is differences, which gives standard errors of differences of means. The setting means produces standard errors of means, LSD produces least significant differences and by setting PSE=* the standard errors can be suppressed altogether. The LSDLEVEL option allows the significance level for the least significant differences to be changed from the default of 5%. ANOVA also has a FACTORIAL option which can be used to specify the maximum order (that is, number of factors) in the treatment terms to be fitted in the analysis; default 3.

To show a more complicated example, these statements were generated to analyse the split-plot design in Section 5.1

```
"Split-Plot Design."
BLOCK blocks/wplots/subplots
TREATMENTS nitrogen*variety
COVARIATE "No Covariate"
ANOVA [PRINT=aovtable,information,mean; FACT=3; FPROB=yes]\
  yield
```

The block formula

```
blocks/wplots/subplots
```

expands, as explained in Section 3.4, to give the three terms

```
block + block.wplot + block.wplot.subplot
```

each of which defines a stratum for the analysis. Similarly, the treatment formula

```
nitrogen*variety
```

expands to

```
nitrogen + variety + nitrogen.variety
```

to request that Genstat fits the main effects of nitrogen and variety, and their interaction. Again there are no covariates.

The Further Output menu uses the ADISPLAY directive to produce the output, procedure APLOT to produce the plots of residuals, procedure AGRAPH to plot tables of means, procedure APERMTEST for permutation tests, and procedure AMCOMPARISON for multiple-comparison tests. ADISPLAY has options PRINT, FPROBABILITY, PSE and LSDLEVEL like those of ANOVA. However, with ADISPLAY the default for PRINT is to print nothing.

The summaries of results are produced by the ARESULTSUMMARY procedure; see part 3 of the *Genstat Reference Manual* for details.

Finally, the AKEEP directive is used by the ANOVA Save Options menu to save the residuals and fitted values after an analysis. This is done by two options called RESIDUALS and FITTEDVALUES. AKEEP also allows information to be saved for any of the individual terms in the analysis. The terms are defined by a formula which is specified using the TERMS parameter. The formula is expanded into a list of model terms, subject to the limit defined by the FACTORIAL option which operates like the FACTORIAL option of ANOVA; the other parameters then specify data structures in parallel with this list, to store the information required. Tables of means are saved using the MEANS parameter. Other useful parameters of AKEEP are EFFECTS (tables of effects for treatment terms), REPLICATIONS (replication tables), RESIDUALS (tables of residuals for block terms), DF (degrees of freedom) and SS (sums of squares).

Below we use AKEEP to save the sum of squares and degrees of freedom for nitrogen and variety from the analysis of the split-plot design in Section 5.1.

```
47  AKEEP nitrogen+variety; SS=N_ss,V_ss; DF=N_df,V_df
48  PRINT N_ss,N_df,V_ss,V_df; DECIMALS=1,0

      N_ss        N_df        V_ss        V_df
   20020.5           3      1786.4           2
```

The One and two-way Analysis of Variance menu uses the A2WAY procedure, which uses the ANOVA directive for balanced designs, and the regression facilities for unbalanced designs. This has a Y parameter that supplies a variate containing the data values to be analysed. The treatment factor or factors are specified by the TREATMENTS option. The FACTORIAL option sets a limit in the number of factors in each treatment term. So you can set FACTORIAL=1 to fit only the main effects when there are two treatment factors; the default FACTORIAL=2 also fits their interaction. The BLOCKS option can supply a blocking factor, for example to define a randomized-block design. There is also a COVARIATES option which can supply one or more variates to be used as covariates in an analysis of covariance.

Printed output from A2WAY is controlled by its PRINT option, with settings aovtable, information, covariates, effects, means, %cv and missingvalues, that operate like those of the ANOVA directive, above.

The PSE option of A2WAY controls the standard errors printed with the tables of means. The default setting is differences, which gives standard errors of differences of means. The setting means produces standard errors of means, lsd produces least significant differences, and by setting PSE=* the standard errors can be suppressed altogether. The significance level to use in the calculation of least significant differences can be changed from the default of 5% using the LSDLEVEL option.

