DESIGN AND MIXED-MODEL ANALYSIS OF EXPERIMENTS

IX. Factorial experiments

So far in discussing experiments we have considered there to be a set of treatments to be applied and we have been concerned to group the plots and apply these treatments so as to make the experiment as sensitive as possible.

Now I will turn to more sophisticated selection of treatments. The first point is that there will often be more than one factor of interest to the experimenter.

Definition IX.1: Experiments that involve more than one randomized or treatment factor are called **factorial experiments**. ■

IX.A Design of factorial experiments

Example IX.1 Fertilizing oranges

Suppose an experimenter is interested in investigating the effect of nitrogen and phosphorus fertilizer on yield of oranges. It was decided to investigate 3 levels of Nitrogen (viz 0,30,60 kg/ha) and 2 levels of Phosphorus (viz. 0,20 kg/ha). The yield after six months was measured.

For a factorial experiment, the treatments are all possible combinations of the 3 Nitrogen $\times$ 2 Phosphorus levels: $3 \times 2 = 6$ treatments. The treatment combinations are:

Treatment	N	P
1	0	0
2	0	20
3	30	0
4	30	20
5	60	0
6	60	20

The experiment could be laid out as a Completely Randomized Design, a Randomized Complete Block Design or a Latin Square with 6 treatments, but not a BIBD or a Youden Square.

The layout for factorial experiments can be obtained using the *Stats > Design > Select Design* command in Genstat. In answer to the question *Which type of design would you like to generate*, you need to choose a) *orthogonal hierarchical designs (randomized blocks, split-plots)* to generate a factorial design in a CRD or an RCBD or, b) to choose *Latin squares (also Graeco-Latin squares etc as feasible)* to generate it in a Latin Square. You will then be asked a series of questions which you would answer as described in the chapter on that design, except that for the number of treatments factors. If your factorial experiment is to investigate *k* factors, you specify the number of treatment factors to be a) for a CRD, $k + 1$ because you need an extra dummy factor whose number of levels is equal to the replication that you are planning for each treatment combination, b) for an RCBD, *k*, and c) for a Latin square, 1 with its number of levels equal to the number of treatment combinations.

Example IX.1 Fertilizing oranges (continued)

A layout for this experiment in a CRD is generated in Genstat as shown in the following output. It involved the following steps:

1. specifying in the *orthogonal hierarchical designs (randomized blocks, split-plots)* option of the *Stats > Design > Select Design* command:
   - one block factor Seedling;
   - three treatment factors N, P and Dummy with 3, 2 and 3 levels, respectively;

- a seed of, say, 413226;

2. loading the factors Seedling, N and P into a Genstat spreadsheet using the *Spread > New > Data in Genstat* menu command;
3. using *Spread > Factor > Edit Levels* to change the levels of N and P to those to be used in the experiment;
4. using the PDESIGN procedure to print the design.

```
Genstat 5  Release 4.1  (PC/Windows NT)              03 April 2000 11:11:29
Copyright 1998, Lawes Agricultural Trust (Rothamsted Experimental Station)

                ____________________________________________________

                Genstat 5 Fourth Edition - (for Windows)
                Genstat 5 Procedure Library Release PL11
                ____________________________________________________

  3  DESIGN
  4  %WSPREAD Seedling,N,P
  5  DELETE [redefine=yes] N,P
  6  FACTOR [modify=yes;nvalues=18;levels=!(0,30,60)] N
  7  READ N; frepresentation=ordinal

   Identifier    Values   Missing    Levels
            N        18         0         3

  9  FACTOR [modify=yes;nvalues=18;levels=!(0,20)] P
 10  READ P; frepresentation=ordinal

   Identifier    Values   Missing    Levels
            P        18         0         2

 12
 13  PDESIGN [BLOCK=Seedling; TREAT=N,P]

*** Treatment combinations on each unit of the design ***


Seedling
       1    0 20
       2   30  0
       3   60  0
       4   30 20
       5   60  0
       6    0  0
       7    0  0
       8   30 20
       9   60 20
      10   60  0
      11   60 20
      12   30 20
      13    0 20
      14    0  0
      15   30  0
      16   60 20
      17   30  0
      18    0 20

Treatment factors are listed in the order: N P
```

Suppose we decide on a RCBD with three blocks — how many plots per block would be required? **Answer** 6. Again Genstat can be used to obtain the layout as shown in the following output. It involved the following steps:

1. specifying in the *orthogonal hierarchical designs (randomized blocks, split-plots)* option of the *Stats > Design > Select Design* command:

- two block factors, Blocks and Seedling;
- two treatment factors, N and P with 3 and 2 levels, respectively, associated with the Seedling stratum;
- a seed of, say, 413226;

2. loading all 4 factors into a Genstat spreadsheet using the *Spread > New > Data in Genstat* menu command;
3. using *Spread > Factor > Edit Levels* to change the levels of N and P to those to be used in the experiment;
4. using the PDESIGN procedure to print the design.

```
Genstat 5  Release 4.1  (PC/Windows NT)              03 April 2000 11:46:31
Copyright 1998, Lawes Agricultural Trust (Rothamsted Experimental Station)

                ____________________________________________________

                Genstat 5 Fourth Edition - (for Windows)
                Genstat 5 Procedure Library Release PL11
                ____________________________________________________

  3  DESIGN
  4  %WSPREAD Blocks,Seedling,N,P
  5  DELETE [redefine=yes] N,P
  6  FACTOR [modify=yes;nvalues=18;levels=!(0,30,60)] N
  7  READ N; frepresentation=ordinal

   Identifier    Values   Missing    Levels
            N        18         0         3

  9  FACTOR [modify=yes;nvalues=18;levels=!(0,20)] P
 10  READ P; frepresentation=ordinal

   Identifier    Values   Missing    Levels
            P        18         0         2

 12
 13  PDESIGN [BLOCK=Blocks/Seedling; TREAT=N,P]

*** Treatment combinations on each unit of the design ***


Seedling    1       2       3       4       5       6
  Blocks
       1   60 20   60  0    0  0    0 20   30 20   30  0
       2    0 20   60 20    0  0   30  0   30 20   60  0
       3   60  0   30 20   60 20    0 20    0  0   30  0

Treatment factors are listed in the order: N P
```

Finally, the experiment could be laid out in a 6×6 Latin square. Again, Genstat can be used to obtain the layout as shown in the following output. It involved the following steps:

1. specifying in the *Latin squares (also Graeco-Latin squares etc as feasible)* option of the *Stats > Design > Select Design* command:
   - 6 treatments;
   - a seed of, say, 413226;
2. using the *NEWLEVELS* function in *CALCULATE* to specify which of the 6 treatments corresponds to which N-P combination;
3. loading the factors Rows, and Columns and the variates N and P into a Genstat spreadsheet using the *Spread > New > Data in Genstat* menu command;

4. using the *Spread > Column > Convert* command to convert the variates to factors;
5. using PDESIGN used to print the design.

```
Genstat 5  Release 4.1  (PC/Windows NT)              03 April 2000 13:24:45
Copyright 1998, Lawes Agricultural Trust (Rothamsted Experimental Station)

                ____________________________________________

                Genstat 5 Fourth Edition - (for Windows)
                Genstat 5 Procedure Library Release PL11
                ____________________________________________

  3  DESIGN
  4  CALC N=Newlevels(Treats; !v(0,60,120,0,60,120))
  5  CALC P=Newlevels(Treats; !v(0,0,0,20,20,20))
  6  %WSPREAD Rows,Columns,N,P
  7  DELETE [redefine=yes] N,P
  8  FACTOR [modify=yes;nvalues=36;levels=!(0,60,120)] N
  9  READ N; frepresentation=ordinal

   Identifier    Values   Missing    Levels
            N        36         0         3

 11  FACTOR [modify=yes;nvalues=36;levels=!(0,20)] P
 12  READ P; frepresentation=ordinal

   Identifier    Values   Missing    Levels
            P        36         0         2

 14
 15  PDESIGN [BLOCK=Rows*Columns; TREAT=N,P]

*** Treatment combinations on each unit of the design ***


Columns     1        2        3        4        5        6
   Rows
      1   120 20     0  0     0 20   120  0    60 20    60  0
      2     0 20    60  0    60 20     0  0   120 20   120  0
      3    60  0   120 20   120  0    60 20     0  0     0 20
      4   120  0     0 20     0  0   120 20    60  0    60 20
      5    60 20   120  0   120 20    60  0     0 20     0  0
      6     0  0    60 20    60  0     0 20   120  0   120 20

Treatment factors are listed in the order: N P
```

In factorial experiments we are not limited to two factors — thus we may have looked at Potassium at 2 levels as well. How many treatments in this case? Answer $3\times2\times2 = 12$.

IX.B Advantages of factorial experiments

(Mead & Curnow sec. 6.2, 6.3 and 14.6)

a) Interaction in factorial experiments

The major advantage of factorial experiments is that they allow the detection of interaction.

Definition IX.2: Two factors are said to **interact** if the effect of one, on the response variable, depends upon the level of the other. If they do not interact, they are said to be **independent**. ■

To investigate whether two factors interact, the simple effects are computed.

Definition IX.3: A **simple effect** for the means computed for each combination of at least two factors is the difference between two of these means having different levels of one of the factors but the same levels for all other factors. ■

We talk of the simple effects of a factor for the levels of the other factors.

Example IX.2 Chemical reactor experiment

Consider a factorial experiment to investigate the effect of catalyst and temperature on the yield of chemical from a chemical reactor. Suppose there were two levels of each of the factors and that the table of means from the experiment was as follows:

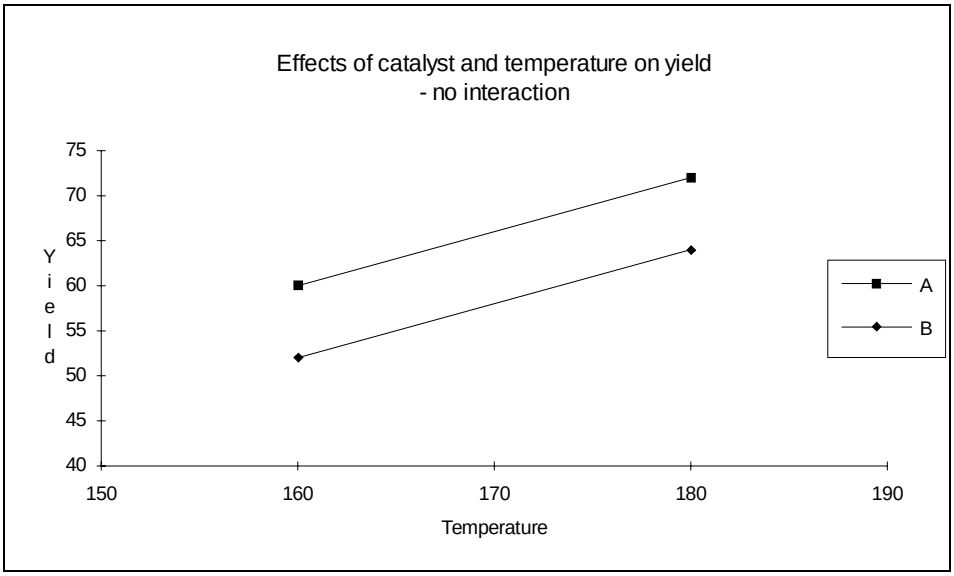
		Temperature (°C)	
		160	180
Catalyst	A	60	72
	B	52	64

For A the temperature effect is 72–60 = 12
For B the temperature effect is 64–52 = 12

These are called the simple effects of temperature.

Clearly, the difference between (effect of) the temperatures is **independent** of which catalyst is used.

The situation can be illustrated using an **interaction plot**, in which the means for the combinations of the factors are plotted.



A set of parallel lines indicates that there is no interaction — the slope of these lines is proportional to the simple effects which are equal. If there was an interaction the simple effects would not be equal and so the lines would not be parallel.

Note that the statement about the independence of two factors is symmetrical in the two factors. Thus,

the simple catalyst effect at 160°C is 52–60 = –8
the simple catalyst effect at 180°C is 64–72 = –8

Thus the difference between (effect of) the catalysts is **independent** of which temperature is used.

So we can say that temperature and catalyst are independent in their effects on yield. In reality, we need to qualify this to say they are *additively* independent (the ratios are not the same).

The practical import of this statement is that we can consider each factor separately. Indeed looking at the overall means for a factor will indicate what is happening in the experiment. For this experiment, the overall means are:

	Temperature (°C)	
	160	180
Mean	56	68

	Catalyst	
	A	B
Mean	66	58

Definition IX.4: A **main effect** of a factor is the difference between two means with different levels of that factor, each mean having been formed from all observations having the same level of the factor. ■

So the differences between the means in these tables are the main effects of the factors. That is, the main effect of Temperature is 12 and that of Catalyst is –8. Note that in this case the main effects are the same as the individual difference calculated above and so they summarize the effect of the factors. Having used the two-way table of means to work out that there is an interaction, it can be abandoned for the purposes of summarizing the results of the analysis.

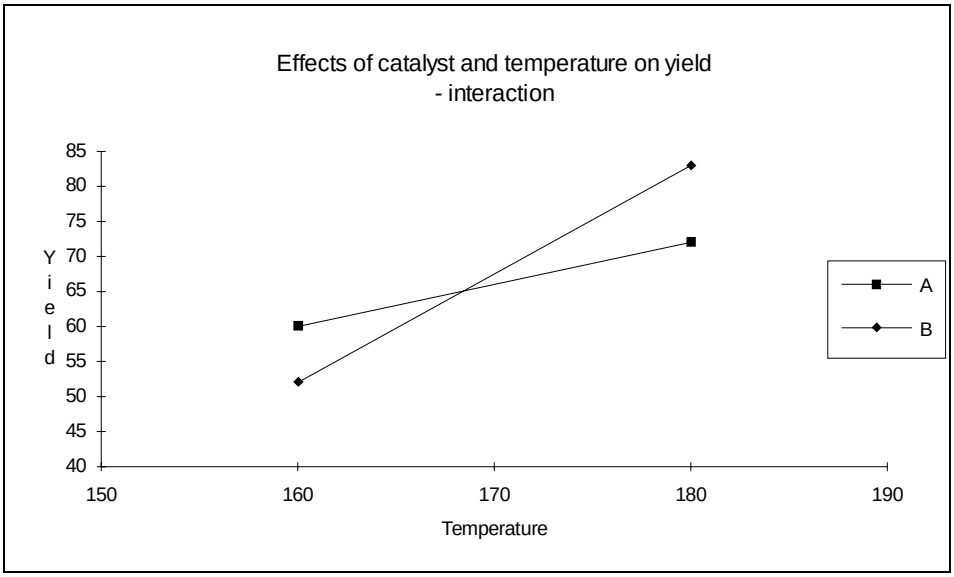
Example IX.3 Second chemical reactor experiment

Suppose the experiment was run with a second reactor and the results were as follows:

		Temperature (°C)	
		160	180
Catalyst	A	60	72
	B	52	83

The simple temperature effect for A is 72–60 = 12
The simple temperature effect for B is 83–52 = 31

Thus the difference between (effect of) the temperatures depends on which catalyst is used. Again, this statement is symmetrical in the factors and so we say that the two factors **interact**. The interaction plot for this example is shown in the following figure.



There is clearly an interaction as the lines have different slopes.

Interactions can be of two types as given in the following definitions.

Definition IX.5: An interaction is a **difference in magnitude** interaction if the direction of the difference is always the same but the size of the difference is not. ■

Definition IX.6: An interaction is a **difference in direction** interaction if the differences between levels of one factor show an increase sometimes and a decrease on others. ■

Also, we can measure the interaction using interaction effects.

Definition IX.7: An **interaction effect** is half the difference of two simple effects for two different levels of just one factor or is half the difference of two interaction effects. ■

Example IX.3 Second chemical reactor experiment (continued)

In this experiment 180°C always yielded more chemical than 160°C, but the size of the increase depended on the catalyst. The interaction is a difference in magnitude interaction. For the interaction to be a difference in direction interaction, the mean for 180°C would have to have been greater than that for 160°C for one catalyst and less for the other catalyst.

In an experiment in which two factors interact, it is not appropriate to summarize the results using the overall means. In the case of the example, the overall means are:

	Temperature (°C)	
	160	180
Mean	56	77.5

	Catalyst	
	A	B
Mean	66	67.5

The main effects cannot be equal to the simple effects in this case because the simple effects differ for the levels of the other factor. The main effects have no practical interpretation. It is the means for the combinations of the factors that must be examined to reveal the behaviour of the experimental material:

		Temperature (°C)	
		160	180
Catalyst	A	60	72
	B	52	83

Note that the interaction effect is computed as the half-difference between the catalyst effects at each temperature or vice-a-versa:

$$[(72–60) – (83–52)]/2 = [12 – 31]/2 = –9.5$$

or

$$[(52–60) – (83–72)]/2 = [–8 – 9]/2 = –9.5$$

Clearly, the interaction effect will be zero when there is no interaction. The situation in which the two factors do not interact is the simpler of the two possibilities.

b) Advantages over one-factor-at-a-time experiments

It is sometimes suggested that rather than complicate things by putting several factors into a single experiment, it would be better to run several experiments each of which investigate one of the factors of interest. However, this is not the case as you will be unable to determine whether or not there is an interaction. Take our temperature-catalyst experiment again. One could run an experiment to examine just the temperature effect and would have to keep the catalyst used constant during the experiment. Suppose it is decided to use catalyst B. The implication of this is that the results of this experiment apply only to the conditions under which the experiment was run. That is, when catalyst B is used. To look at the other factor, a second experiment with 2 different catalysts would be run, but the same temperature used throughout say 160°C. WELL YOU HAVE ONLY APPLIED THREE OF THE FOUR POSSIBLE COMBINATIONS OF THE TWO FACTORS — catalyst A at 180°C has not been tested but catalyst B at 160°C has been tested twice as indicated in the following table:

			Experiment 1 Temperature (°C)	
			160	180
Experiment 2	Catalyst	A	60	?
		B	52	83

The results of the experiments would indicate that temperature increases yield by 31 gms and that the catalysts differ by 8 gms in yield. However, these conclusions are restricted — if we presume the factors act additively we would predict the yield for catalyst A at 160°C to be 60+31 = 83 + 8 = 91. This is quite clearly erroneous — we need the factorial experiment to determine if there is an interaction; exactly the same total amount of resources are involved in the two alternative strategies, assuming the number of replicates is the same in all the experiments.

In addition to allowing the detection of interaction, factorial experiment also have the advantage that, if the factors are additive then the main effects are estimated with greater precision. In the one-factor-at-a time experiments the effect of a particular factor is estimated as the difference between two means each based on *r* observations where *r* is the number of replicates of each treatments. In the factorial experiment the main effects of the factors are the difference between two means based on 2r observations which represents a $\sqrt{2}$ increase in precision. The improvement in precision will be greater for more factors and more levels; for a $3\times3\times3 = 3^3$ experiment the main effects are based on 9r observations leading to a 3 fold increase in precision over one-factor-at-a-time experiments.

To summarize, relative to one-factor-at-a-time experiments, factorial experiments have the advantages that:

1. if the factors interact, factorial experiments allow this to be detected and estimates of the interaction effect can be obtained, and

2. if the factors are independent, factorial experiments result in the estimation of the main effects with greater precision.

IX.C Models for factorial experiments and the estimation of their parameters

The models for the factorial experiments will depend on the design used in assigning the treatments — that is, CRD, RCBD or LS. The design will determine the unrandomized factors and the terms to be included involving those factors. They will also depend on the number of randomized factors.

a) Models for two-factor CRD experiments

We will consider the models for the two-factor CRD. Let the total number of observations be n and the factors be A and B with a and b levels, respectively. Suppose that the combinations of A and B are each replicated r times — that is, $n = a \times b \times r$.

The maximal model for a two-factor CRD experiment, where the two randomized factors A and B are fixed, is:

$$\psi = E[\mathbf{Y}] = \mathbf{X}_{AB}(\alpha\beta) \text{ and } \mathbf{V} = \sigma_U^2 \mathbf{I}_n,$$

where $\mathbf{Y}$ is the n-random vector of random variables representing the response variable,

$(\alpha\beta)$ is the ab-vector of parameters for the A-B combinations,

$\mathbf{X}_{AB}$ is the $n \times ab$ matrix giving the combinations of A and B that occurred on each unit, and

σ_U^2 is the variability arising from different units.

Our model also involves assuming $\mathbf{Y} \sim N(\psi, \mathbf{V})$.

The alternative models for the expectation that are usually considered are:

$$E[\mathbf{Y}] = \mathbf{X}_G \mu \qquad (\text{no factors affect response})$$
$$E[\mathbf{Y}] = \mathbf{X}_A \alpha \qquad (\text{A only affects response})$$
$$E[\mathbf{Y}] = \mathbf{X}_B \beta \qquad (\text{B only affects response})$$
$$E[\mathbf{Y}] = \mathbf{X}_A \alpha + \mathbf{X}_B \beta \qquad (\text{A and B independently affect response})$$
$$E[\mathbf{Y}] = \mathbf{X}_{AB}(\alpha\beta) \qquad (\text{A and B interact in effect on response})$$

We can give expressions for the **X** matrices in terms of direct products of **I**s and **1**s. To do this requires that the elements of the **Y** vector be ordered so that the values of the factors A and B are in standard order.

Definition IX.8: Suppose we have the factors A, B, …, Z with *a*, *b*, …, *z* levels, respectively, each of which has $n = ab\ldots z$ values. The values of the factors are in **standard order** when

1. for any factor, the values of a factor consist of repetitions of the sequence of its levels that begins with the first level and goes to the last level, with the factor A consisting of just a single sequence; and
2. the number of consecutive values with the same level of a factor is the product of the numbers of levels of all the factors to its right, with factor Z changing with every value. ■

That is the values of the factors are systematically ordered in a hierarchical fashion — they are ordered according to A, then B, then C, … and then Z.

Suppose, the elements of the **Y** vector for our two-factor CRD are arranged so that the values of the factors A, B and a dummy factor, say Reps, are in standard order. Then the **X** matrices can be written as the following direct products:

$$\mathbf{X}_{\mathrm{G}} = \mathbf{1}_a \otimes \mathbf{1}_b \otimes \mathbf{1}_r = \mathbf{1}_{abr},\ \mathbf{X}_{\mathrm{A}} = \mathbf{I}_a \otimes \mathbf{1}_b \otimes \mathbf{1}_r,\ \mathbf{X}_{\mathrm{B}} = \mathbf{1}_a \otimes \mathbf{I}_b \otimes \mathbf{1}_r \text{ and } \mathbf{X}_{\mathrm{AB}} = \mathbf{I}_a \otimes \mathbf{I}_b \otimes \mathbf{1}_r$$

The thing about these matrices is that $\mathbf{X}_{\mathrm{G}}$ can be written as a linear combination of the columns of each of the other three and that $\mathbf{X}_{\mathrm{A}}$ and $\mathbf{X}_{\mathrm{B}}$ can be written as linear combinations of the columns of $\mathbf{X}_{\mathrm{AB}}$. Consequently the expectation model $\psi = \mathbf{X}_{\mathrm{G}}\mu$ is marginal to $\psi = \mathbf{X}_{\mathrm{A}}\alpha$, $\psi = \mathbf{X}_{\mathrm{B}}\beta$ and $\psi = \mathbf{X}_{\mathrm{AB}}(\alpha\beta)$ and the models $\psi = \mathbf{X}_{\mathrm{A}}\alpha$ and $\psi = \mathbf{X}_{\mathrm{B}}\beta$ are marginal to $\psi = \mathbf{X}_{\mathrm{AB}}(\alpha\beta)$ or more loosely, that G is marginal to A, B and A.B and that A and B are marginal to A.B.

Note that this set of expectation models is different from any other set that has been presented so far. The maximal model is not just the sum of all terms under consideration. Consequently, choosing between the models is not just a matter of deciding for each term in the maximal model whether or not it can be deleted.

Example IX.4 2×2 Factorial experiment

Suppose A and B have two levels each and that each combination of A and B is replicated 3 times. Hence, $a = b = 2$, $r = 3$ and $n = 12$. Then

$$\alpha' = (\alpha_1 \quad \alpha_2),\ \beta' = (\beta_1 \quad \beta_2) \text{ and } (\alpha\beta)' = (\alpha\beta_{11} \quad \alpha\beta_{12} \quad \alpha\beta_{21} \quad \alpha\beta_{22})$$

Suppose **Y** is arranged so that the values of A, B and Reps are in standard order — that is

$$\mathbf{Y}' = (Y_{11} \quad Y_{11} \quad Y_{11} \quad Y_{12} \quad Y_{12} \quad Y_{12} \quad Y_{21} \quad Y_{21} \quad Y_{21} \quad Y_{22} \quad Y_{22} \quad Y_{22}).$$

Then,

$$\mathbf{X}_G = \mathbf{1}_2 \otimes \mathbf{1}_2 \otimes \mathbf{1}_3 = \mathbf{1}_{12},\ \mathbf{X}_A = \mathbf{I}_2 \otimes \mathbf{1}_2 \otimes \mathbf{1}_3,\ \mathbf{X}_B = \mathbf{1}_2 \otimes \mathbf{I}_2 \otimes \mathbf{1}_3 \text{ and } \mathbf{X}_{AB} = \mathbf{I}_2 \otimes \mathbf{I}_2 \otimes \mathbf{1}_3$$

so that the **X** matrices are as follows:

$$\begin{array}{cccc} \mathbf{X}_G & \mathbf{X}_A & \mathbf{X}_B & \mathbf{X}_{AB} \\ & \text{A} \quad 1 \quad 2 & & 1 \quad 1 \quad 2 \quad 2 \\ & \text{B} & 1 \quad 2 & 1 \quad 2 \quad 1 \quad 2 \\ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} & \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} & \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

Notice that, as previously suggested, $\mathbf{X}_G$ can be written as a linear combination of the columns of each of the other three and that $\mathbf{X}_A$ and $\mathbf{X}_B$ can be written as linear combinations of the columns of $\mathbf{X}_{AB}$.

Also note that, for the maximal model,

$$\psi = E[\mathbf{Y}] = \mathbf{X}_{AB}(\alpha\beta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} (\alpha\beta)_{11} \\ (\alpha\beta)_{12} \\ (\alpha\beta)_{21} \\ (\alpha\beta)_{22} \end{bmatrix} = \begin{bmatrix} (\alpha\beta)_{11} \\ (\alpha\beta)_{11} \\ (\alpha\beta)_{11} \\ (\alpha\beta)_{12} \\ (\alpha\beta)_{12} \\ (\alpha\beta)_{12} \\ (\alpha\beta)_{21} \\ (\alpha\beta)_{21} \\ (\alpha\beta)_{21} \\ (\alpha\beta)_{22} \\ (\alpha\beta)_{22} \\ (\alpha\beta)_{22} \end{bmatrix}$$

That is, the maximal model allows for a different response for each combination of A and B.

b) Estimation of the model parameters for two-factor CRD experiments

In the following theorems, which are given without proof, we give the estimators of the fitted values for the models considered for the two-factor CRD.

Theorem IX.1: Let **Y** be a *n*-vector of jointly-distributed random variables with

$$\psi = E[\mathbf{Y}] = \mathbf{X}_G\mu \text{ and } \text{var}[\mathbf{Y}] = \sigma_U^2\mathbf{I}$$

where μ is the overall mean response and $\mathbf{X}_G$ is the $n\times1$ matrix of ones.

Then $\hat{\psi} = \mathbf{G}$ where **G** is the *n*-vector containing just the grand mean.

Proof: left as an exercise ■

Theorem IX.2: Let **Y** be a *n*-vector of jointly-distributed random variables with

$$\psi = E[\mathbf{Y}] = \mathbf{X}_A\alpha \text{ and } \text{var}[\mathbf{Y}] = \sigma_U^2\mathbf{I}$$

where α is the *a*-vector of parameters specifying a different mean response for each level of the factor A, and
$\mathbf{X}_A$ is the $n\times a$ matrix indicating the levels of A for the elements of **Y**,
σ_U^2 is the variability arising from different units.

Then $\hat{\psi} = \mathbf{A}$ where **A** is the *n*-vector of A means.

Proof: left as an exercise ■

Theorem IX.3: Let **Y** be a *n*-vector of jointly-distributed random variables with

$$\psi = E[\mathbf{Y}] = \mathbf{X}_B\beta \text{ and } \text{var}[\mathbf{Y}] = \sigma_U^2\mathbf{I}$$

where β is the *b*-vector of parameters specifying a different mean response for each level of the factor B,
$\mathbf{X}_B$ is the $n\times b$ matrix indicating the levels of B for the elements of **Y**,
σ_U^2 is the variability arising from different units.

Then $\hat{\psi} = \mathbf{B}$ where **B** is the *n*-vector of B means.

Proof: left as an exercise ■

Theorem IX.4: Let **Y** be a *n*-vector of jointly-distributed random variables with

$$\psi = E[\mathbf{Y}] = \mathbf{X}_A\alpha + \mathbf{X}_B\beta \text{ and } \text{var}[\mathbf{Y}] = \sigma_U^2\mathbf{I}$$

where α is the *a*-vector of parameters specifying a different mean response for each level of the factor A,
$\mathbf{X}_A$ is the $n \times a$ matrix indicating the levels of A for the elements of **Y**,
β is the *b*-vector of parameters specifying a different mean response for each level of the factor B,
$\mathbf{X}_B$ is the $n \times b$ matrix indicating the levels of B for the elements of **Y**,
σ_U^2 is the variability arising from different units.

Then $\hat{\psi} = \mathbf{A} + \mathbf{B} - \mathbf{G}$ where **A**, **B** and **G** are the *n*-vectors of A, B and grand means, respectively.

Proof: left as an exercise ■

Theorem IX.5: Let **Y** be a *n*-vector of jointly-distributed random variables with

$$\psi = E[\mathbf{Y}] = \mathbf{X}_{AB}(\alpha\beta) \text{ and } \operatorname{var}[\mathbf{Y}] = \sigma_U^2 \mathbf{I},$$

where $(\alpha\beta)$ is the *ab*-vector of parameters specifying a different mean response for each A-B combinations,
$\mathbf{X}_{AB}$ is the $n \times ab$ matrix giving the combinations of A and B for each element of **Y**,
σ_U^2 is the variability arising from different units.

Then $\hat{\psi} = \mathbf{A}.\mathbf{B}$ where **A.B** is the *n*-vector of means for the combinations of A and B, respectively.

Proof: left as an exercise ■

Note that all the ψ estimators are functions of means and so can be written in terms of mean operators, **P**s. Further, if **Y** is arranged so that the associated factors A, B and the dummy factor Reps are in standard order, the **P** operators can be written as the direct product of **I** and **J** matrices. These expressions are summarized in the following table.

Model	Estimator
$\psi = \mathbf{X}_G \mu$	$\hat{\psi} = \mathbf{G} = \mathbf{P}_G \mathbf{Y} = n^{-1} \mathbf{J}_a \otimes \mathbf{J}_b \otimes \mathbf{J}_r \mathbf{Y}$
$\psi = \mathbf{X}_A \alpha$	$\hat{\psi} = \mathbf{A} = \mathbf{P}_A \mathbf{Y} = (br)^{-1} \mathbf{I}_a \otimes \mathbf{J}_b \otimes \mathbf{J}_r \mathbf{Y}$
$\psi = \mathbf{X}_B \beta$	$\hat{\psi} = \mathbf{B} = \mathbf{P}_B \mathbf{Y} = (ar)^{-1} \mathbf{J}_a \otimes \mathbf{I}_b \otimes \mathbf{J}_r \mathbf{Y}$
$\psi = \mathbf{X}_A \alpha + \mathbf{X}_B \beta$	$\hat{\psi} = \mathbf{A} + \mathbf{B} - \mathbf{G}$
$\psi = \mathbf{X}_{AB}(\alpha\beta)$	$\hat{\psi} = \mathbf{A}.\mathbf{B} = \mathbf{P}_{AB} \mathbf{Y} = r^{-1} \mathbf{I}_a \otimes \mathbf{I}_b \otimes \mathbf{J}_r \mathbf{Y}$

That is, $\mathbf{P}_G$, $\mathbf{P}_A$ and $\mathbf{P}_B$ are equal to the direct product of $\mathbf{P}_G$, $\mathbf{P}_B$ and $\mathbf{P}_T$, from the RCBD, with $\mathbf{J}_r$.

IX.D Hypothesis testing using the ANOVA method for factorial experiments

Like the models, the analysis of a factorial experiment is going to depend on the basic design that has been employed — that is, CRD, RCBD or LS. The design will determine the unrandomized factors and structure so that you basically perform the analysis appropriate to that design. The modification is that, instead of having just a single source in the analysis of variance corresponding to treatments, you will have a source for each factor and one for each possible combinations of factors.

We will first use the procedure outlined in section VII.A for determining the analysis of variance table to establish the general form of the analysis for a two-factor CRD. Recall that, in general, such an experiment would have a total of *n* observations on two factors A and B with *a* and *b* levels, respectively.

A. Description of pertinent features of the study

1. Observational unit – a unit
2. Response variable – Y
3. Unrandomized factors – Units
4. Randomized factors – A, B
5. Type of study – Two-factor CRD

B. The experimental structure

Structure	Formula
unrandomized	*n* Units
randomized	*a* A**b* B

C. Terms derived from the structure formulae

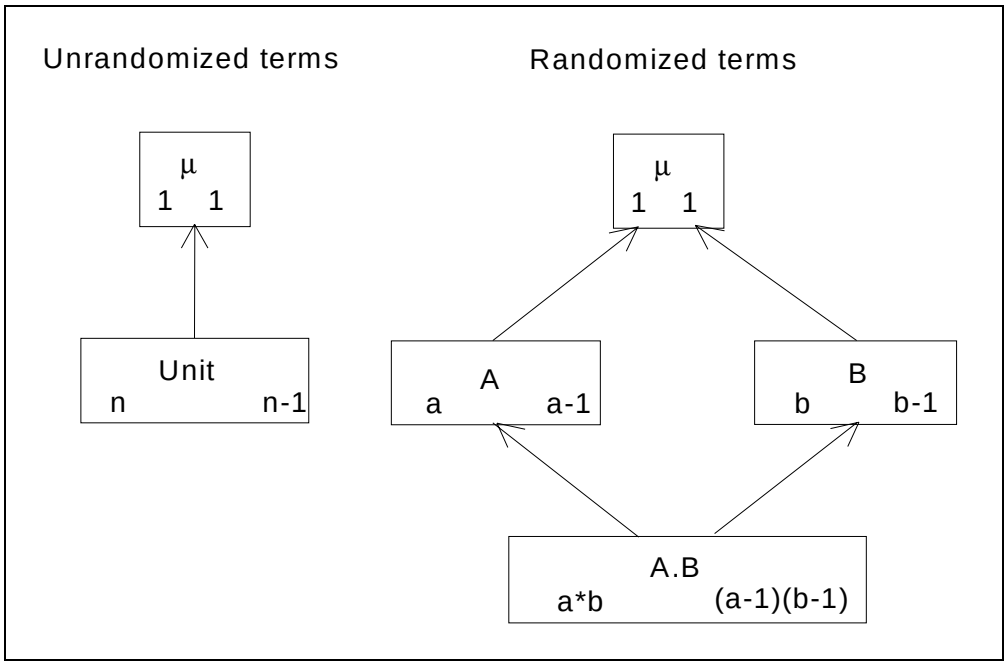
Units = Units

A*B = A + B + A.B.

(Note that for a three factor experiment, with factors A, B, and C, the sources will be: A*B*C = A + B + C + A.B + A.C + B.C + A.B.C.)

D. Degrees of freedom

Hasse diagrams for two-factor CRD



E. The analysis of variance table

Source	Df
Units	n–1
A	a–1
B	b–1
A.B	(a–1)(b–1)
Residual	ab(r–1)

F. Maximal expectation and variation models

Assume the randomized factors are fixed and that the unrandomized factor is a random factor. Then the expectation terms are A, B and A.B. The variation term is: Units.

The maximal expectation model is

$$E[Y] = A.B$$

and the variation model is

$$var[Y] = Units$$

G. The expected mean squares.

The variation component is σ^2_U with multiplier 1.

Source	df	E[MSq]
Units	n–1	
A	a–1	$\sigma^2_U + f_A(\psi)$
B	b–1	$\sigma^2_U + f_B(\psi)$
A.B	(a–1)(b–1)	$\sigma^2_U + f_{AB}(\psi)$
Residual	ab(r–1)	σ^2_U

This is the general form of the analysis of variance table. It is based on fitting models in a sequence such that marginal terms are fitted before the terms to which they are marginal. Thus the following sequence of models could be used in the case of the two-factor CRD:

$$E[\mathbf{Y}] = \mathbf{X}_G\mu$$
$$E[\mathbf{Y}] = \mathbf{X}_A\alpha$$
$$E[\mathbf{Y}] = \mathbf{X}_A\alpha + \mathbf{X}_B\beta$$
$$E[\mathbf{Y}] = \mathbf{X}_{AB}(\alpha\beta)$$

The sequence of models fitted here is arbitrary only to the extent that the sequence of fitting of A and B is immaterial, provided the number of observations for each treatment is the same; however, A.B must be fitted after A and B have been fitted since A and B are marginal to A.B. Note also that we have always fitted G first since it is marginal to all other terms.

a) Expressions for the sums of squares

The sums of squares that we require are R(α| μ), R(β| μ), R($\alpha\beta$| α,β) and D($\alpha\beta$), the A, B and A.B reduction sums of squares and the deviance from the maximal model, respectively. Now, by definition, R(α| μ) = D(μ) – D(α), R(β| μ) = D(μ) – D(β) and R($\alpha\beta$| α,β) = D(α,β) – D($\alpha\beta$). We conjecture that R(α| μ) = R(α| β) and R(β| μ) = R(β| α) so that, by definition, alternative expressions for the A and B reduction sums of squares are R(α| μ) = D(β) – D(α,β), R(β| μ) = D(α) – D(α,β). Consequently, the sums of squares for the analysis can be expressed in terms of the following residual sums of squares (or unscaled deviance) D(μ), D(α), D(α,β) and D($\alpha\beta$), corresponding to the above sequence of models.