For unbalanced designs, the means are produced for A2WAY by the PREDICT directive. The first step (A) of the calculation forms the full table of predictions, classified by all the treatment and blocking factors. The second step (B) averages the full table of over the factors that do not occur in the table of means. The COMBINATIONS option specifies which cells of the full table are to be formed in Step A. The default setting, estimable, fills in all the cells other than those that involve parameters that cannot be estimated. Alternatively, setting COMBINATIONS=present excludes the cells for factor

combinations that do not occur in the data. The ADJUSTMENT option then defines how the averaging is done in Step B. The default setting, marginal, forms a table of marginal weights for each factor, containing the proportion of observations with each of its levels; the full table of weights is then formed from the product of the marginal tables. The setting equal weights all the combinations equally. Finally, the setting observed uses the WEIGHTS option of PREDICT to weight each factor combination according to its own individual replication in the data.

The PLOT option of A2WAY allows up to four of the following residual plots to be requested:

fittedvalues	for a plot of residuals against fitted values;
normal	for a Normal plot;
halfnormal	for a half-Normal plot;
histogram	for a histogram of residuals; and
absresidual	for a plot of the absolute values of the residuals against the fitted values.

By default the first four are produced. The GRAPHICS option determines the type of graphics that is used, with settings highresolution (the default) and lineprinter.

The RESIDUALS parameter of A2WAY can save the residuals from the analysis, and the FITTEDVALUES parameter can save the fitted values. The SAVE parameter can save a "save" structure that can be used as input to procedure A2DISPLAY to produce further output, or to procedure A2KEEP to copy output into Genstat data structures.

The Unbalanced ANOVA menu uses procedure AUNBALANCED, which uses the Genstat regression facilities. The method of use is similar to that for ANOVA. The treatment terms to be fitted must be specified, before calling the procedure, by the TREATMENTSTRUCTURE directive. Similarly, any covariates must be indicated by the COVARIATE directive. The procedure also takes account of any blocking structure specified by the BLOCKSTRUCTURE directive. However, it cannot produce stratified analyses like those generated by ANOVA, and is able to estimate treatments and covariates only in the "bottom stratum". So, for example, the full analysis can be produced for a randomized block design, where the treatments are all estimated on the plots within blocks, but it cannot produce the whole-plot analysis in a split-plot design. The parameters of AUNBALANCED are identical to those of ANOVA, and there are also FACTORIAL and FPROBABILITY options like those of ANOVA. Printed output is controlled by the PRINT option, with settings: aovtable to print the analysis-of-variance table, effects to print the effects (as estimated by Genstat regression), means to print tables of predicted means with standard errors, residuals to print residuals and fitted values, screen to print "screening" tests for treatment terms, and %cv to print the coefficient of variation. The default is to print the analysis-of-variance table and tables of means.

AUNBALANCED calls procedure RSCREEN to provide the screening tests for the treatment terms: marginal tests to assess the effect of adding each term to the simplest possible model (i.e. a model containing any blocks and covariates, and any terms marginal to the term); conditional tests to assess the effect of adding each term to the fullest possible model (i.e. a model containing all terms other than those to which the term is marginal). For example, if we have

```
BLOCKSTRUCTURE Blocks
```

and

```
TREATMENTSTRUCTURE A + B + A.B
```

the marginal test for A will show the effect of adding A to a model containing only Blocks, while the conditional test will show the effect of adding A to a model containing Blocks and B. (The terms A and B are marginal to A.B.)

Like A2WAY, AUNBALANCED forms tables of means using the PREDICT directive and again has options COMBINATIONS and ADJUSTMENT to control how this is done. The PSE option controls the types of standard errors that are produced to accompany the tables of means, with settings: differences for a summary of the standard errors for differences between pairs of means, alldifferences for standard errors for differences between all pairs of means, lsd for a summary of the least significant differences between pairs of means, alllsd for all the least significant differences between pairs of means, and means for standard errors of the means (relevant for comparing them with zero). The default is differences. The NOMESSAGE option allows various warning messages (produced by the FIT directive) to be suppressed, and the PLOT option allows various residual plots to be requested: fittedvalues for a plot of residuals against fitted values, normal for a Normal plot, halfnormal for a half Normal plot, and histogram for a histogram of residuals.

Procedure AUDISPLAY is used to produce further output for an unbalanced design. It has options PRINT, FPROBABILITY, COMBINATIONS, ADJUSTMENT, PSE and LSDLEVEL like those of AUNBALANCED, except that no screening tests are available.