As for previous orthogonal analyses, we would expect to be able to obtain the sums of squares by decomposing the data vector using successive applications of mean and residual operators. It would be expected that the sequence of mean operators to be applied, given the sequence of models listed above, is as follows:

1. the grand mean operator — $\mathbf{P}_G$
2. the A mean operator — $\mathbf{P}_A$
3. the B mean operator — $\mathbf{P}_B$
4. the A.B combinations mean operator — $\mathbf{P}_{AB}$

Applying these operators using the recursive procedures would produce the following vectors:

1.	$\mathbf{g} = \mathbf{P}_G\mathbf{y}$	$\mathbf{e}_G = \mathbf{R}_G\mathbf{y}$
2.	$\mathbf{a}_e = \mathbf{P}_A\mathbf{R}_G\mathbf{y}$	$\mathbf{e}_A = \mathbf{R}_A\mathbf{R}_G\mathbf{y}$
3.	$\mathbf{b}_e = \mathbf{P}_B\mathbf{R}_A\mathbf{R}_G\mathbf{y}$	$\mathbf{e}_{A+B} = \mathbf{R}_B\mathbf{R}_A\mathbf{R}_G\mathbf{y}$
4.	$\mathbf{a.b}_e = \mathbf{P}_{AB}\mathbf{R}_B\mathbf{R}_A\mathbf{R}_G\mathbf{y}$	$\mathbf{e}_{AB} = \mathbf{R}_{AB}\mathbf{R}_B\mathbf{R}_A\mathbf{R}_G\mathbf{y}$

By considering what happens at each of the above steps one can straightforwardly show that the sequence decomposes the data vector as follows:

$$\mathbf{y} = \mathbf{g} + \mathbf{a}_e + \mathbf{b}_e + \mathbf{a.b}_e + \mathbf{e}_{A.B}$$

We conjecture that the estimators of the required sums of squares are given by $R(\alpha \mid \mu) = \mathbf{A}_e'\mathbf{A}_e$, $R(\beta \mid \mu) = \mathbf{B}_e'\mathbf{B}_e$, $R(\alpha\beta \mid \alpha, \beta) = (\mathbf{A.B}_e)'(\mathbf{A.B}_e)$ and $D(\alpha\beta) = \mathbf{E}_{AB}'\mathbf{E}_{AB}$. It is these results that are the subject of the theorems in this section. However, before proceeding to the theorems we note that, given the similarity of the operators $\mathbf{P}_G$, $\mathbf{P}_A$ and $\mathbf{P}_B$ to $\mathbf{P}_G$, $\mathbf{P}_B$ and $\mathbf{P}_T$ from the RCBD, it is to be expected that the same relationship exists between operators within each set. The following lemma, which parallels lemma V.1 for $\mathbf{P}_G$, $\mathbf{P}_B$ and $\mathbf{P}_T$, provides us with the relationship between $\mathbf{P}_G$, $\mathbf{P}_A$ and $\mathbf{P}_B$.

Lemma IX.1: Let $\mathbf{P}_G = n^{-1}\mathbf{J}_a \otimes \mathbf{J}_b \otimes \mathbf{J}_r$, $\mathbf{P}_A = (br)^{-1}\mathbf{I}_a \otimes \mathbf{J}_b \otimes \mathbf{J}_r$ and $\mathbf{P}_B = (ar)^{-1}\mathbf{J}_a \otimes \mathbf{I}_b \otimes \mathbf{J}_r$. Then

$$\mathbf{P}_A\mathbf{P}_B = \mathbf{P}_B\mathbf{P}_A = \mathbf{P}_A\mathbf{P}_G = \mathbf{P}_G\mathbf{P}_A = \mathbf{P}_B\mathbf{P}_G = \mathbf{P}_G\mathbf{P}_B = \mathbf{P}_G.$$

where $\mathbf{P}_G$, $\mathbf{P}_A$ and $\mathbf{P}_B$ are symmetric and idempotent.

Proof: similar to that for lemma V.1 ■

A further useful result involving the relationship between $\mathbf{P}_{AB}$ and the other operators is contained in the following lemma.

Lemma IX.2: Let $\mathbf{P}_G = n^{-1}\mathbf{J}_a \otimes \mathbf{J}_b \otimes \mathbf{J}_r$, $\mathbf{P}_A = (br)^{-1}\mathbf{I}_a \otimes \mathbf{J}_b \otimes \mathbf{J}_r$, $\mathbf{P}_B = (ar)^{-1}\mathbf{J}_a \otimes \mathbf{I}_b \otimes \mathbf{J}_r$ and $\mathbf{P}_{AB} = r^{-1}\mathbf{I}_a \otimes \mathbf{I}_b \otimes \mathbf{J}_r$. Also let $\mathbf{R}_G = \mathbf{I} - \mathbf{P}_G$, $\mathbf{R}_A = \mathbf{I} - \mathbf{P}_A$ and $\mathbf{R}_B = \mathbf{I} - \mathbf{P}_B$ and $\mathbf{R}_{AB} = \mathbf{I} - \mathbf{P}_{AB}$. Then

$$\mathbf{P}_{AB}\mathbf{P}_A = \mathbf{P}_A\mathbf{P}_{AB} = \mathbf{P}_A$$
$$\mathbf{P}_{AB}\mathbf{P}_B = \mathbf{P}_B\mathbf{P}_{AB} = \mathbf{P}_B$$
$$\mathbf{P}_{AB}\mathbf{P}_G = \mathbf{P}_G\mathbf{P}_{AB} = \mathbf{P}_G$$

and so

$$\mathbf{R}_{AB}\mathbf{R}_B\mathbf{R}_A\mathbf{R}_G = \mathbf{R}_{AB}$$

where $\mathbf{P}_{AB}$ is symmetric and idempotent.

Proof: left as an exercise ■

For our conjectures about the equivalence of various sums of squares to hold requires the equivalence of various sequences of mean and residual operators. These are the subject of the following lemma, as are expressions in terms of linear combinations of mean operators.

Lemma IX.3: Let $\mathbf{P}_G$, $\mathbf{P}_A$, $\mathbf{P}_B$, $\mathbf{P}_{AB}$, $\mathbf{R}_G$, $\mathbf{R}_A$ and $\mathbf{R}_B$ be as defined in lemma IX.2. Then

$$\mathbf{P}_A\mathbf{R}_B\mathbf{R}_G = \mathbf{P}_A\mathbf{R}_G = \mathbf{P}_A - \mathbf{P}_G$$
$$\mathbf{P}_B\mathbf{R}_A\mathbf{R}_G = \mathbf{P}_B\mathbf{R}_G = \mathbf{P}_B - \mathbf{P}_G$$
$$\mathbf{P}_{AB}\mathbf{R}_B\mathbf{R}_A\mathbf{R}_G = \mathbf{P}_{AB}\mathbf{R}_A\mathbf{R}_B\mathbf{R}_G = \mathbf{P}_{AB} - \mathbf{P}_A - \mathbf{P}_B + \mathbf{P}_G$$

and $\mathbf{P}_A\mathbf{R}_G$, $\mathbf{P}_B\mathbf{R}_G$ and $\mathbf{P}_{AB}\mathbf{R}_B\mathbf{R}_A\mathbf{R}_G$ are symmetric, idempotent.

Proof: left as an exercise ■

Corollary IX.1: Let $\mathbf{A}_e = \mathbf{P}_A\mathbf{R}_G\mathbf{Y} = (\mathbf{P}_A - \mathbf{P}_G)\mathbf{Y}$, $\mathbf{B}_e = \mathbf{P}_B\mathbf{R}_G\mathbf{Y} = (\mathbf{P}_B - \mathbf{P}_G)\mathbf{Y}$ and $\mathbf{A.B}_e = \mathbf{P}_{AB}\mathbf{R}_A\mathbf{R}_B\mathbf{R}_G\mathbf{Y} = (\mathbf{P}_{AB} - \mathbf{P}_A - \mathbf{P}_B + \mathbf{P}_G)\mathbf{Y}$. Then

$$\mathbf{A}_e = \mathbf{A} - \mathbf{G},\ \mathbf{B}_e = \mathbf{B} - \mathbf{G} \text{ and } \mathbf{A.B}_e = \mathbf{A.B} - \mathbf{A} - \mathbf{B} + \mathbf{G}$$

Proof: follows immediately from lemmas IX.2 and IX.3. ■

This corollary leads to the following expressions for the effects vectors:

$$\mathbf{G} = \{\bar{Y}_{..}\},\ \mathbf{A}_e = \{\bar{Y}_{i.} - \bar{Y}_{..}\},\ \mathbf{B}_e = \{\bar{Y}_{.j} - \bar{Y}_{..}\} \text{ and } \mathbf{A.B}_e = \{\bar{Y}_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..}\}$$

where $\bar{Y}_{..}$ is the grand mean or mean of all the random variables

$\bar{Y}_{i.}$ is the mean of all random variables with the ith level of A

$\bar{Y}_{.j}$ is the mean of all random variables with the *j*th level of B

$\bar{Y}_{ij}$ is the mean of all random variables with the *i*th level of A and the *j*th level of B

Also, $\mathbf{R}_{AB}\mathbf{Y} = \left\{Y_k - \bar{Y}_{ij.}\right\} = \mathbf{E}_{AB}\mathbf{Y}$.

Consequently each observation is decomposed as follows:

$$\left\{Y_k\right\} = \left\{\bar{Y}_{..}\right\} + \left\{\bar{Y}_{i.} - \bar{Y}_{..}\right\} + \left\{\bar{Y}_{.j} - \bar{Y}_{..}\right\} + \left\{\bar{Y}_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..}\right\} + \left\{Y_k - \bar{Y}_{ij}\right\}$$

Theorem IX.6: Let **Y** be a *n*-vector of jointly-distributed random variables and let $D(\mu) = \mathbf{Y}'\mathbf{R}_G\mathbf{Y}$, $D(\alpha) = \mathbf{Y}'\mathbf{R}_A\mathbf{R}_G\mathbf{Y}$, $D(\beta) = \mathbf{Y}'\mathbf{R}_B\mathbf{R}_G\mathbf{Y}$, $D(\alpha,\beta) = \mathbf{Y}'\mathbf{R}_B\mathbf{R}_A\mathbf{R}_G\mathbf{Y}$ and $D(\alpha\beta) = \mathbf{Y}'\mathbf{R}_{AB}\mathbf{R}_B\mathbf{R}_A\mathbf{R}_G\mathbf{Y}$ be the estimators for the residual sums of squares for the expectation models $E[\mathbf{Y}] = \mathbf{X}_G\mu$, $E[\mathbf{Y}] = \mathbf{X}_A\alpha$, $E[\mathbf{Y}] = \mathbf{X}_B\beta$, $E[\mathbf{Y}] = \mathbf{X}_A\alpha + \mathbf{X}_B\beta$ and $E[\mathbf{Y}] = \mathbf{X}_{AB}(\alpha\beta)$, respectively.

Then, for R(α| μ) = D(μ) – D(α), R(α| β) = D(β) – D(α,β), R(β| μ) = D(μ) – D(β), R(β| α) = D(α) – D(α,β) and R($\alpha\beta$| α,β) = D(α,β) – D($\alpha\beta$),

$$R(\alpha|\,\mu) = R(\alpha|\,\beta) = \mathbf{A}_e'\mathbf{A}_e$$
$$R(\beta|\,\mu) = R(\beta|\,\alpha) = \mathbf{B}_e'\mathbf{B}_e$$
$$R(\alpha\beta\,|\,\alpha,\beta) = (\mathbf{A.B}_e)'(\mathbf{A.B}_e)$$
$$D(\alpha\beta) = \mathbf{E}_{AB}'\mathbf{E}_{AB}$$

where $\mathbf{A}_e$, $\mathbf{B}_e$ and $\mathbf{A.B}_e$ are as defined in corollary IX.1 and $\mathbf{E}_{AB} = \mathbf{R}_{AB}\mathbf{R}_B\mathbf{R}_A\mathbf{R}_G\mathbf{Y} = \mathbf{R}_{AB}\mathbf{Y}$.

Proof: left as an exercise ■

b) Degrees of freedom

Theorem IX.7: Let **Y** be a *n*-vector of jointly-distributed random variables and $D(\mu) = \mathbf{Y}'\mathbf{R}_G\mathbf{Y}$. Let $R(\alpha|\,\mu) = R(\alpha|\,\beta) = \mathbf{A}_e'\mathbf{A}_e$, $R(\beta|\,\mu) = R(\beta|\,\alpha) = \mathbf{B}_e'\mathbf{B}_e$, $R(\alpha\beta\,|\,\alpha,\beta) = (\mathbf{A.B}_e)'(\mathbf{A.B}_e)$ and $D(\alpha\beta) = \mathbf{E}_{AB}'\mathbf{E}_{AB}$ be as defined in theorem IX.6.

Then the degrees of freedom of D(μ), R(α| μ), R(β| μ), R($\alpha\beta$| α,β) and D($\alpha\beta$) are *n*–1, *a*–1, *b*–1, (*a*–1)(*b*–1) and *n*–*ab*, respectively, where *n* is the number of observations, *a* is the number of levels of A and *b* is the number of levels of B.

Proof: First establish the traces of $\mathbf{P}_G$, $\mathbf{P}_A$, $\mathbf{P}_B$ and $\mathbf{P}_{AB}$. Then use the expressions given in lemmas IX.2 and IX.3 for the idempotents as linear combinations of these **P** matrices to derive the required degrees of freedom. ■

c) Expected mean squares

Theorem IX.8: Let **Y** be a *n*-vector of jointly-distributed random variables with $\psi = \mathbf{X}_G\mu$, $\psi = \mathbf{X}_A\alpha$, $\psi = \mathbf{X}_B\beta$, $\psi = \mathbf{X}_A\alpha + \mathbf{X}_B\beta$ and $\psi = \mathbf{X}_{AB}(\alpha\beta)$ being alternative models for the expectation and $\text{var}[\mathbf{Y}] = \sigma_U^2\mathbf{I}_n$. Also let R($\alpha$| μ), R(β| μ), R($\alpha\beta$| α,β) and D($\alpha\beta$) be as defined in theorem IX.6.

Then

$$E\left[R(\alpha|\,\mu)/(a-1)\right] = \sigma_U^2 + f_A(\psi),$$
$$E\left[R(\beta|\,\mu)/(b-1)\right] = \sigma_U^2 + f_B(\psi),$$
$$E\left[R(\alpha\beta|\,\mu)/(a-1)(b-1)\right] = \sigma_U^2 + f_{AB}(\psi) \text{ and}$$
$$E\left[D(\alpha\beta)/(n-ab)\right] = \sigma_U^2$$

where $f_A(\psi)$, $f_B(\psi)$ and $f_{AB}(\psi)$ under all models except the minimal model are summarized in the following table. For $\psi = \mathbf{X}_G\mu$, $f_A(\psi) = f_B(\psi) = f_{AB}(\psi) = 0$.

Expectation functions of expected mean squares under different expectation models

Source	Function	Expectation model $\psi = \mathbf{X}_A\alpha$	$\psi = \mathbf{X}_B\beta$	$\psi = \mathbf{X}_A\alpha + \mathbf{X}_B\beta$
A	$f_A(\psi)$	$\dfrac{rb\sum_{i=1}^{a}(\alpha_i - \bar{\alpha}_.)^2}{(a-1)}$	0	$\dfrac{rb\sum_{i=1}^{a}(\alpha_i - \bar{\alpha}_.)^2}{(a-1)}$
B	$f_B(\psi)$	0	$\dfrac{ra\sum_{j=1}^{b}(\beta_j - \bar{\beta}_.)^2}{(b-1)}$	$\dfrac{ra\sum_{j=1}^{b}(\beta_j - \bar{\beta}_.)^2}{(b-1)}$
A.B	$f_{AB}(\psi)$	0	0	0

Source	$\psi = \mathbf{X}_{AB}(\alpha\beta)$
A	$f_A(\psi) = \dfrac{rb\sum_{i=1}^{a}\left(\overline{(\alpha\beta)}_{i.} - \overline{(\alpha\beta)}_{..}\right)^2}{(a-1)}$
B	$f_B(\psi) = \dfrac{ra\sum_{j=1}^{b}\left(\overline{(\alpha\beta)}_{.j} - \overline{(\alpha\beta)}_{..}\right)^2}{(b-1)}$
A.B	$f_{AB}(\psi) = \dfrac{r\sum_{i=1}^{a}\sum_{j=1}^{b}\left((\alpha\beta)_{ij} - \overline{(\alpha\beta)}_{i.} - \overline{(\alpha\beta)}_{.j} + \overline{(\alpha\beta)}_{..}\right)^2}{(a-1)(b-1)}$

where $\bar{\alpha}_{.} = \sum_{i=1}^{a}\alpha_i / a$, $\bar{\beta}_{.} = \sum_{j=1}^{b}\beta_j / b$, $\overline{(\alpha\beta)}_{i.} = \sum_{j=1}^{b}(\alpha\beta)_{ij} / b$, $\overline{(\alpha\beta)}_{.j} = \sum_{i=1}^{a}(\alpha\beta)_{ij} / a$ and $\overline{(\alpha\beta)}_{..} = \sum_{i=1}^{a}\sum_{j=1}^{b}(\alpha\beta)_{ij} / ab$.

Note that the $f(\psi)$ for a particular expected mean square only involves either the term corresponding to that mean square or a term to which it is marginal and will only be zero when none of these occurs in the model. For example, $f_A(\psi)$ is nonzero for models involving α or $(\alpha\beta)$.

Proof: We first use theorem II.11 to show that for any sums of squares, $S = \mathbf{Y'EY}$ say, where **E** is symmetric and idempotent and $trace(\mathbf{E}) = \nu$,

$$\begin{aligned} E[S/\nu] &= E(\mathbf{Y'EY}/\nu) \\ &= \left(trace(\mathbf{EV}) + \psi'\mathbf{E}\psi\right)/\nu \\ &= \left(\sigma_U^2 trace(\mathbf{E}) + \psi'\mathbf{E}\psi\right)/\nu \\ &= \sigma_U^2 + \psi'\mathbf{E}\psi/\nu \\ &= \sigma_U^2 + f(\psi) \end{aligned}$$

It now only remains to establish that the functions $f(\psi)$ for the various sums of squares under the alternative models are as given in the above table. This is best done using the expressions in terms of linear combinations of **P**s given for matrices of the sums of squares quadratic forms in lemmas IX.2 and IX.3.

For example, for $E\left[R(\alpha\beta \mid \mu)/(a-1)(b-1)\right]$ under the model $\psi = \mathbf{X}_{AB}(\alpha\beta)$ we require to establish $f_{AB}(\psi)$ under this model.

Now as the matrix of the quadratic form for this sums of squares is $\mathbf{P}_{AB}-\mathbf{P}_A-\mathbf{P}_B+\mathbf{P}_G$,

$$f_{\mathrm{AB}}(\psi)=\psi'(\mathbf{P}_{AB}-\mathbf{P}_A-\mathbf{P}_B+\mathbf{P}_G)\psi/(a-1)(b-1)$$

So $f_{\mathrm{AB}}(\psi)$ is the sum of squares of the elements of $(\mathbf{P}_{AB}-\mathbf{P}_A-\mathbf{P}_B+\mathbf{P}_G)\psi$ divided by the degrees of freedom. But under the model being considered an element of ψ is $\{(\alpha\beta)_{ij}\}$, each $(\alpha\beta)_{ij}$ being repeated *r* times in ψ. Recalling that the **P** matrices are just mean operators it is clear that an element of $(\mathbf{P}_{AB}-\mathbf{P}_A-\mathbf{P}_B+\mathbf{P}_G)\psi$ involving the *i*th level of A and *j*th level of B is

$$\left((\alpha\beta)_{ij}-\overline{(\alpha\beta)}_{i.}-\overline{(\alpha\beta)}_{.j}+\overline{(\alpha\beta)}_{..}\right)$$

and the sum of squares of a vector containing *r* of these for each combination of *i* and *j* is

$$r\sum_{i=1}^{a}\sum_{j=1}^{b}\left((\alpha\beta)_{ij}-\overline{(\alpha\beta)}_{i.}-\overline{(\alpha\beta)}_{.j}+\overline{(\alpha\beta)}_{..}\right)^2$$

and so $f_{\mathrm{AB}}(\psi)$ under this model is as given in the table.

Deriving the other expressions for the functions $f(\psi)$ is left as an exercise. ■

Example IX.4 2×2 Factorial experiment (continued)

We have already demonstrated for this example that

$$\psi=\begin{bmatrix}(\alpha\beta)_{11}\\(\alpha\beta)_{11}\\(\alpha\beta)_{11}\\(\alpha\beta)_{12}\\(\alpha\beta)_{12}\\(\alpha\beta)_{12}\\(\alpha\beta)_{21}\\(\alpha\beta)_{21}\\(\alpha\beta)_{21}\\(\alpha\beta)_{22}\\(\alpha\beta)_{22}\\(\alpha\beta)_{22}\end{bmatrix}.\ \text{Hence}\ \mathbf{P}_{\mathrm{A}}\psi=\mathbf{P}_{\mathrm{A}}\begin{bmatrix}(\alpha\beta)_{11}\\(\alpha\beta)_{11}\\(\alpha\beta)_{11}\\(\alpha\beta)_{12}\\(\alpha\beta)_{12}\\(\alpha\beta)_{12}\\(\alpha\beta)_{21}\\(\alpha\beta)_{21}\\(\alpha\beta)_{21}\\(\alpha\beta)_{22}\\(\alpha\beta)_{22}\\(\alpha\beta)_{22}\end{bmatrix}=\begin{bmatrix}\overline{(\alpha\beta)}_{1.}\\\overline{(\alpha\beta)}_{1.}\\\overline{(\alpha\beta)}_{1.}\\\overline{(\alpha\beta)}_{1.}\\\overline{(\alpha\beta)}_{1.}\\\overline{(\alpha\beta)}_{1.}\\\overline{(\alpha\beta)}_{2.}\\\overline{(\alpha\beta)}_{2.}\\\overline{(\alpha\beta)}_{2.}\\\overline{(\alpha\beta)}_{2.}\\\overline{(\alpha\beta)}_{2.}\\\overline{(\alpha\beta)}_{2.}\end{bmatrix}$$

where $\overline{(\alpha\beta)}_{i.} = r\sum_{j=1}^{b}(\alpha\beta)_{ij} \Big/ rb = \sum_{j=1}^{b}(\alpha\beta)_{ij} \Big/ b$. That is $\overline{(\alpha\beta)}_{i.}$ is the mean of all parameters that have subscript *i* — it is said to be the mean over the subscript that has been replaced by a dot.

Similar vectors could be established for $\mathbf{P}_G\psi$, $\mathbf{P}_B\psi$ and $\mathbf{P}_{AB}\psi$.

d) Distribution of the F statistic

The analysis of variance involve three F statistics. It can be proved that the F statistic that tests for whether or not one of the terms should be in the model term is distributed as Snedecor's (central) F. The proofs will be similar to those from the analyses for other experiments.

e) Hypothesis test

The hypothesis test is a follows:

Step 1: Set up hypotheses

a) H_0: $\alpha_1 = \alpha_2 = ... = \alpha_a$
H_1: at least one pair of population A means is different

b) H_0: $\beta_1 = \beta_2 = ... = \beta_b$
H_1: at least one pair of population B means is different

c) H_0: there is no interaction between A and B
$\left((\alpha\beta)_{ij} - \overline{(\alpha\beta)}_{i.} - \overline{(\alpha\beta)}_{.j} + \overline{(\alpha\beta)}_{..} = 0 \quad \text{for all i,j}\right)$

H_1: there is an interaction between A and B
$\left((\alpha\beta)_{ij} - \overline{(\alpha\beta)}_{i.} - \overline{(\alpha\beta)}_{.j} + \overline{(\alpha\beta)}_{..} \neq 0 \quad \text{for some i,j}\right)$

Step 2: Calculate test statistics

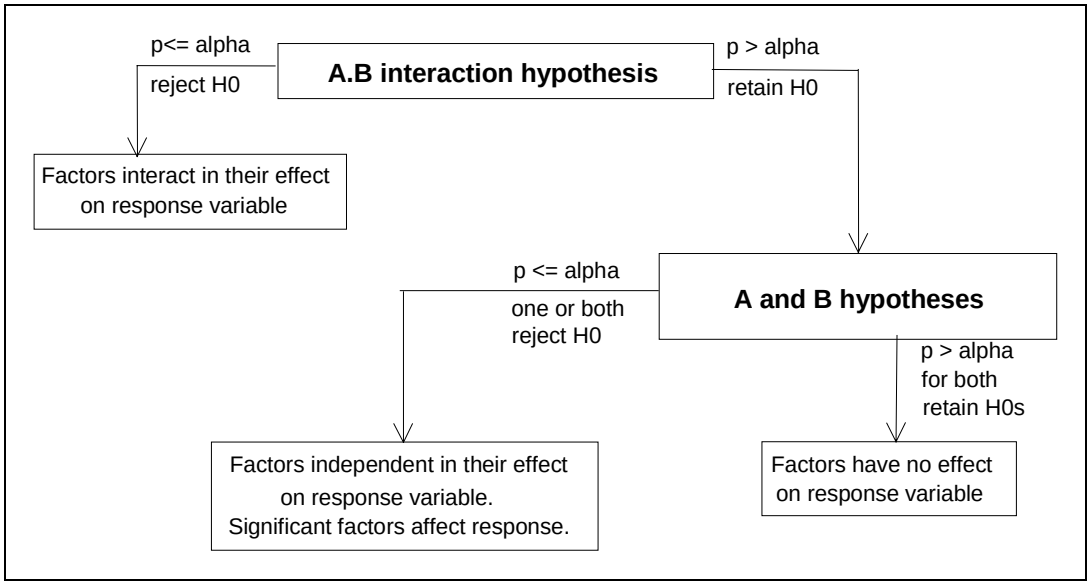
The analysis of variance table is as follows:

Source	df	SSq	E[MSq]	F
Units	n–1	$D(\mu)=(\mathbf{y}-\mathbf{g})'(\mathbf{y}-\mathbf{g})$		
A	a–1	$R(\alpha \mid \mu)=\mathbf{a}_e'\mathbf{a}_e$ †	$\sigma_U^2+f_A(\psi)$	s_A^2/s_R^2
B	b–1	$R(\beta \mid \mu)=\mathbf{b}_e'\mathbf{b}_e$ †	$\sigma_U^2+f_B(\psi)$	s_B^2/s_R^2
A.B	(a–1)(b–1)	$R(\alpha\beta \mid \alpha,\beta)=(\mathbf{ab}_e)'(\mathbf{ab}_e)$ †	$\sigma_U^2+f_{AB}(\psi)$	s_{AB}^2/s_R^2
Residual	ab(r–1)	$D(\alpha\beta)=\mathbf{e}_{AB}'\mathbf{e}_{AB}$	σ_U^2	

† $R(\beta|\mu) = D(\mu) - D(\beta) = R(\beta|\alpha)$
$R(\alpha|\mu) = D(\mu) - D(\alpha) = R(\alpha|\beta)$
$R(\alpha\beta|\alpha,\beta) = D(\alpha,\beta) - D(\alpha\beta)$

Step 3: Decide between hypotheses

The steps for choosing the model that best describes the expectation of the response is illustrated in the following diagram. The fundamental point of this diagram is that one first examines the A.B interaction and, if it is not significant, the tests for A and B effects are not proceeded with. For any hypothesis, if $P(F \geq F_{calc}) \leq 0.05$, then the evidence suggests that the null hypothesis be rejected.



The rationale for this procedure arises from the expected mean squares under the alternative expectation models. We noted above that the $f(\psi)$ for a particular expected mean square only involves either the term corresponding to that mean square or a term to which it is marginal and will only be zero when none of these occurs in the model. This pattern can be seen in the table below. From this it can be concluded that the $f(\psi)$ for a

particular mean square will only go to zero when the term corresponding to it is dropped from the model, because all terms marginal to it will have also been dropped from the model.

	True model for $\psi = E[Y]$				
Function of ψ	$\mathbf{X}_G\mu$	$\mathbf{X}_A\alpha$	$\mathbf{X}_B\beta$	$\mathbf{X}_A\alpha + \mathbf{X}_B\beta$	$\mathbf{X}_{AB}(\alpha\beta)$
$f_A(\psi)$	$=0$	$\neq 0$	$=0$	$\neq 0$	$\neq 0$
$f_B(\psi)$	$=0$	$=0$	$\neq 0$	$\neq 0$	$\neq 0$
$f_{AB}(\psi)$	$=0$	$=0$	$=0$	$=0$	$\neq 0$

For example, an examination of the expressions for $f_{AB}(\psi)$ reveals that it is zero for all models except the model $\psi = \mathbf{X}_{AB}(\alpha\beta)$. Now the significance test for A.B is a test of whether $f_{AB}(\psi)$ is zero. If A.B is significant, we conclude that $f_{AB}(\psi) \neq 0$ and the model that best describes the data is $\psi = \mathbf{X}_{AB}(\alpha\beta)$ — no other model is consistent with a non-zero $f_{AB}(\psi)$. Having determined the model to be used for the data, there is no point to doing any further hypothesis tests; they will not provide any further information about the model.

However, if A.B is not significant, we conclude that $f_{AB}(\psi) = 0$ and the model $\psi = \mathbf{X}_{AB}(\alpha\beta)$ is not the best model for the data. Examination of the expected mean squares under the models $\psi = \mathbf{X}_A\alpha$, $\psi = \mathbf{X}_B\beta$ and $\psi = \mathbf{X}_A\alpha + \mathbf{X}_B\beta$ reveals that the choice between these models depends on which of A and B are not significant. A term corresponding to the significant source must be included in the model. For example, if only B is significant, the model for the data would be $\psi = \mathbf{X}_B\beta$ because this is the only model for which $f_{AB}(\psi) = f_A(\psi) = 0$ and $f_B(\psi) \neq 0$. If neither A nor B is significant, then the model for the data would be $\psi = \mathbf{X}_G\mu$ because this is the only model for which $f_{AB}(\psi) = f_A(\psi) = f_B(\psi) = 0$.

IX.E Computation in Genstat and analysis of an example

Example IX.5 Animal survival experiment

To demonstrate the analysis I will use the example from Box, Hunter and Hunter (sec. 7.7). In this experiment three poisons and four treatments were investigated. The 12 treatments were applied to animals using a CRD and the survival times of the animals measured (10 hours). The data are as follows:

	Treatment

		1	2	3	4
	I	0.31	0.82	0.43	0.45
		0.45	1.10	0.45	0.71
		0.46	0.88	0.63	0.66
		0.43	0.72	0.76	0.62
	II	0.36	0.92	0.44	0.56
Poison		0.29	0.61	0.35	1.02
		0.40	0.49	0.31	0.71
		0.23	1.24	0.40	0.38
	III	0.22	0.30	0.23	0.30
		0.21	0.37	0.25	0.36
		0.18	0.38	0.24	0.31
		0.23	0.29	0.22	0.33

A. Description of pertinent features of the study

1. Observational unit – an animal
2. Response variable – Survival Time
3. Unrandomized factors – Animals
4. Randomized factors – Treatments, Poisons
5. Type of study – Two-factor CRD

B. The experimental structure

Structure	Formula
unrandomized	*48* Animals
randomized	*4* Treatments**3* Poisons

These are the steps that need to be performed before Genstat is used to obtain the analysis. The remaining steps are left as an exercise for you.

The data has been saved in a Genstat spreadsheet and the following instructions will produce the analysis of variance for this example:

```
PRINT Animals,Treat,Poison,SurvTime
BLOCK Animals
TREAT Treat*Poison
ANOVA [FPROB=Y; PSE=LSD] SurvTime
```

The Genstat output produced by these instructions is:

```
Genstat 5  Release 4.1  (PC/Windows NT)               03 April 2000 15:37:30
Copyright 1998, Lawes Agricultural Trust (Rothamsted Experimental Station)

                ____________________________________________________

                Genstat 5 Fourth Edition - (for Windows)
                Genstat 5 Procedure Library Release PL11
                ____________________________________________________

  3  "Data taken from File: D:/ANALYSES/LM/MULTIFAC/FAC2POIS.GSH"
  4  DELETE [redefine=yes] Animals,Poison,Treat,SurvTime
  5  FACTOR [modify=yes;nvalues=48;levels=48] Animals
  6  READ Animals; frepresentation=ordinal

   Identifier    Values   Missing    Levels
      Animals        48         0        48

  9  FACTOR [modify=yes;nvalues=48;levels=3] Poison
 10  READ Poison; frepresentation=ordinal

   Identifier    Values   Missing    Levels
       Poison        48         0         3

 13  FACTOR [modify=yes;nvalues=48;levels=4] Treat
 14  READ Treat; frepresentation=ordinal

   Identifier    Values   Missing    Levels
        Treat        48         0         4

 17  VARIATE [nvalues=48] SurvTime
 18  READ SurvTime

   Identifier   Minimum      Mean   Maximum    Values   Missing
     SurvTime    0.1800    0.4794    1.2400        48         0
```

```
 23
 24  PRINT Animals,Treat,Poison,SurvTime

    Animals       Treat      Poison    SurvTime
          1           1           1      0.3100
          2           2           1      0.8200
          3           3           1      0.4300
          4           4           1      0.4500
          5           1           1      0.4500
          6           2           1      1.1000
          7           3           1      0.4500
          8           4           1      0.7100
          9           1           1      0.4600
         10           2           1      0.8800
         11           3           1      0.6300
         12           4           1      0.6600
         13           1           1      0.4300
         14           2           1      0.7200
         15           3           1      0.7600
         16           4           1      0.6200
         17           1           2      0.3600
         18           2           2      0.9200
         19           3           2      0.4400
         20           4           2      0.5600
         21           1           2      0.2900
         22           2           2      0.6100
         23           3           2      0.3500
         24           4           2      1.0200
         25           1           2      0.4000
         26           2           2      0.4900
         27           3           2      0.3100
         28           4           2      0.7100
         29           1           2      0.2300
         30           2           2      1.2400
         31           3           2      0.4000
         32           4           2      0.3800
         33           1           3      0.2200
         34           2           3      0.3000
         35           3           3      0.2300
         36           4           3      0.3000
         37           1           3      0.2100
         38           2           3      0.3700
         39           3           3      0.2500
         40           4           3      0.3600
         41           1           3      0.1800
         42           2           3      0.3800
         43           3           3      0.2400
         44           4           3      0.3100
         45           1           3      0.2300
         46           2           3      0.2900
         47           3           3      0.2200
         48           4           3      0.3300

 25  BLOCK Animals
 26  TREAT Treat*Poison
 27  ANOVA [FPROB=Y; PSE=LSD] SurvTime


27.....................................................................

***** Analysis of variance *****

Variate: SurvTime

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

Animals stratum
Treat                      3    0.92121    0.30707   13.81  <.001
Poison                     2    1.03301    0.51651   23.22  <.001
Treat.Poison               6    0.25014    0.04169    1.87  0.112
Residual                  36    0.80073    0.02224
```

```
Total                     47    3.00508

* MESSAGE: the following units have large residuals.

Animals 24          0.353   s.e. 0.129
Animals 26         -0.325   s.e. 0.129
Animals 30          0.425   s.e. 0.129
Animals 32         -0.288   s.e. 0.129


***** Tables of means *****

Variate: SurvTime

Grand mean  0.479

    Treat        1        2        3        4
            0.314    0.677    0.393    0.534

   Poison        1        2        3
            0.617    0.544    0.276

    Treat   Poison        1        2        3
        1             0.413    0.320    0.210
        2             0.880    0.815    0.335
        3             0.567    0.375    0.235
        4             0.610    0.667    0.325


*** Least significant differences of means (5% level) ***

Table                Treat      Poison       Treat
                                             Poison
rep.                    12          16           4
d.f.                    36          36          36
l.s.d.              0.1235      0.1069       0.213
```

The hypothesis test for the example is a follows:

Step 1: Set up hypotheses

a) H_0: $\tau_A = \tau_B = \tau_C = \tau_D$
   H_1: at least one pair of population Treatment means is different

b) H_0: $\rho_I = \rho_{II} = \rho_{III}$
   H_1: at least one pair of population Poison means is different

c) H_0: there is no interaction between Treatment and Poison
   $\left((\tau\rho)_{ij} - \overline{(\tau\rho)}_{i.} - \overline{(\tau\rho)}_{.j} + \overline{(\tau\rho)}_{..} = 0 \quad \text{for all i,j}\right)$

   H_1: there is an interaction between Treatment and Poison
   $\left((\tau\rho)_{ij} - \overline{(\tau\rho)}_{i.} - \overline{(\tau\rho)}_{.j} + \overline{(\tau\rho)}_{..} \neq 0 \quad \text{for some i,j}\right)$

Step 2: Calculate test statistics

The analysis of variance table for a two-factor CRD, with random factors being the unrandomized factors and fixed factors the randomized factors, is:

Source	df	SSq	MSq	E[MSq]	F	Prob
Animals	47	3.0051				
Treatment	3	0.9212	0.3071	$\sigma^2_{\mathrm{A}} + f_{\mathrm{T}}(\psi)$	13.81	<.001
Poison	2	1.0330	0.5165	$\sigma^2_{\mathrm{A}} + f_{\mathrm{P}}(\psi)$	23.22	<.000
Treat.Poison	6	0.2501	0.0417	$\sigma^2_{\mathrm{A}} + f_{\mathrm{TP}}(\psi)$	1.87	0.112
Residual	36	0.8007	0.0222	σ^2_{A}		

Step 3: Decide between hypotheses

> Although the interaction of Treatment and Poison is not significant, there is some evidence of an interaction. If we presume that there is no interaction then it can be concluded that both main effects are highly significant.

However, it remains to perform the usual diagnostic checking.

IX.F Diagnostic checking

The assumptions underlying a factorial experiment will be the same as for the basic design employed.