The menus described in Chapter 8 use the REML directive. Before using REML we first need to define the model that is to be fitted in the analysis. For straightforward linear mixed models, the only directive that needs to be specified is VCOMPONENTS. The FIXED option specifies a model formula defining the fixed model terms to be fitted, and the RANDOM parameter specifies another model formula defining the random terms. There are two other parameters. INITIAL provides initial values for the estimation of each variance component. These are supplied as the ratio of the component to the residual variance, but the default value of one is usually satisfactory. The CONSTRAINT parameter can be used to indicate whether each variance component is to be constrained in any way. The default setting, none, leaves them unconstrained. The positive setting forces a variance component to be kept positive, the fixrelative fixes the relative value of the component to be equal to that specified by the INITIAL parameter, and the fixabsolute setting fixes it to the absolute value specified by INITIAL. The FACTORIAL option sets a limit on the number of factors and variates allowed in each fixed term (default 3); any term containing more than that number is deleted from the model.

Usually, only FIXED and RANDOM need to be set. For example, the statement below defines the models for the split-plot example in Section 7.1.

```
VCOMPONENTS [FIXED=variety*nitrogen] \
  RANDOM=blocks/wplots/subplots
```

Once the models have been defined, the REML directive can be used to perform the analysis. The first parameter of REML specifies the y-variate to be analysed. The PRINT option is set to a list of strings to select the output to be printed. These are similar to the check boxes of the Further Output menu. The most commonly used settings are:

model description of model fitted,

components	estimates of variance components and estimated parameters of covariance models,
effects	estimates of parameters in the fixed and random models,
means	predicted means for factor combinations,
vcovariance	variance-covariance matrix of the estimated components,
deviance	deviance of the fitted model,
waldtests	Wald tests for all fixed terms in model,
missingvalue	estimates of missing values,
covariancemodels	estimated covariance models.

The default setting of PRINT=model,components,Wald,cova gives a description of the model and covariance models that have been fitted, together with estimates of the variance components and the Wald tests. By default if tables of means and effects are requested, tables for all terms in the fixed model are given together with a summary of the standard error of differences between effects/means. Options PTERMS and PSE can be used to change the terms or obtain different types of standard error. For example,

```
REML [PRINT=means; PTERMS=nitrogen.variety; \
     PSE=allestimates]
```

will produce a nitrogen by variety table of predicted means with a standard error for each cell.

Further output is produced by the VDISPLAY directive, which has options PRINT, PTERMS and PSE like those of REML.

Information from the analysis can be saved using the VKEEP directive. For example this has options RESIDUALS and FITTEDVALUES to save the residuals and fitted values respectively. It also has parameters to allow you to save variance components, predicted means, standard errors and so on. Full details are given in Section 5.9 of Part 2 of the *Guide to the Genstat Command Language*.

The Analysis of variance by ANOVA, regression or REML menu uses the AOVANYHOW procedure; see part 3 of the *Genstat Reference Manual* for details.

Index

A2DISPLAY procedure 123
A2WAY procedure 122
Additive model 60
ADISPLAY directive 121
AGRAPH procedure 121
AKEEP directive 19, 122
Alpha design 103
Analysis of covariance 120
Analysis of variance
 menu 12, 65
 one-way 11, 16, 120
 skeleton 80
Analysis of variance by ANOVA, regression or REML menu 118, 125
Analysis of Variance menu 17, 32, 40, 50, 71, 107
Analysis-of-variance table 11, 52, 71
Angular transformation 61
Anova Contrasts menu 17, 40, 42
ANOVA directive 120
ANOVA Further Output menu 13, 18, 22
ANOVA options menu 41
ANOVA Save Options menu 15, 64
AOVANYHOW procedure 125
APLOT procedure 121
ARESULTSUMMARY procedure 121
Assumptions 57
Asterisk
 as crossing operator 46
AUDISPLAY procedure 124
AUNBALANCED procedure 123
Balance
 first-order 92
Balanced design 112, 120
Bias 31
Binomial data 61
Block structure 46
Blocking 26, 30
 in two directions 31
Blocking structure 25
BLOCKSTRUCTURE directive 73, 120
Calculate menu 49
Combined effects 115
Combined means 115
COMPARISON function 48
Completely randomized design 24, 26
Conditional test 98
Confounded interaction 87
Confounded treatment effect 87
Contrast 40, 48
 comparison 40
 function 40
 polynomial 17, 42
Control treatment 48, 83
Correlation model 106
Count data 61
cov. ef. 52
Covariance efficiency factor 52
Covariate 51, 120
COVARIATE directive 120
Crossing operator 46
Cyclic design 103
degrees of freedom 4
 effective 115
Design of experiments 79
Designs analysable by ANOVA 92
Deviations from fitted contrasts 18, 44
Display Data in Output Window menu 64
Dot character
 as operator 46
Effective degrees of freedom 115
Effective standard errors
 for contrasts 42
Effects 9, 10
Efficiency 114
Efficiency factor 88, 90, 92, 114
Equivalence test 22
Error (as residual)
 in analysis of variance 120
Exact test 69
Experimental unit 25
Extracting results
 from analysis of variance 122
Factorial model 38
Factorial plus added control 48, 85
Field experiment 103, 109
First-order balance 92
Fitted values 37, 38
 from analysis of variance 122
Fixed effect 104
Function
 for contrast 48
General Analysis of Variance menu 46
Generally balanced design 120
Generally-balanced design 91
Generate a Standard Design menu 79
Generate a Standard Design Options menu 80
Generating a standard design 79
Grand mean 9
Graphics
 fitted analysis of variance model 121
 model checking 121
Half-Normal plot 59, 60
Higher-order term 46, 47
Histogram of residuals 60
Homogeneous variance 57, 58
Information summary 88, 114