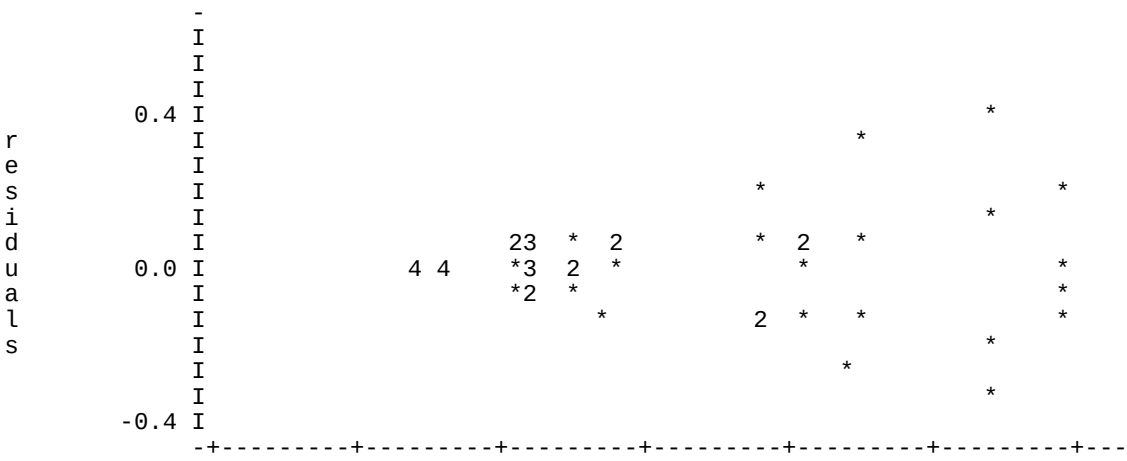
Example IX.5 Animal survival experiment (continued)

As the experiment was set up as a CRD, the assumptions underlying its analysis will be the same as for the CRD. The Genstat instructions that do diagnostic checking of the residuals for this example are as follows:

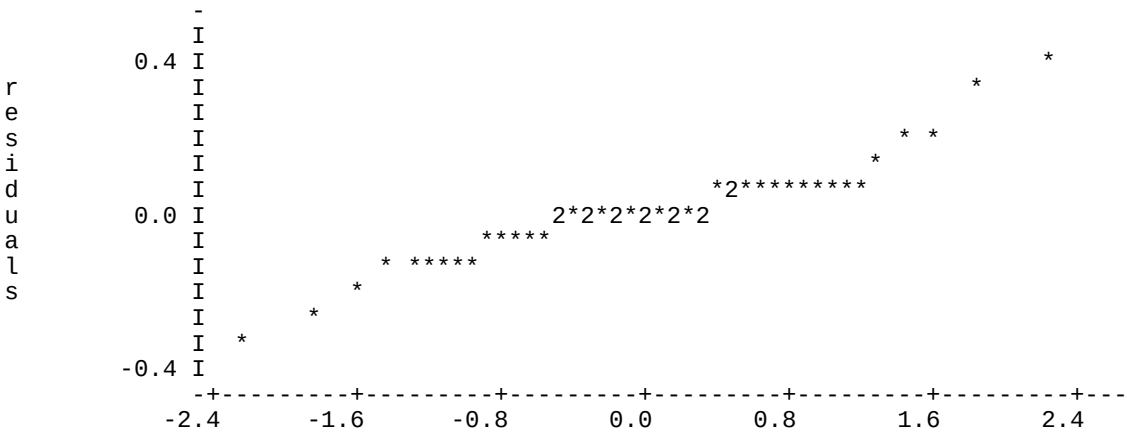
```
APLOT METHOD=fit,normal
"
**** Tukey''s one-degree-of-freedom-for-non-additivity.
**** It is the term designated covariate in the following analysis
"
AKEEP [FIT=Fit]
CALC ResSq=Fit*Fit
ANOVA [PRINT=*]  ResSq; RES=ResSq
COVAR ResSq                              "A computational trick"
ANOVA [PRINT=A; FPROB=Y] SurvTime
```

The Genstat output produced by the instructions that deal with diagnostic checking is as follows:

```
28  APLOT METHOD=fit,normal
```



Normal plot




```
 29  "
-30  **** Tukey''s one-degree-of-freedom-for-non-additivity.
-31  **** It is the term designated covariate in the following analysis
-32  "
 33  AKEEP [FIT=Fit]
 34  CALC ResSq=Fit*Fit
 35  ANOVA [PRINT=*]  ResSq; RES=ResSq
 36  COVAR ResSq                          "A computational trick"
 37  ANOVA [PRINT=A; FPROB=Y] SurvTime


37.....................................................................

***** Analysis of variance (adjusted for covariate) *****

Variate: SurvTime
Covariate: ResSq

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

Animals stratum
Treat                      3    0.92121    0.30707   13.81    1.00  <.001
Poison                     2    1.03301    0.51651   23.22    1.00  <.001
Treat.Poison               6    0.25014    0.04169    1.87    1.00  0.112
Residual                  36    0.80073    0.02224            1.00
Total                     47    3.00508
```

Note that the analysis of variance table in the above output does not include the Covariate line. This is because Tukey's one-degree-of-freedom-for-nonadditivity cannot be computed for the CRD. It would be better to leave out the instructions to compute this one degree-of-freedom.

The residual-versus-fitted-values and normal probability plots both indicate a problem with the assumptions. The residual-versus-fitted-values plot is displaying strong funnel-shape so that variance is increasing with fitted values. The normal probability plot is not displaying a straight-line trend so that the data is displaying nonnormality. The question is would a transformation fix the problem and, if so, which one?

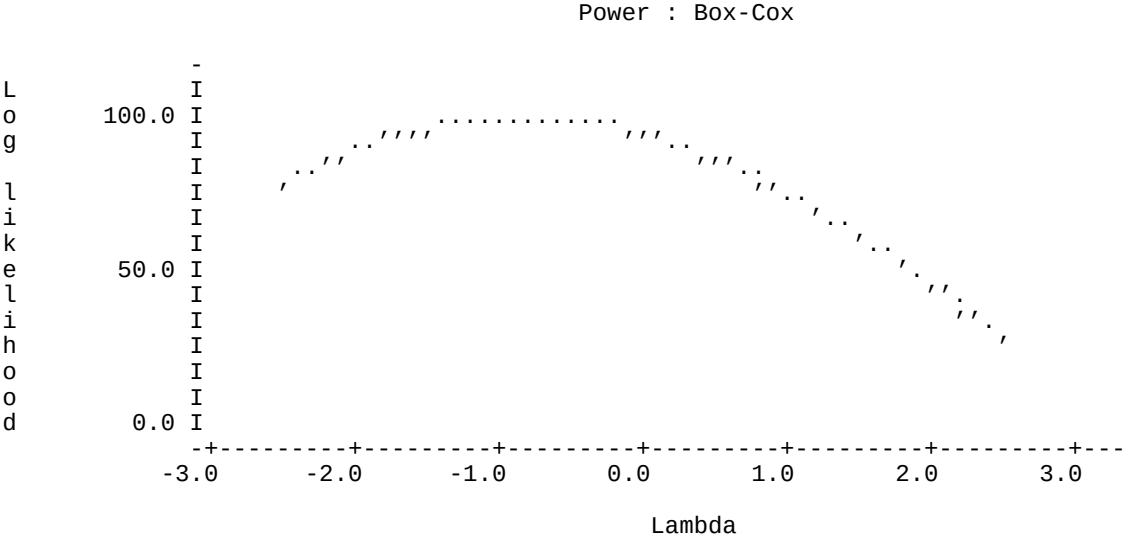
a) Box-Cox transformations for correcting transformable non-additivity

Box, Hunter and Hunter (sec. 7.9) describe the Box-Cox procedure for determining the appropriate power transformation for a set of data. It has been implemented in the Genstat procedure YTRANSFORM. When you run this procedure you obtain a plot of the log-likelihood of λ, the power of the transformation to be used (for $\lambda = 0$ use the ln transformation).

Example IX.5 Animal survival experiment (continued)

The following output has been obtained for the example and it indicates that, as the log likelihood is a maximum around $\lambda = -1$, the reciprocal transformation should be used.

```
 38  YTRANSFORM [TERMS=Treat*Poison; LOWER=-2.5; UPPER=2.5; GRAPHICS=line] \
 39                              SurvTime; SAVE=s
```



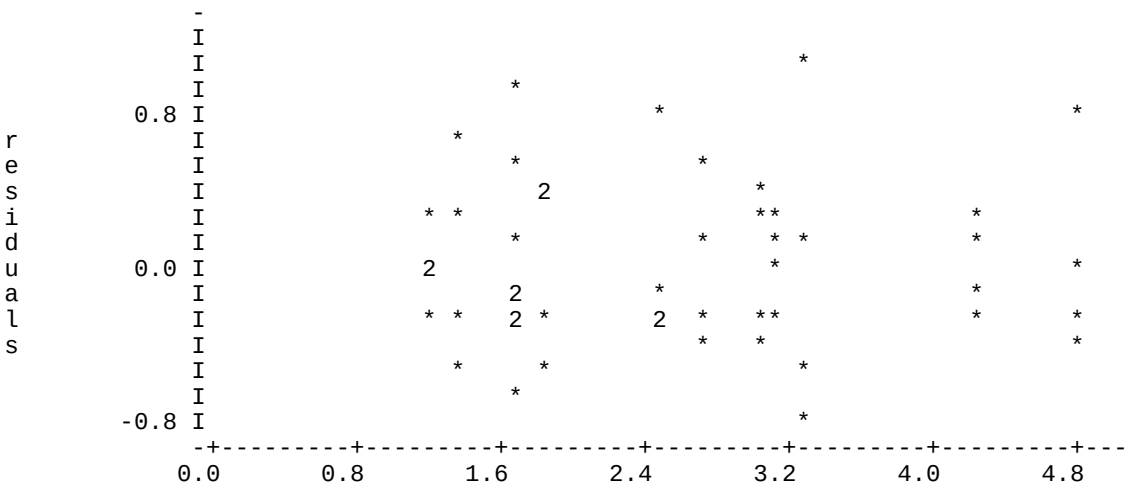
To repeat the analysis on the reciprocals I merely changed the file to include the following instruction after the data:

```
CALC SurvTime=1/SurvTime
```

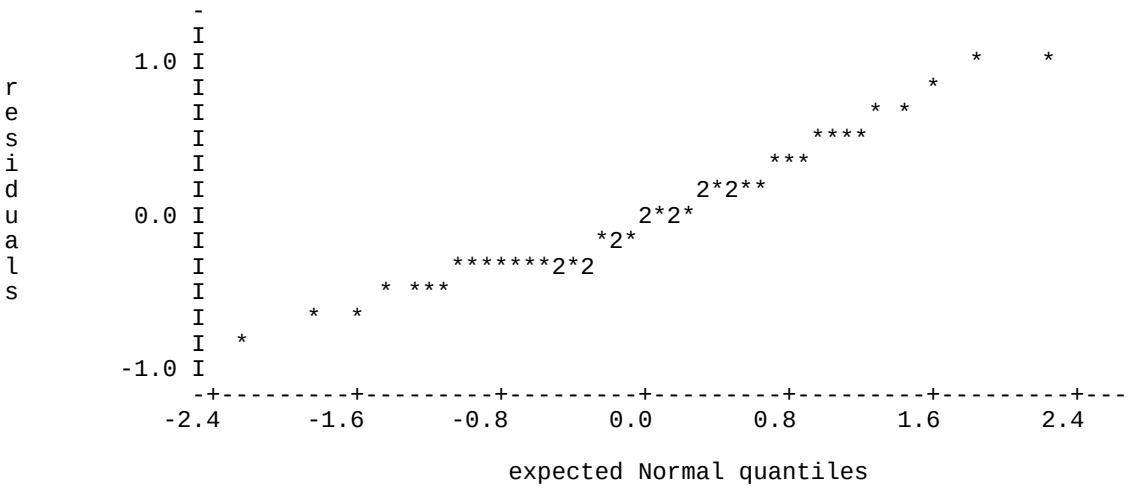
The Genstat output produced by these instructions, after data input, is:

```
  40  CALC SurvTime=1/SurvTime
  41  BLOCK Animals
  42  TREAT Treat*Poison
  43  COVAR
  44  ANOVA [FPROB=Y; PSE=LSD] SurvTime


44.....................................................................

***** Analysis of variance *****

Variate: SurvTime

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

Animals stratum
Treat                      3    20.4143     6.8048   28.34  <.001
Poison                     2    34.8771    17.4386   72.63  <.001
Treat.Poison               6     1.5708     0.2618    1.09  0.387
Residual                  36     8.6431     0.2401
Total                     47    65.5053


* MESSAGE: the following units have large residuals.

Animals 29          1.079   s.e. 0.424


***** Tables of means *****

Variate: SurvTime

Grand mean  2.622

    Treat        1        2        3        4
            3.519    1.862    2.947    2.161

   Poison        1        2        3
            1.801    2.269    3.797

    Treat   Poison        1        2        3
        1            2.487    3.268    4.803
        2            1.163    1.393    3.029
        3            1.863    2.714    4.265
        4            1.690    1.702    3.092


*** Least significant differences of means (5% level) ***

Table                Treat      Poison       Treat
                                             Poison
rep.                    12          16           4
d.f.                    36          36          36
l.s.d.              0.4057      0.3513      0.7027
```

```
  45  APLOT METHOD=fit,normal
```



Normal plot



A comparison of the untransformed and transformed analyses follows:

		UNTRANSFORMED				TRANSFORMED		
Source	df	MSq	F	Prob	df	MSq	F	Prob
Animals	47				47			
Treatment	3	0.3071	13.81	0.0000	3	6.8048	28.34	<0.001
Poison	2	0.5165	23.22	0.0000	2	17.4386	72.63	<0.001
Treat.Poison	6	0.0417	1.87	0.1112	6	0.2618	1.09	0.387
Residual	36	0.0222			36	0.2401		

The analysis of the transformed data indicates that there is no interaction on the transformed scale with the interaction mean square being nearly equal to the Residual mean squares. The main effect mean squares are even larger than before indicating that we are able to separate the treatments even more on the transformed scale.

The diagnostic checking now indicates that all the assumptions are met.

IX.G Treatment differences

As usual the examination of treatment differences can be based on multiple comparisons or submodels. If all the factors are qualitative, multiple comparison procedures would be performed on the appropriate tables of means. If one or more of the factors are quantitative then submodels would be appropriate.

a) Multiple comparison procedures

For two factor experiments, there will be altogether three tables of means, namely one for each of A, B and A.B. Which table is of interest depends on the results of the hypothesis tests outlined above.

A.B Interaction significant

In this case you look at the table of means for the A.B combinations.

		A						
		1	2	3	.	.	.	*a*
	1	x	x	x	.	.	.	x
	2	x	x	x	.	.	.	x
	.	.	.	.	.	.	.	.
B	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.
	b	x	x	x	.	.	.	x

$$w(5\%) = \frac{q_{ab,\upsilon,0.05}}{\sqrt{2}}\, s_{\bar{x}_d} = \frac{q_{ab,\upsilon,0.05}}{\sqrt{2}}\, s\sqrt{\frac{2}{r}}$$

$$LSD(5\%) = t_{\nu,0.05}\, s_{\bar{x}_d} = t_{\nu,0.05}\, s\sqrt{\frac{2}{r}}$$

A.B interaction not significant

In this case examine the A and B tables of means for the significant lines.

	A						
	1	2	3	.	.	.	*a*
Means	x	x	x	.	.	.	x

$$w(5\%) = \frac{q_{a,\upsilon,0.05}}{\sqrt{2}}\, s_{\bar{x}_d} = \frac{q_{a,\upsilon,0.05}}{\sqrt{2}}\, s\sqrt{\frac{2}{rb}}$$

$$LSD(5\%) = t_{\nu,0.05}\, s_{\bar{x}_d} = t_{\nu,0.05}\, s\sqrt{\frac{2}{rb}}$$

	B						
	1	2	3	.	.	.	*b*
Means	x	x	x	.	.	.	x

$$w(5\%) = \frac{q_{b,\upsilon,0.05}}{\sqrt{2}}\, s_{\bar{x}_d} = \frac{q_{b,\upsilon,0.05}}{\sqrt{2}}\, s\sqrt{\frac{2}{ra}}$$

$$LSD(5\%) = t_{\nu,0.05}\, s_{\bar{x}_d} = t_{\nu,0.05}\, s\sqrt{\frac{2}{ra}}$$

That is, we examine each factor separately.

For our example, as the interaction is not significant, the overall tables of means are examined.

Poison			
I	II	III	w(5%)
1.801	2.269	3.797	0.4298

$$\left[\text{Note: } w(5\%) = \frac{3.46}{\sqrt{2}}\sqrt{\frac{0.2469\times 2}{16}}\right]$$

Treatment				
1	3	4	2	w(5%)
3.519	2.947	2.161	1.862	0.5465

$$\left[\text{Note: } w(5\%) = \frac{3.81}{\sqrt{2}}\sqrt{\frac{0.2469\times 2}{12}}\right]$$

b) Polynomial submodels

As stated previously, the formal expression for maximal model for a two-factor CRD experiment, where the two randomized factors A and B are fixed, is:

$$\psi = E[\mathbf{Y}] = \mathbf{X}_{AB}(\alpha\beta) \text{ and } \mathbf{V} = \sigma_U^2 \mathbf{I}_n,$$

where $\mathbf{Y}$ is the *n*-random vector of random variables representing the response variable,

$(\alpha\beta)$ is the *ab*-vector of parameters for the A-B combinations, and

σ_U^2 is the variability arising from different units.

In respect of fitting polynomial submodels, two situations are possible:

i) one factor only is quantitative, or
ii) both factors are quantitative.

One quantitative and one qualitative factor

In investigating submodels for a two-factor factorial with one factor, B say, quantitative, the following set of models for the expectation is considered:

$E[\mathbf{Y}] = \mathbf{X}_{AB}(\alpha\beta)$	depends on combination of A and B
$E[\mathbf{Y}] = \mathbf{X}_{A}\alpha + \mathbf{X}_{A1}(\alpha\theta)_1 + \mathbf{X}_{A2}(\alpha\theta)_2$	*quadratic* response to B, differing for A
$E[\mathbf{Y}] = \mathbf{X}_{A}\alpha + \mathbf{X}_{A1}(\alpha\theta)_1$	linear response to B, differing for A
$E[\mathbf{Y}] = \mathbf{X}_{A}\alpha + \mathbf{X}_{B}\beta$	nonsmooth, independent response to A & B
$E[\mathbf{Y}] = \mathbf{X}_{A}\alpha + \mathbf{X}_{2}\theta_2$	quadratic response to B, intercept differs for A
$E[\mathbf{Y}] = \mathbf{X}_{A}\alpha + \mathbf{X}_{1}\gamma_1$	linear response to B, intercept differs for A
$E[\mathbf{Y}] = \mathbf{X}_{A}\alpha$	nonsmooth response, depends on A only
$E[\mathbf{Y}] = \mathbf{X}_{B}\beta$	nonsmooth response, depends on B only
$E[\mathbf{Y}] = \mathbf{X}_{G}\mu + \mathbf{X}_{2}\theta_2$	quadratic response to B, A has no effect
$E[\mathbf{Y}] = \mathbf{X}_{G}\mu + \mathbf{X}_{1}\gamma_1$	*linear* response to B, A has no effect
$E[\mathbf{Y}] = \mathbf{X}_{G}\mu$	neither factor affects the response

where $(\alpha\beta) = \{(\alpha\beta)_{ij}\}$

$(\alpha\theta)_1 = \{(\alpha\gamma)_{i1}\}$

$(\alpha\theta)_2 = \{(\alpha\gamma)_{i2}\}$

$\alpha = \{\alpha_i\}$

$\theta_2' = [\gamma_1 \quad \gamma_2]$

The above models are ordered from the most complex to the simplest. Why this set of expectation models? The following definition and rules are used in formulating them.

Definition IX.9: A **polynomial term** is one in which the **X** matrix involves the quantitative levels of a factor(s). ■

Definition IX.10: The **degree** for a polynomial term with respect to a quantitative factor is the power to which levels of that factor are to be raised in this term. ■

Definition IX.11: A polynomial term is said to be of **lower degree** than a second polynomial term if the degrees of all quantitative factors in first term are less than or

equal to those of the second term and the power of at least one factor in the first term is less than that of the same factor in the second term. ■

Definition IX.12: A polynomial term is said to be of **higher degree** than a second polynomial term if the degrees of all quantitative factors in first term are greater than or equal to those of the second term and the power of at least one factor in the first term is higher than that of the same factor in the second term. ■

As before, γs are used for the coefficients of polynomial terms and a numeric subscript for each quantitative fixed factor in the experiment is placed on the γs to indicate the degree(s) to which the factor(s) is(are) raised.

Rule IX.1: An expectation model must include all terms of lower degree than a term that has been put in the model. ■

In addition to rule IX.1, rule VII.7 specifies that any terms marginal to other terms in the model are removed from the model. Note that the term $\mathbf{X}_1\gamma_1$ is not marginal to $\mathbf{X}_2\gamma_2$ — the column $\mathbf{X}_1$ is not a linear combination of the column $\mathbf{X}_2$. However, the rules above imply that if term $\mathbf{X}_2\gamma_2$ is included in the model, so must the term $\mathbf{X}_1\gamma_1$. As far as the marginality of *models* is concerned, the model involving just $\mathbf{X}_1\gamma_1$ is marginal to the model consisting of $\mathbf{X}_1\gamma_1$ and $\mathbf{X}_2\gamma_2$. Also note that the term $\mathbf{X}_1\gamma_1$ is marginal to $\mathbf{X}_{A1}(\alpha\theta)_1$ since $\mathbf{X}_1$ is the sum of the columns of $\mathbf{X}_{A1}$. Consequently, a model containing $\mathbf{X}_{A1}(\alpha\theta)_1$ will not contain $\mathbf{X}_1\gamma_1$.

In general, the models to which a particular model is marginal will be found above it in the list; however, a model is marginal to only some, not all, of the models above it in this list. Note that the last four models differ from the four models immediately above them only in not including α.

The expressions for these models, in terms of a single observation, are as follows:

$$E[Y_k] = (\alpha\beta)_{ij}$$
$$E[Y_k] = \alpha_i + (\alpha\gamma)_{i1} x_{\beta_j} + (\alpha\gamma)_{i2} x_{\beta_j}^2$$
$$E[Y_k] = \alpha_i + (\alpha\gamma)_{i1} x_{\beta_j}$$
$$E[Y_k] = \alpha_i + \beta_j$$
$$E[Y_k] = \alpha_i + \gamma_1 x_{\beta_j} + \gamma_2 x_{\beta_j}^2$$
$$E[Y_k] = \alpha_i + \gamma_1 x_{\beta_j}$$
$$E[Y_k] = \alpha_i$$
$$E[Y_k] = \beta_j$$
$$E[Y_k] = \mu + \gamma_1 x_{\beta_j} + \gamma_2 x_{\beta_j}^2$$
$$E[Y_k] = \mu + \gamma_1 x_{\beta_j}$$
$$E[Y_k] = \mu$$

where Y_k is the random variable representing the response variable for the *k*th plot,

μ is the overall level of the response variable in the experiment,

α_i is the overall effect of the *i*th level of factor A on the response,

x_{β_j} is the value of the *j*th level of factor B,

γ_1 and γ_2 are the linear and quadratic coefficients of the equation describing the change in response as the level of B changes,

β_j is the overall effect of the *j*th level of factor B on the response,

$(\alpha\gamma)_{i1}$ and $(\alpha\gamma)_{i2}$ are the linear and quadratic coefficients of the equation describing, for the *i*th level of A, the change in response as the level of B changes,

$(\alpha\beta)_{ij}$ is the interaction between the *i*th level of A and the *j*th level of B.

Below is a sequence of models that can be fitted to produce the analysis of variance for this situation. The sequence of models fitted here is restricted in that: linear terms must be fitted before corresponding quadratic terms; main effects must be fitted before interactions. The deviances after fitting the models given below are, respectively: D(μ), D(α), D(α,γ_1), D(α,γ_1,γ_2), D(α,β), D($\alpha,\beta,(\alpha\gamma)_1$), D($\alpha,\beta,(\alpha\gamma)_1,(\alpha\gamma)_2$), D($\alpha\beta$).

$$E[\mathbf{Y}] = \mathbf{X}_{G}\mu$$
$$E[\mathbf{Y}] = \mathbf{X}_{A}\alpha$$
$$E[\mathbf{Y}] = \mathbf{X}_{A}\alpha + \mathbf{X}_{1}\theta_{1}$$
$$E[\mathbf{Y}] = \mathbf{X}_{A}\alpha + \mathbf{X}_{2}\theta_{2}$$
$$E[\mathbf{Y}] = \mathbf{X}_{A}\alpha + \mathbf{X}_{B}\beta$$
$$E[\mathbf{Y}] = \mathbf{X}_{A}\alpha + \mathbf{X}_{B}\beta + \mathbf{X}_{A1}(\alpha\theta)_{1}$$
$$E[\mathbf{Y}] = \mathbf{X}_{A}\alpha + \mathbf{X}_{B}\beta + \mathbf{X}_{A1}(\alpha\theta)_{1} + \mathbf{X}_{A2}(\alpha\theta)_{2}$$
$$E[\mathbf{Y}] = \mathbf{X}_{AB}\ (\alpha\beta\ )$$

The analysis of variance table for a two-factor CRD with one quantitative factor is:

Source	df	SSq
Individuals	$n-1$	
A	$a-1$	$R(\alpha\|\mu)$
B	$b-1$	$R(\beta\|\mu)$
Linear	1	$R(\gamma_1\|\mu)$
Quadratic	1	$R(\gamma_2\|\mu,\gamma_1)$
Deviations	$b-3$	$R(\beta\|\mu,\gamma_1,\gamma_2)$
A.B	$(a-1)(b-1)$	$R(\alpha\beta\|\alpha,\beta)$
$A.B_{Linear}$	$a-1$	$R((\alpha\gamma)_1\|\alpha,\beta)$
$A.B_{Quadratic}$	$a-1$	$R((\alpha\gamma)_2\|\alpha,\beta,(\alpha\gamma)_1)$
Deviations	$(a-1)(b-3)$	$R(\alpha\beta\|\alpha,\beta,(\alpha\gamma)_1,(\alpha\gamma)_2)$
Residual	$(r-1)(ab-1)$	$D(\alpha\beta)$

In deciding between the various hypotheses one must take into account rules IX.1 and VI.7 about the degrees and marginality of terms in the models. The following strategy should be employed in determining which of the models is to be used to describe the data.

For Deviations

Only if the terms to which a term is marginal are not significant then, if $P(F \geq F_{calc}) \leq 0.05$, the evidence suggests that the null hypothesis be rejected and the term must be incorporated in the model. Deviations for B is marginal to Deviations for A.B so that if the latter is significant, the Deviations for B is not tested; indeed no further testing occurs as the maximal model has to be used to describe the data.

For $A.B_{Linear}$ and $A.B_{Quadratic}$

Only if the polynomial terms of higher degree are not significant then, if $P(F \geq F_{calc}) \leq 0.05$, the evidence suggests that the null hypothesis be rejected and the term be incorporated in the model. $A.B_{Linear}$ is of lower degree than $A.B_{Quadratic}$ so that if the latter is significant, $A.B_{Linear}$ is not tested.

For A, Linear for B, Quadratic for B

Only if the terms to which a term is marginal and the polynomial terms of higher degree are not significant then, if $P(F \geq F_{calc}) \leq 0.05$, the evidence suggests that the null hypothesis be rejected and the term be incorporated in the model. For example, for the Linear term for B, the Quadratic term for B is of higher degree and it is marginal to $A.B_{Linear}$ so that if either of these is significant, Linear for B is not tested.

Both factors quantitative

The maximal model for a two-factor factorial with both factors quantitative is the same as when there is only one qualitative factor. However, the following set of expectation models is considered when both factors are quantitative:

$E[\mathbf{Y}] = \mathbf{X}_{AB}(\alpha\beta)$	depends on combination of A and B
$E[\mathbf{Y}] = \mathbf{X}_A\alpha + \mathbf{X}_B\beta + \mathbf{X}^*_{22}\theta^*_{22}$	nonsmooth main effects,
$\left(\text{or some subset of } \mathbf{X}^*_{22}\theta^*_{22}\right)$	smooth interaction in A and B
$E[\mathbf{Y}] = \mathbf{X}_G\mu + \mathbf{X}_{22}\theta_{22}$	smooth response in A and B
$\left(\text{or some subset of } \mathbf{X}_{22}\theta_{22}\right)$	
$E[\mathbf{Y}] = \mathbf{X}_A\alpha + \mathbf{X}_B\beta$	nonsmooth, independent response to A & B
$E[\mathbf{Y}] = \mathbf{X}_A\alpha + +\mathbf{X}_{22}\theta^*_{22}$	smooth response for B, smooth interaction
$\left(\text{or some subset of } \mathbf{X}^*_{22}\theta^*_{22}\right)$	in A and B, intercept differs for A
$E[\mathbf{Y}] = \mathbf{X}_A\alpha + \left(\mathbf{X}_{02}\theta_{02} \text{ or } \mathbf{X}_{01}\gamma_{01}\right)$	smooth response for B, intercept differs for A
$E[\mathbf{Y}] = \mathbf{X}_A\alpha$	nonsmooth response, depends on A only
$E[\mathbf{Y}] = \mathbf{X}_{22}\theta^*_{22} + \mathbf{X}_B\beta$	smooth response for A, smooth interaction
$\left(\text{or some subset of } \mathbf{X}^*_{22}\theta^*_{22}\right)$	in A and B, intercept differs for B
$E[\mathbf{Y}] = \left(\mathbf{X}_{20}\theta_{20} \text{ or } \mathbf{X}_{10}\gamma_{10}\right) + \mathbf{X}_B\beta$	smooth response for A, intercept differs for B
$E[\mathbf{Y}] = \mathbf{X}_B\beta$	nonsmooth response, depends on B only

where $(\alpha\beta) = \{(\alpha\beta)_{ij}\}$

$\alpha = \{\alpha_i\}$

$\beta = \{\beta_j\}$

$\theta'_{22} = [\gamma_{10} \quad \gamma_{20} \quad \gamma_{01} \quad \gamma_{02} \quad \gamma_{11} \quad \gamma_{12} \quad \gamma_{21} \quad \gamma_{22}]$

$\theta'^*_{22} = [\gamma_{11} \quad \gamma_{12} \quad \gamma_{21} \quad \gamma_{22}], [\gamma_{01} \quad \gamma_{02} \quad \gamma_{11} \quad \gamma_{12} \quad \gamma_{21} \quad \gamma_{22}]$ or

$[\gamma_{10} \quad \gamma_{20} \quad \gamma_{11} \quad \gamma_{12} \quad \gamma_{21} \quad \gamma_{22}]$

$\theta'_{20} = [\gamma_{10} \quad \gamma_{20}]$

$\theta'_{02} = [\gamma_{01} \quad \gamma_{02}]$

Again, rules IX.1 and VI.7 were used in deriving this set of models. Note that the version of θ'^*_{22} that applies to models involving α or β is the one that excludes coefficients whose terms are marginal to the α or β terms in the model. Also, the subsets of terms from $\mathbf{X}_{22}\theta_{22}$ mentioned above include the null subset and must

conform to rule IX.1 so that whenever a term from $\mathbf{X}_{22}\theta_{22}$ is added to the subset, all terms of lower degree must also be included in the subset.

The expression for the maximal polynomial submodel $E[\mathbf{Y}] = \mathbf{X}_G\mu + \mathbf{X}_{22}\theta_{22}$, in terms of a single observation, is:

$$E[Y_k] = \mu + \gamma_{10}x_{\alpha_i} + \gamma_{20}x_{\alpha_i}^2 + \gamma_{01}x_{\beta_j} + \gamma_{02}x_{\beta_j}^2 + \gamma_{11}x_{\alpha_i}x_{\beta_j} + \gamma_{12}x_{\alpha_i}x_{\beta_j}^2 + \gamma_{21}x_{\alpha_i}^2x_{\beta_j} + +\gamma_{22}x_{\alpha_i}^2x_{\beta_j}^2$$

where Y_k is the random variable for the response variable for the *k*th plot,

μ is the overall level of the response variable in the experiment,

x_{α_i} is the value of the *i*th level of factor A,

x_{β_j} is the value of the *j*th level of factor B,

γs are the coefficients of the equation describing the change in response as the levels of A and/or B changes with the first subscript indicating the degree with respect to factor A and the second subscript indicating the degree with respect to factor B,

In interpreting the fitted models, the following observations apply (for more see Box, Hunter and Hunter, section 15.4)

- models in which there are only main effect terms define

  $\Rightarrow$ a plane if both terms linear
  $\Rightarrow$ a parabolic tunnel if one term is linear and the other quadratic
  $\Rightarrow$ a paraboloid if both involve quadratic terms

- models including interaction submodels define nonlinear surfaces

  $\Rightarrow$ they will be monotonic for factors involving only linear terms,

  $\Rightarrow$ for interactions involving quadratic terms, some candidate shapes are:

Alinear.Bquadratic Aquadratic.Bquadratic

Again, the sequence of models fitted here is restricted in that: terms of lower degree must be fitted before terms of higher degree; main effects must be fitted before interactions. Thus the terms can be fitted in the sequence that yields the following deviances: $D(\mu)$, $D(\mu,\gamma_{10})$, $D(\mu,\gamma_{10},\gamma_{20})$, $D(\alpha)$, $D(\alpha,\gamma_{01})$, $D(\alpha,\gamma_{01},\gamma_{02})$, $D(\alpha,\beta)$, $D(\alpha,\beta,\gamma_{11})$, $D(\alpha,\beta,\gamma_{11},\gamma_{12})$, $D(\alpha,\beta,\gamma_{11},\gamma_{12},\gamma_{21})$, $D(\alpha,\beta,\gamma_{11},\gamma_{12},\gamma_{21},\gamma_{22})$, $D(\alpha\beta)$

The analysis of variance table for a two-factor CRD with both factors quantitative is:

Source	df	SSq
Individuals	$n-1$	
A	$a-1$	$R(\alpha\|\mu)$
Linear	1	$R(\gamma_{10}\|\mu)$
Quadratic	1	$R(\gamma_{20}\|\mu,\gamma_{10})$
Deviations	$a-3$	$R(\alpha\|\mu,\gamma_{10},\gamma_{20})$
B	$b-1$	$R(\beta\|\mu)$
Linear	1	$R(\gamma_{01}\|\mu)$
Quadratic	1	$R(\gamma_{02}\|\mu,\gamma_{01})$
Deviations	$b-3$	$R(\beta\|\mu,\gamma_{01},\gamma_{02})$
A.B	$(a-1)(b-1)$	$R(\alpha\beta\|\alpha,\beta)$
$A_{Linear}.B_{Linear}$	1	$R(\gamma_{11}\|\alpha,\beta)$
$A_{Linear}.B_{Quadratic}$	1	$R(\gamma_{12}\|\alpha,\beta,\gamma_{11})$
$A_{Quadratic}.B_{Linear}$	1	$R(\gamma_{21}\|\alpha,\beta,\gamma_{11})$
$A_{Quadratic}.B_{Quadratic}$	1	$R(\gamma_{22}\|\alpha,\beta,\gamma_{11},\gamma_{12},\gamma_{21})$
Deviations	$(a-1)(b-1)-4$	$R(\alpha\beta\|\alpha,\beta,\gamma_{11},\gamma_{12},\gamma_{21},\gamma_{22})$
Residual	$(r-1)(ab-1)$	$D(\alpha\beta)$

Step 3: Decide between hypotheses

For Deviations

Only if the terms to which a term is marginal are not significant then, if $P(F \geq F_{calc}) \leq 0.05$, the evidence suggests that the null hypothesis be rejected and the term must be incorporated in the model. Deviations for A and B are marginal to Deviations for A.B so that if the latter is significant, neither the Deviations for A nor for B is tested; indeed no further testing occurs as the maximal model has to be used to describe the data.

For all Linear and Quadratic terms

Only if the polynomial terms of higher degree are not significant and the terms to which the term is marginal are not significant then, if $P(F \geq F_{calc}) \leq 0.05$, the evidence suggests that the null hypothesis be rejected; the term and all polynomial terms of lower degree must be incorporated in the model. For example, $A_{Linear}.B_{Linear}$ is marginal to A.B and is of lower degree than all

other polynomial interaction terms and so is not tested if any of them is significant.

In using Genstat to fit these submodels you just include a POLND function for each quantitative factor. For example A*POLND(B; 2) or POLND(A; 2)*POLND(B; 2).

IX.H Nested factorial structures

Nested factorial structures commonly arise when a control treatment is included or when an interaction can be described in terms of one cell being different to the others. In these situations one sets up a factor (One say) with two levels, one for the control treatment or the different cell and two for the other treatments or cells. A second factor (Treats say) is set up with the same number of levels as there are treatments or cells. Then the structure for these two factors is One/Treats so that the terms in the analysis are One + One.Treats. One compares the control or single cell with the mean of the others. One.Treats reflects the differences between the other treatments or cells. Note that this can also be achieved using an orthogonal contrast, but the output using the nested factors is more convenient.

More generally one set up a nested factorial structure so that in the analysis there is: a) a term that reflects the average differences between *g* groups; and b) a term that reflects the differences within groups or several terms each one of which reflects the differences within a group.

Example IX.6 Grafting experiment (Daniel, 1977, p.27)

For example, consider the following RCBD experiment involving two factors each at two levels. The response is the percent grafts that take.