Interaction 37, 38
 between contrasts 44, 45
 in analysis of variance 46, 47
Keeping results
 from analysis of variance 122
Latin square
 with split plots 75
Latin square design 32, 33
Lattice square 112
Least significant difference 12, 121
Level 35
Linear contrast 48
Linear mixed model 104, 109, 117, 118
Linear Mixed Model Options menu 105
Linear Mixed Models Further Output menu 108
Linear Mixed Models menu 104, 111
Linear model 8, 37
 additive 57, 60
 factorial 38
 with covariate 51
Logarithmic transformation 61
Logit transformation 61
Main effect 37
Marginal test 98
Mean square 10
Means Plot menu 14
Missing value 76
Mixed Model 103
Model
 for analysis of variance 120
Model checking 121
Model formula 46, 48
Model term 46
Multiple-comparison tests 20
 enabling 20
Nested structure 48
Nesting operator 47
Non-additivity 60
Non-inferiority test 22
Non-orthogonal 88
Non-superiority 22
Normal distribution 3, 59, 60
Normal plot 59, 60
Normality of residuals 57, 59
Nuisance term 98
One- and two-way Analysis of variance menu 12, 15, 28, 29, 36, 95, 102
One-way analysis of variance 120
Orthogonal 88, 94
Orthogonal polynomial 18, 19
Outlier 57, 60
Partially balanced design 120
Percentage data 61
Permutation test 68
POL function 48
Polynomial contrast 17, 48
Power
 in analysis of variance 83
Power for Design menu 83
Predicted mean 95
Predicted value
 from analysis of variance 122
Probability
 for F-statistic 121
Proportions 61
Pseudo-factor 93
Quadratic contrast 48
Quantile 59
Random effect 104
Randomization 30, 79
Randomized-block design 26, 33, 36, 60, 80
 generating 30
REG function 48
REML 103
 compared to ANOVA 117
REML directive 124
Replication 4
 required in a design 80
Replications Required menu 81
Residual 8
 from analysis of variance 122
Residual plot 58, 60
 half-Normal 59
 Normal plot 59
RSCREEN procedure 123
Save ANOVA Results in a Spreadsheet File menu 15
Saving
 analysis of variance results 122
Screening test 98
Skeleton analysis of variance 80
Slash symbol 47
Spline 105
Split-plot design 48, 71, 104, 121
Split-split-plot design 74
Stabilize the variance 61
Standard error
 of difference of means 4, 12, 51, 73, 121
 of mean 121
Standard errors 90
 for contrasts 42
Standard errors of differences between means 90
Stats menu 79, 85
Storage
 of results from analysis of variance 122
Stratum 29, 71, 74, 114, 121
Stratum variance 115
Sum of squares 9
 corrected for the grand mean 11
 due to the treatments 10
Summaries of results 55, 121
Summarize Contents of Variates menu 6
t-test
 by hand 4

two-sample 3
T-Tests menu 6
Table of means 108
back-transformed 64
in ANOVA 90
Transformation 61
logarithm 49
logarithmic 61
logit 61
Treatment
confounded 87
types of 35
Treatment factor 35
Treatment structure 46, 50
factorial 46
nested 47
Treatment term 120
TREATMENTSTRUCTURE directive 120
Types of treatment 35
Unbalanced ANOVA menu 95, 97, 102
Unbalanced design 93, 95, 97, 117, 122
Variance component 106, 115, 124
Variance ratio 10
VCOMPONENTS directive 124
VDISPLAY directive 125
Wald statistic 106
Weighted analysis 58
Whole-plot 106