		B	1		2	
		A	1	2	1	2
	I		64	23	30	*
	II		75	14	50	33
Block	III		76	12	41	17
	IV		73	33	25	10

The components of this experiment are:

1. Observational unit – a plot
2. Response variable – % Take
3. Unrandomized factors – Blocks, Plots
4. Randomized factors – A, B
5. Type of study – Two-factor RCBD

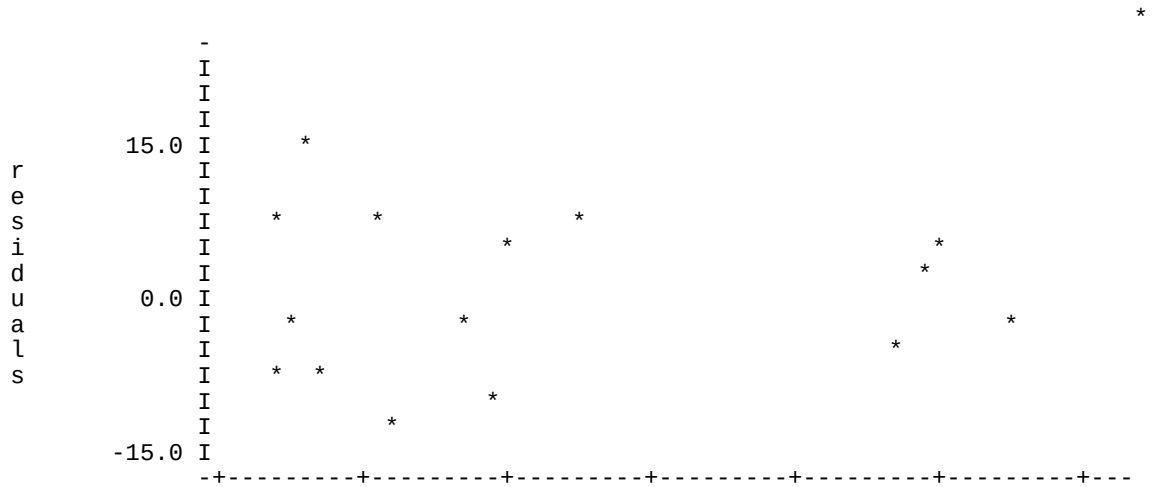
The experimental structure for this experiment is:

Structure	Formula
unrandomized	*4* Blocks/*4* Plots
randomized	*2* A**2* B

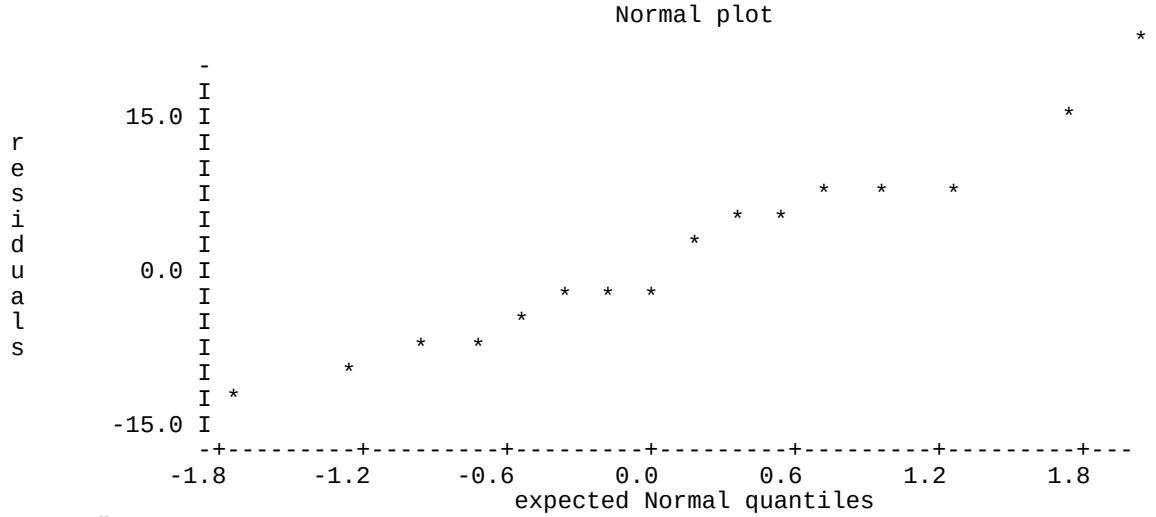
The following is the Genstat output contains the analysis of the data:

```
Genstat 5  Release 4.1  (PC/Windows NT)              03 April 2000 17:08:31
Copyright 1998, Lawes Agricultural Trust (Rothamsted Experimental Station)

                ____________________________________________________

                Genstat 5 Fourth Edition - (for Windows)
                Genstat 5 Procedure Library Release PL11
                ____________________________________________________

   3  "Data taken from File: D:/ANALYSES/LM/MULTIFAC/FAC2TAKE.GSH"
   4  DELETE [redefine=yes] Blocks,Plots,A,B,Take
   5  FACTOR [modify=yes;nvalues=16;levels=4] Blocks
   6  READ Blocks; frepresentation=ordinal

    Identifier    Values   Missing    Levels
        Blocks        16         0         4

   8  FACTOR [modify=yes;nvalues=16;levels=4] Plots
   9  READ Plots; frepresentation=ordinal

    Identifier    Values   Missing    Levels
         Plots        16         0         4

  11  FACTOR [modify=yes;nvalues=16;levels=2] A
  12  READ A; frepresentation=ordinal

    Identifier    Values   Missing    Levels
             A        16         0         2

  14  FACTOR [modify=yes;nvalues=16;levels=2] B
  15  READ B; frepresentation=ordinal

    Identifier    Values   Missing    Levels
             B        16         0         2

  17  VARIATE [nvalues=16] Take
  18  READ Take

    Identifier   Minimum      Mean   Maximum    Values   Missing
          Take     10.00     38.40     76.00        16         1

  20
  21  PRINT Blocks,Plots,A,B,Take

      Blocks       Plots           A           B        Take
           1           1           1           1       64.00
           1           2           2           1       23.00
           1           3           1           2       30.00
           1           4           2           2           *
           2           1           1           1       75.00
           2           2           2           1       14.00
           2           3           1           2       50.00
           2           4           2           2       33.00
           3           1           1           1       76.00
           3           2           2           1       12.00
           3           3           1           2       41.00
           3           4           2           2       17.00
           4           1           1           1       73.00
           4           2           2           1       33.00
           4           3           1           2       25.00
           4           4           2           2       10.00
```

```
  22  BLOCK Blocks/Plots
  23  TREAT A*B
  24  ANOVA [FPROB=Y; PSE=LSD] Take


24.....................................................................

***** Analysis of variance *****

Variate: Take

Source of variation     d.f.(m.v.)      s.s.       m.s.    v.r.  F pr.

Blocks stratum              3           223.8       74.6    0.73

Blocks.Plots stratum
A                           1          4807.1     4807.1   46.93  <.001
B                           1          1393.8     1393.8   13.61  0.006
A.B                         1          1133.4     1133.4   11.06  0.010
Residual                    8(1)        819.5      102.4

Total                      14(1)       7849.6


***** Tables of means *****

Variate: Take

Grand mean  36.9

        A        1        2
             54.2     19.6

        B        1        2
             46.2     27.6

        A        B        1        2
        1                72.0     36.5
        2                20.5     18.7


*** Least significant differences of means (5% level) ***

Table                     A           B           A
                                                  B
rep.                      8           8           4
d.f.                      8           8           8
l.s.d.                11.67       11.67       16.50

(Not adjusted for missing values)


***** Missing values *****

Variate: Take

 Unit  estimate
    4      14.7

Max. no. iterations 2

  25  APLOT METHOD=fit,normal
```



Normal plot




```
  26  "
 -27  **** Tukey''s one-degree-of-freedom-for-non-additivity.
 -28  **** It is the term designated covariate in the following analysis
 -29  "
  30  AKEEP [FIT=Fit]
  31  CALC ResSq=Fit*Fit
  32  ANOVA [PRINT=*]  ResSq; RES=ResSq
  33  COVAR ResSq                          "A computational trick"
  34  ANOVA [PRINT=A; FPROB=Y] Take

34.....................................................................

***** Analysis of variance (adjusted for covariate) *****

Variate: Take
Covariate: ResSq

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

Blocks stratum
Covariate                    1           85.3       85.3    1.23          0.383
Residual                     2          138.5       69.3    0.59    1.08

Blocks.Plots stratum
A                            1         4724.2     4724.2   40.50    0.99  <.001
B                            1         1392.7     1392.7   11.94    0.99  0.011
A.B                          1         1133.1     1133.1    9.71    0.99  0.017
Covariate                    1            3.0        3.0    0.03          0.876
Residual                     7(1)        816.5      116.6            0.88
```

```
Total                    14(1)       7849.6
```

The hypothesis test for this example is a follows:

Step 1: Set up hypotheses

a) H_0: $\alpha_1 = \alpha_2$
   H_1: $\alpha_1 = \alpha_2$

c) H_0: $\beta_1 = \beta_2$
   H_1: $\beta_1 = \beta_2$

c) H_0: $(\alpha\beta)_{21} - (\alpha\beta)_{11} - (\alpha\beta)_{22} + (\alpha\beta)_{12} = 0$
   H_1: $(\alpha\beta)_{21} - (\alpha\beta)_{11} - (\alpha\beta)_{22} + (\alpha\beta)_{12} \neq 0$

Step 2: Calculate test statistics

The analysis of variance table for the two-factor RCBD is:

Source	df	SSq	MSq	E[MSq]	F	Prob
Blocks	3	223.8	74.6	$\sigma^2_{BP} + 4\sigma^2_B$	0.73	
Blocks.Plots	12	8153.8				
A	1	4807.1	4807.1	$\sigma^2_{BP} + f_A(\psi)$	46.93	<0.001
B	1	1393.8	1393.8	$\sigma^2_{BP} + f_B(\psi)$	13.61	0.006
A.B	1	1133.4	1133.4	$\sigma^2_{BP} + f_{AB}(\psi)$	11.06	0.010
Residual	8†	819.5	102.4	σ^2_S		
Nonadditivity	1	3.0	3.0		0.03	0.876
Deviations	7	816.5	116.6			

†the Residual degrees of freedom have been reduced by one to allow for the missing observation

Step 3: Decide between hypotheses

Note that the plot of residuals-versus-fitted-values reveals nothing untoward, the test for nonadditivity is not significant and the normal probability plot also appears to be satisfactory. There is a significant interaction between A and B. The table of means for the combination of A and B needs to be examined.

		B		
		1	2	LSD(5%)
A	1	72.00	36.50	16.51
	2	20.50	18.75	

$$LSD(5\%) = 2.3060\sqrt{\frac{102.475\times 2}{4}}$$

It would appear that the results of this experiment are best described in terms of A and B both at level 1 is different from either A or B not at level 1. This can be achieved by setting up a factor for the 4 treatments and a two-level factor that compares the cell with A and B both at level 1 with the remaining factors. The four-level factor for treatments is then specified as nested within the two-level factor.

The re-analysis for these new factors is achieved in Genstat as follows:

```
 35  FACPRODUCT FACTORS=!p(A,B); PRODUCT=Treats
 36  FACTOR [LEV=2] Cell_1_1
 37  CALCULATE Cell_1_1=NEWLEVELS(Treats; !v(1,2,2,2))
 38  PRINT Treats,Cell_1_1

     Treats    Cell_1_1
          1           1
          3           2
          2           2
          4           2
          1           1
          3           2
          2           2
          4           2
          1           1
          3           2
          2           2
          4           2
          1           1
          3           2
          2           2
          4           2

 39  BLOCK Blocks/Plots
 40  TREAT Cell_1_1/Treats
 41  COVAR
 42  ANOVA [FPROB=Y; PSE=LSD] Take


42.....................................................................

***** Analysis of variance *****

Variate: Take

Source of variation     d.f.(m.v.)      s.s.       m.s.    v.r.  F pr.

Blocks stratum             3           223.8       74.6    0.73

Blocks.Plots stratum
Cell_1_1                   1          6564.5     6564.5   64.08  <.001
Cell_1_1.Treats            2           769.9      384.9    3.76  0.071
Residual                   8(1)        819.5      102.4

Total                     14(1)       7849.6


***** Tables of means *****

Variate: Take

Grand mean  36.9

 Cell_1_1        1        2
             72.0     25.2
     rep.        4       12
```

```
 Cell_1_1   Treats         1         2         3         4
        1               72.0
        2                         36.5      20.5      18.7


*** Least significant differences of means (5% level) ***

Table             Cell_1_1    Cell_1_1
                                Treats
rep.               unequal           4
d.f.                     8           8
l.s.d.               13.47       16.50

(Not adjusted for missing values)


***** Missing values *****

Variate: Take

 Unit  estimate
    4      14.7

Max. no. iterations 2
```

The revised analysis of variance table is:

Source	df	SSq	MSq	F	Prob
Blocks	3	221.0	74.6	0.73	
Blocks.Plots	12	8153.8			
Cell 1,1 vs rest	1	6564.5	6564.5	64.08	<0.001
Among rest	2	769.9	384.9	3.76	0.071
Residual	8†	819.5	102.4		
Total	14†	8362.9			

† the Residual and Total degrees of freedom have been reduced by one to allow for the missing observation

Clearly, the difference between the treatments is best summarized in terms of this single degree of freedom contrast between cell1,1 and the others. The mean for cell 1,1 is 72.0 and, for the other three treatments, the mean is 25.2, a difference of 46.8.

Such *one-cell interactions* are a very common form of interaction.

Example IX.7 Spraying sultanas

An experiment was conducted to investigate the effects of tractor speed and spray pressure on the quality of dried sultanas. The response was the lightness of the dried sultanas which is measured using a Hunterlab D25 L colour difference meter. Lighter sultanas are considered to be of better quality and these will have a higher lightness measurement (L). There were three tractor speeds and two spray pressures resulting in 6 treatment combinations which were applied to 6 plots, each consisting of 12 vines, using a randomized complete block design with three blocks. However, these 6 treatment combinations resulted in only 4 rates of spray application as indicated in the following table.

Table of application rates for the sprayer experiment

	Tractor Speed (km hour^{-1})		
Pressure (kPa)	3.6	2.6	1.8
140	**2090**	2930	**4120**
330	2930	**4120**	5770

To analyze this experiment we set up a factor, Rates, with four levels to compare the means of the four rates and two factors with three levels, Rate2 and Rate3, each of which compares the means of two treatment combinations with the same application rate. The levels of the factors Rate2 and Rate3 are shown in the following table.

Table of factor levels for Rate2 and Rate3 in the sprayer experiment

	Rate2			Rate3		
Tractor Speed (km hour^{-1})	3.6	2.6	1.8	3.6	2.6	1.8
Pressure (kPa)						
140	**1**	2	**1**	**1**	1	**2**
330	3	**1**	1	1	**3**	1

The experimental structure for this experiment is:

Structure	Formula
unrandomized	*3* Blocks/ *6* Plots
randomized	*4* Rates/(*3* Rate2+*3* Rate3)

The sources in the analysis of variance table are:

Source	df	E[MSq]
Blocks	2	$\sigma^2_{\mathrm{BP}} + \sigma^2_{\mathrm{B}}$
Blocks.Plots	15	
Rates	3	$\sigma^2_{\mathrm{BP}} + f_{\mathrm{R}}(\psi)$
Rates.Rate2	1	$\sigma^2_{\mathrm{BP}} + f_{\mathrm{R2}}(\psi)$
Rates.Rate3	1	$\sigma^2_{\mathrm{BP}} + f_{\mathrm{R3}}(\psi)$
Residual	10	σ^2_{BP}
Total	17	

IX.I Sample size

Sample size for a two-way factorial experiment can be computed using the Excel worksheet ANOVAPower.xls. It is a matter of adapting the procedures described in V.H, *Sample size*, to this situation. That is the cell's have to be set up as described in the following table:

Col-umn	Column heading	Modification to cell under heading
A	**sample size (r)**	enter a value
B	**alpha**	enter a value
C	**DF numerator**	enter a value
D	**DF denominator**	enter the formula '=a*(A5-b)' where a and b are constants that you work out for your proposed design
E	**central F**	leave this formula as it is '=FINV(B5,C5,D5)'
F	**no. values in a mean (m)**	enter the formula '=c*A5' where c is a constant that you work out for your proposed design
G	**delta**	enter a value
H	**standard deviation**	enter a value
I	**lambda**	leave this formula as it is '=F5*G5*G5/2/H5/H5'
J	**power**	leave this formula as it is '=1-NCF(E5,C5,D5,I5)'

The first step is to determine which effects you wish to specify to be determined with a nominated power: A main, B main and/or interaction effects. Then you will need to determine the power for each set of effects. As shown in the following table, the cells in this table that vary for the different types of effects are: DF numerator, no. values in a mean (m) and delta.

Source	DF numerator	m	Δ
A	$a-1$	br	$\alpha_i - \alpha_k$
B	$b-1$	ar	$\beta_j - \beta_\ell$
A.B	$(a-1)(b-1)$	r	$\left\{\left[(\alpha\beta)_{ij} - (\alpha\beta)_{i\ell}\right] - \left[(\alpha\beta)_{kj} - (\alpha\beta)_{k\ell}\right]\right\}$

where $i \neq k$ and $j \neq \ell$.

Note that for A and B, Δ is the difference between a pair of A or B means. For the A.B interaction, Δ is the difference between the differences of two pairs of means where these means involve only two levels of A and two levels of B as indicated in the above table — it is twice an interaction effect.

The DF denominator depends on the design employed in randomizing the treatments as given in the following table:

Design	DF denominator
CRD	$ab(r-1)$
RCBD	$(ab-1)(r-1)$
LS	$(r-1)(r-2)$

Example IX.1 Fertilizing oranges (continued)

Suppose that, for the experiment to investigate 3 levels of N and 2 levels of P, you wish to determine the number of blocks to use in an RCBD. When the difference between two levels of P is observed at each of two levels of N, you would like to be able to detect with 80% power, a change of more than 10 in the P difference. You believe that the standard deviation will be about 7.5 and you will use a significance level of 5%. How many replicates are required to achieve the desired power?

So we require the number of blocks in an RCBD to detect an interaction effect.

The entries in the Excel worksheet should be completed as follows:

Column	Column heading	Cell contents
A	**sample size (r)**	12
B	**alpha**	0.05
C	**DF numerator**	2
D	**DF denominator**	=5*(A5-1)
E	**central F**	=FINV(B5,C5,D5)
F	**no. values in a mean (m)**	=1*A5
G	**delta**	10
H	**standard deviation**	7.5
I	**lambda**	=F5*G5*G5/2/H5/H5
J	**power**	=1-NCF(E5,C5,D5,I5)

The completed worksheet has the following values:

sample size (r)	alpha	DF numer-ator	DF denomin-ator	central F	no. values in a mean (m)	delta	standard deviation	lambda	power
12	0.05	2	55	3.1650	12	10	7.500000	10.6667	0.8193

So 12 blocks needed to detect, with 80% power, a minimum change of 10, between two levels of N in the P difference.

IX.J Models and hypothesis testing for three-factor experiments

In this section we sketch the analysis of the general three factor experiment. That is an experiment with factors A, B and C with *a*, *b* and *c* levels, respectively, and each of the *abc* combinations of A, B and C replicated *r* times. That is, there will be $n = abcr$ observations. We do this in preparation for discussing experiments with more than two factors in the next section. The analysis is an extension of that for a two-factor CRD.

a) Using rules to determine the ANOVA table

The general components of a 3-factor CRD experiment are:

1. Observational unit – a unit
2. Response variable – Response
3. Unrandomized factors – Units
4. Randomized factors – A, B, C
5. Type of study – Three-factor CRD

The experimental structure for this experiment is:

Structure	Formula
unrandomized	*n* Units
randomized	*a* A**b* B**c* C

The terms derived from the randomized structure formula are:

$$\begin{aligned} A*B*C &= A + (B*C) + A.(B*C) \\ &= A + B + C + B.C + A.B + A.C + A.B.C \end{aligned}$$

The degrees of freedom of these terms can be derived by the cross product rule.

> For each factor in the term, calculate the number of levels minus one and multiply these together.

Given that the only random factor is Units, the following are the symbolic expressions for the maximal expectation and variation models:

$$E[Y] = \text{A.B.C}$$
$$\text{var}[Y] = \text{Units}$$

Hence the analysis of variance table with expected mean squares is:

Source	df	E[MSq]
Units	n–1	
A	a–1	$\sigma_U^2 + f_A(\psi)$
B	b–1	$\sigma_U^2 + f_B(\psi)$
A.B	(a–1)(b–1)	$\sigma_U^2 + f_{AB}(\psi)$
C	c–1	$\sigma_U^2 + f_C(\psi)$
A.C	(a–1)(c–1)	$\sigma_U^2 + f_{AC}(\psi)$
B.C	(b–1)(c–1)	$\sigma_U^2 + f_{BC}(\psi)$
A.B.C	(a–1)(b–1)(c–1)	$\sigma_U^2 + f_{ABC}(\psi)$
Residual	abc(r–1)	σ_U^2
Total	$abcr$–1	

b) Models

The models for the expectation that are considered for a three-factor factorial experiment are as follows:

$$E[\mathbf{Y}] = \mathbf{X}_{ABC}(\alpha\beta\gamma)$$

$$E[\mathbf{Y}] = \mathbf{X}_{AB}(\alpha\beta) + \mathbf{X}_{AC}(\alpha\gamma) + \mathbf{X}_{BC}(\beta\gamma)$$

$$E[\mathbf{Y}] = \mathbf{X}_{AB}(\alpha\beta) + \mathbf{X}_{AC}(\alpha\gamma) \quad \left(\begin{array}{l}\text{and equivalent models with a pair of}\\ \text{two-factors interactions}\end{array}\right)$$

$$E[\mathbf{Y}] = \mathbf{X}_{A}(\alpha) + \mathbf{X}_{BC}(\beta\gamma) \quad \left(\begin{array}{l}\text{and equivalent models with two factors}\\ \text{interacting and one factor independent}\end{array}\right)$$

$$E[\mathbf{y}] = \mathbf{X}_{A}(\alpha) + \mathbf{X}_{B}(\beta) + \mathbf{X}_{C}(\gamma) \quad \left(\begin{array}{l}\text{and other models consisting of}\\ \text{only main effects}\end{array}\right)$$

$$E[\mathbf{Y}] = \mathbf{X}_{G}\mu$$

where

$$(\alpha\beta\gamma) = \{(\alpha\beta\gamma)_{ijk}\}$$

$$(\alpha\beta) = \{(\alpha\beta)_{ij}\}$$

$$(\beta\gamma) = \{(\beta\gamma)_{jk}\}$$

$$(\alpha\gamma) = \{(\alpha\gamma)_{ik}\}$$

$$\alpha = \{\alpha_i\}$$

$$\beta = \{\beta_j\}$$

$$\gamma = \{\gamma_k\}$$

Altogether there are 18 different models that are to be considered. Expressions for the fitted values for each of the models would be derived and these would be given in terms of the vectors of means **A**, **B**, **A.B**, **C**, **A.C**, **B.C** and **A.B.C** where the means in the vector are those for each combination of the factors in the vector's name.

c) Sums of squares

As far as the hypothesis test is concerned we require the following reduction sums of squares to decide which model best describes the data:

$R(\alpha|\,\mu)$, $R(\beta|\,\mu)$, $R(\alpha\beta|\,\alpha,\beta)$, $R(\gamma|\,\mu)$, $R(\alpha\gamma|\,\alpha,\gamma)$, $R(\beta\gamma|\,\beta,\gamma)$ and $R(\alpha\beta\gamma|\,\alpha\beta,\alpha\gamma,\beta\gamma)$.

That is, only 7 models in addition to the model $E[\mathbf{Y}] = \mathbf{X}_G\mu$ have to be fitted to obtain the sums of squares for the analysis. However, there are many different orders in which terms can be added to the model to obtain these reduction sums of squares. The above set of reduction sums of squares gives the reduction for adding a term given that just the terms marginal to it are already in the model. This restriction of not fitting a term until all terms marginal to it have been fitted is required to get the correct reduction sums of squares. An equivalent set of reduction sums of squares, which corresponds to one possible order in which models might be fitted, is as follows:

$R(\alpha|\,\mu)$, $R(\beta|\,\alpha)$, $R(\alpha\beta|\,\alpha,\beta)$, $R(\gamma|\,\alpha\beta)$, $R(\alpha\gamma|\,\alpha\beta,\gamma)$, $R(\beta\gamma|\,\alpha\beta,\alpha\gamma)$ and $R(\alpha\beta\gamma|\,\alpha\beta,\alpha\gamma,\beta\gamma)$.

Of course, we also require $D(\mu)$ and $D(\alpha\beta\gamma)$ to obtain the Total and Residual sums of squares respectively.

As for previous orthogonal analyses this analysis is based on the decomposition of the data vector by successive application of mean and residual operators. Thus, the decomposition may be written:

$$\mathbf{y} = \mathbf{g} + \mathbf{a}_e + \mathbf{b}_e + \mathbf{a.b}_e + \mathbf{c}_e + \mathbf{a.c}_e + \mathbf{b.c}_e + \mathbf{a.b.c}_e + \mathbf{e}_{A.B.C}$$

Theorem IX.9: Let **Y** be a *n*-vector of jointly-distributed random variables and let

$$\begin{aligned}
\mathbf{P}_G &= \tfrac{1}{abcr}\ \mathbf{J}_a \otimes \mathbf{J}_b \otimes \mathbf{J}_c \otimes \mathbf{J}_r \\
\mathbf{P}_A &= \tfrac{1}{bcr}\ \mathbf{I}_a \otimes \mathbf{J}_b \otimes \mathbf{J}_c \otimes \mathbf{J}_r \\
\mathbf{P}_B &= \tfrac{1}{acr}\ \mathbf{J}_a \otimes \mathbf{I}_b \otimes \mathbf{J}_c \otimes \mathbf{J}_r \\
\mathbf{P}_{AB} &= \tfrac{1}{cr}\ \mathbf{I}_a \otimes \mathbf{I}_b \otimes \mathbf{J}_c \otimes \mathbf{J}_r \\
\mathbf{P}_C &= \tfrac{1}{abr}\ \mathbf{J}_a \otimes \mathbf{J}_b \otimes \mathbf{I}_c \otimes \mathbf{J}_r \\
\mathbf{P}_{AC} &= \tfrac{1}{br}\ \mathbf{I}_a \otimes \mathbf{J}_b \otimes \mathbf{I}_c \otimes \mathbf{J}_r \\
\mathbf{P}_{BC} &= \tfrac{1}{ar}\ \mathbf{J}_a \otimes \mathbf{I}_b \otimes \mathbf{I}_c \otimes \mathbf{J}_r \\
\mathbf{P}_{ABC} &= \tfrac{1}{r}\ \mathbf{I}_a \otimes \mathbf{I}_b \otimes \mathbf{I}_c \otimes \mathbf{J}_r
\end{aligned}$$

Then,

$$
\begin{aligned}
D(\mu) &= \mathbf{Y}'(\mathbf{I}-\mathbf{P}_{\mathrm{G}})\mathbf{Y} \\
R(\alpha|\,\mu) &= \mathbf{Y}'(\mathbf{P}_{\mathrm{A}}-\mathbf{P}_{\mathrm{G}})\mathbf{Y} \\
R(\beta|\,\mu) &= \mathbf{Y}'(\mathbf{P}_{\mathrm{B}}-\mathbf{P}_{\mathrm{G}})\mathbf{Y} \\
R(\alpha\beta|\,\alpha,\beta) &= \mathbf{Y}'(\mathbf{P}_{\mathrm{AB}}-\mathbf{P}_{\mathrm{A}}-\mathbf{P}_{\mathrm{B}}+\mathbf{P}_{\mathrm{G}})\mathbf{Y} \\
R(\gamma|\,\mu) &= \mathbf{Y}'(\mathbf{P}_{\mathrm{C}}-\mathbf{P}_{\mathrm{G}})\mathbf{Y} \\
R(\alpha\gamma|\,\alpha,\gamma) &= \mathbf{Y}'(\mathbf{P}_{\mathrm{AC}}-\mathbf{P}_{\mathrm{A}}-\mathbf{P}_{\mathrm{C}}+\mathbf{P}_{\mathrm{G}})\mathbf{Y} \\
R(\beta\gamma|\,\beta,\gamma) &= \mathbf{Y}'(\mathbf{P}_{\mathrm{BC}}-\mathbf{P}_{\mathrm{B}}-\mathbf{P}_{\mathrm{C}}+\mathbf{P}_{\mathrm{G}})\mathbf{Y} \\
R(\alpha\beta\gamma|\,\alpha\beta,\alpha\gamma,\beta\gamma) &= \mathbf{Y}'(\mathbf{P}_{\mathrm{ABC}}-\mathbf{P}_{\mathrm{AB}}-\mathbf{P}_{\mathrm{AC}}-\mathbf{P}_{\mathrm{BC}}+\mathbf{P}_{\mathrm{A}}+\mathbf{P}_{\mathrm{B}}+\mathbf{P}_{\mathrm{C}}-\mathbf{P}_{\mathrm{G}})\mathbf{Y} \\
D(\alpha\beta\delta) &= \mathbf{Y}'(\mathbf{I}-\mathbf{P}_{\mathrm{ABC}})\mathbf{Y}
\end{aligned}
$$

Proof: This proof is similar to previous proofs given for sums of squares. First, it would be necessary to establish the relationships between pairs of **P** matrices using the direct product expressions for them. Then, expressions for the reductions sums of squares would converted to differences between two deviances or residual sums of squares. Each deviance can then be written as the difference between the random vector **Y** and the fitted values for the model corresponding to the deviance so that the difference of two deviances would be the difference in two sets of fitted values. Since the fitted values are functions of vectors of means, they can be re-expressed in terms of **P** matrices to yield the expressions given above. Also, it would have to be demonstrated that the linear combination of **P** matrices in each reduction sums of squares is symmetric and idempotent ■

A further theorem would look at expressions for the sums of squares in terms of **R** and **P** operators and at the equivalence of various reduction sums of squares for fitting the same term, the different reduction sums of squares corresponding to different models fitted prior to adding the term of interest or different sequences in the recursive mean operator procedure.

d) Degrees of freedom

Theorem IX.10: Let D(μ), R(α| μ), R(β| μ), R($\alpha\beta$| α,β), R(γ| μ), R($\alpha\gamma$| α,γ), R($\beta\gamma$| β,γ), R($\alpha\beta\gamma$| $\alpha\beta,\alpha\gamma,\beta\gamma$) and D($\alpha\beta\gamma$) be as defined in theorem IX.6. Then the degrees of freedom of these sums of squares are n–1, a–1, b–1, (a–1)(b–1), c–1, (a–1)(c–1), (b–1)(c–1), (a–1)(b–1)(c–1) and (n–abc) = abc(r–1), respectively.

Proof: left as an exercise ■

e) Expected mean squares

Lemma IX.4: Let **Y** be a *n*-vector of jointly-distributed random variables with $E[\mathbf{Y}] = \psi$ and $\text{var}[\mathbf{Y}] = \sigma_U^2 \mathbf{I}_n$. Also, let **E** be symmetric and idempotent with $trace(\mathbf{E}) = \nu$. Then,

$$E[\mathbf{Y}'\mathbf{E}\mathbf{Y}/\nu] = \sigma_U^2 + f(\psi)$$

where $f(\psi) = \psi'\mathbf{E}\psi/\nu$

Proof: a straightforward application of theorem II.11. ■

Now, all the models under consideration are of the type described in lemma IX.4 and all of the sums of squares are of the form $\mathbf{Y}'\mathbf{E}\mathbf{Y}$. So, as for the two-factor case, a theorem on the expected mean squares would involve the specification of the form of $f(\psi)$ for each mean square under each model. It would state that the $f(\psi)$ for a particular mean square will only involve either the term corresponding to that mean square or a term to which it is marginal and will only be zero when none of these occurs in the model. From this it would be concluded that the $f(\psi)$ for a particular mean square will only go to zero when the term corresponding to it is dropped from the model. Consequently, the hypothesis test for a mean square is a test for whether the expectation term corresponding to it is zero, when it occurs in a model.

For example, consider the A.B mean square. Its expected value is

$$E\left[\mathbf{Y}'\left(\mathbf{P}_{AB} - \mathbf{P}_A - \mathbf{P}_B + \mathbf{P}_G\right)\mathbf{Y}/(a-1)(b-1)\right] = \sigma_U^2 + f_{AB}(\psi)$$

Now, $f_{AB}(\psi) \neq 0$ for models involving the AB term $\left[\mathbf{X}_{AB}(\alpha\beta)\right]$ or terms to which the AB term is marginal. That is, it is nonzero for the models

$$\begin{aligned}
E[\mathbf{Y}] &= \mathbf{X}_{ABC}(\alpha\beta\gamma) \\
E[\mathbf{Y}] &= \mathbf{X}_{AB}(\alpha\beta) + \mathbf{X}_{AC}(\alpha\gamma) + \mathbf{X}_{BC}(\beta\gamma) \\
E[\mathbf{Y}] &= \mathbf{X}_{AB}(\alpha\beta) + \mathbf{X}_{BC}(\beta\gamma) \\
E[\mathbf{Y}] &= \mathbf{X}_{AB}(\alpha\beta) + \mathbf{X}_{AC}(\alpha\gamma) \\
E[\mathbf{Y}] &= \mathbf{X}_{AB}(\alpha\beta) + \mathbf{X}_C\gamma \\
E[\mathbf{Y}] &= \mathbf{X}_{AB}(\alpha\beta)
\end{aligned}$$

It will only become zero when the AB term is dropped from the model. Dropping the ABC term in the first model leads to a consideration of the next simplest model which is one involving the terms AB, AC, and BC. Dropping terms other than AB from the

other models leaves AB in the model so that $f_{AB}(\psi)$ remains nonzero. It is only when the AB term is dropped that $f_{AB}(\psi)$ becomes zero.

But, $f_{AB}(\psi) = \psi'(\mathbf{P}_{AB} - \mathbf{P}_A - \mathbf{P}_B + \mathbf{P}_G)\psi / (a-1)(b-1)$ so that $f_{AB}(\psi)$ will be zero when $(\mathbf{P}_{AB} - \mathbf{P}_A - \mathbf{P}_B + \mathbf{P}_G)\psi$ is zero. However, it can be shown that for a model involving the AB term, and so not the A, B or ABC terms, an element of $(\mathbf{P}_{AB} - \mathbf{P}_A - \mathbf{P}_B + \mathbf{P}_G)\psi$ is equal to $\left((\alpha\beta)_{ij} - \overline{(\alpha\beta)}_{i.} - \overline{(\alpha\beta)}_{.j} + \overline{(\alpha\beta)}_{..}\right)$ where *i* is the level of A and B the level of *j* for that element. That is, these have to be zero for all *i* and *j* to indicate that the AB term can be omitted from the model.

f) Distribution of the F statistics

The ratio of a mean square for a treatment term to the residual mean square can be shown to be distributed as Snedecor's F.

g) The hypothesis test

The hypothesis test for choosing the model for a 3-factor experiment is as follows:

Step 1: Set up hypotheses

	Term being tested	
a) H_0: $\alpha_1 = \alpha_2 = ... = \alpha_a$ H_1: at least one pair of population A means is different	A	(T)
b) H_0: $\beta_1 = \beta_2 = ... = \beta_b$ H_1: at least one pair of population B means is different	B	(C)
c) H_0: $\left((\alpha\beta)_{ij} - \overline{(\alpha\beta)}_{i.} - \overline{(\alpha\beta)}_{.j} + \overline{(\alpha\beta)}_{..} = 0 \quad \text{for all i,j}\right)$ H_1: $\left((\alpha\beta)_{ij} - \overline{(\alpha\beta)}_{i.} - \overline{(\alpha\beta)}_{.j} + \overline{(\alpha\beta)}_{..} \neq 0 \quad \text{for some i,j}\right)$	A.B	(T.C)
d) H_0: $\gamma_1 = \gamma_2 = ... = \gamma_c$ H_1: at least one pair of population C means is different	C	(K)
e) H_0: $\left((\alpha\gamma)_{ik} - \overline{(\alpha\gamma)}_{i.} - \overline{(\alpha\gamma)}_{.k} + \overline{(\alpha\gamma)}_{..} = 0 \quad \text{for all i,k}\right)$ H_1: $\left((\alpha\gamma)_{ik} - \overline{(\alpha\gamma)}_{i.} - \overline{(\alpha\gamma)}_{.k} + \overline{(\alpha\gamma)}_{..} \neq 0 \quad \text{for some i,k}\right)$	A.C	(T.K)
f) H_0: $\left((\beta\gamma)_{jk} - \overline{(\beta\gamma)}_{j.} - \overline{(\beta\gamma)}_{.k} + \overline{(\beta\gamma)}_{..} = 0 \quad \text{for all j,k}\right)$ H_1: $\left((\beta\gamma)_{jk} - \overline{(\beta\gamma)}_{j.} - \overline{(\beta\gamma)}_{.k} + \overline{(\beta\gamma)}_{..} \neq 0 \quad \text{for some j,k}\right)$	B.C	(C.K)

g) H_0: $\left(\begin{array}{l} (\alpha\beta\gamma)_{ijk} - \overline{(\alpha\beta\gamma)}_{ij.} - \overline{(\alpha\beta\gamma)}_{i.k} - \overline{(\alpha\beta\gamma)}_{.jk} \\ + \overline{(\alpha\beta\gamma)}_{i..} + \overline{(\alpha\beta\gamma)}_{.j.} + \overline{(\alpha\beta\gamma)}_{..k} - \overline{(\alpha\beta\gamma)}_{...} = 0 \quad \text{for all i,j,k} \end{array} \right)$ A.B.C (T.C.K)

H_1: $\left(\begin{array}{l} (\alpha\beta\gamma)_{ijk} - \overline{(\alpha\beta\gamma)}_{ij.} - \overline{(\alpha\beta\gamma)}_{i.k} - \overline{(\alpha\beta\gamma)}_{.jk} \\ + \overline{(\alpha\beta\gamma)}_{i..} + \overline{(\alpha\beta\gamma)}_{.j.} + \overline{(\alpha\beta\gamma)}_{..k} - \overline{(\alpha\beta\gamma)}_{...} \neq 0 \quad \text{for some i,j,k} \end{array} \right)$

Step 2: Calculate test statistics

The form of the analysis of variance table for a three-factor factorial CRD is:

Source	df	E[MSq]
Units	n–1	
A	a–1	$\sigma_U^2 + f_A(\psi)$
B	b–1	$\sigma_U^2 + f_B(\psi)$
A.B	(a–1)(b–1)	$\sigma_U^2 + f_{AB}(\psi)$
C	c–1	$\sigma_U^2 + f_C(\psi)$
A.C	(a–1)(c–1)	$\sigma_U^2 + f_{AC}(\psi)$
B.C	(b–1)(c–1)	$\sigma_U^2 + f_{BC}(\psi)$
A.B.C	(a–1)(b–1)(c–1)	$\sigma_U^2 + f_{ABC}(\psi)$
Residual	abc(r–1)	σ_U^2
Total	$abcr$–1	

Step 3: Decide between hypotheses

For A.B.C interaction

If $P(F \geq F_{calc}) \leq 0.05$, the evidence suggests that the null hypothesis be rejected and the term should be incorporated in the model.

For A.B, A.C and B.C interactions

Only if A.B.C is not significant, then if $P(F \geq F_{calc}) \leq 0.05$, the evidence suggests that the null hypothesis be rejected and the term should be incorporated in the model.

For A, B and C

For each term, only if the interactions involving the term are not significant, then if $P(F \geq F_{calc}) \leq 0.05$, the evidence suggests that the null hypothesis be rejected and the term should be incorporated in the model.