5 Designing experiments—keeping it simple

5.1 Three fundamental principles of experimental design

The concept of blocking in experimental design is introduced. To analyse a blocked experiment, the categorical variable BLOCK is included in the model formula. In this example, we compare two analyses: one in which BLOCK has been statistically eliminated, and one in which it has not.

MINITAB COMMANDS FOR BOX 5.1

Commands
```
glm YIELD = BEAN;
brief 1.
```

Menu route

Stat > ANOVA > General Linear Model
 YIELD → Response
 BEAN → Model
 Results...
 ⊙ Analysis of variance table

MINITAB OUTPUT FOR BOX 5.1 Analysis of bean yields assuming a fully randomised design

General Linear Model: YIELD versus BEAN

```
Factor            Type Levels Values
BEAN             fixed      6  1 2 3 4 5 6

Analysis of Variance for YIELD, using Adjusted SS for Tests

Source          DF       Seq SS      Adj SS      Adj MS        F        P
BEAN             5      444.435     444.435      88.887    14.59    0.000
Error           18      109.690     109.690       6.094
Total           23      554.125
```

MINITAB COMMANDS FOR BOX 5.2 Analysis of a blocked experiment

Commands

```
glm YIELD = BLOCK + BEAN;
brief 1.
```

Menu route

Stat > ANOVA > General Linear Model
    `YIELD` → Response
    `BLOCK + BEAN` → Model
    [Results…]
        ⊙ Analysis of variance table

MINITAB OUTPUT FOR BOX 5.2 Analysis of bean yields assuming a randomised block design

General Linear Model: `YIELD` versus `BLOCK, BEAN`

```
Factor            Type Levels Values
BLOCK            fixed      4 1 2 3 4
BEAN             fixed      6 1 2 3 4 5 6

Analysis of Variance for YIELD, using Adjusted SS for Tests

Source           DF        Seq SS        Adj SS      Adj MS        F          P
BLOCK             3        52.895        52.895      17.632     4.66      0.017
BEAN              5       444.435       444.435      88.887    23.48      0.000
Error            15        56.795        56.795       3.786
Total            23       554.125
```

A Latin Square design is an example of using two blocking factors in an experimental design. Both categorical blocking variables are added to the model formula. The following analysis uses the *oilseed rape* dataset.

MINITAB COMMANDS FOR BOX 5.3 Analysing a Latin Square design

Commands

```
glm SEEDS = COLUMN + ROW + TREATMT;
brief 1.
```

Menu route

Stat > Anova > General Linear Model
    `SEEDS` → Response
    `COLUMN + ROW + TREATMT` → Model
    [Results…]
        ⊙ Analysis of variance table

MINITAB OUTPUT FOR BOX 5.3 Analysis of a Latin Square design

General Linear Model: SEEDS versus COLUMN, ROW, TREATMT

```
Factor          Type Levels Values
COLUMN         fixed      4 1 2 3 4
ROW            fixed      4 1 2 3 4
TREATMT        fixed      4 1 2 3 4

Analysis of Variance for SEEDS, using Adjusted SS for Tests

Source         DF     Seq SS     Adj SS     Adj MS       F       P
COLUMN          3    1332.25    1332.25     444.08    8.79   0.013
ROW             3    1090.25    1090.25     363.42    7.20   0.021
TREAMT          3    2650.25    2650.25     883.42   17.49   0.002
Error           6     303.00     303.00      50.50
Total          15    5375.75
```

5.2 The geometrical analogy for blocking

Calculating the fitted model for two categorical variables

Revisiting the *Beans* dataset, the full coefficient table can be obtained by altering the level of output requested (see Minitab supplement Chapter 3 and below).

MINITAB COMMANDS FOR BOX 5.4 Coefficients table for two categorical variables

Commands

```
glm YIELD = BLOCK + BEAN;
brief 3.
```

Menu route

Stat > ANOVA > General Linear Model
YIELD → Response
BLOCK + BEAN → Model
[Results...]
⊙ In addition, coefficients for all terms

This would give the following ANOVA and coefficients table:

MINITAB OUTPUT FOR BOX 5.4 **ANOVA and coefficient tables for a randomised blocked design**

General Linear Model: YIELD versus BLOCK, BEAN

```
Factor         Type Levels Values
BLOCK         fixed      4 1 2 3 4
BEAN          fixed      6 1 2 3 4 5 6

Analysis of Variance for YIELD, using Adjusted SS for Tests

Source           DF      Seq SS     Adj SS     Adj MS        F        P
BLOCK             3      52.895     52.895     17.632     4.66    0.017
BEAN              5     444.435    444.435     88.887    23.48    0.000
Error            15      56.795     56.795      3.786
Total            23     554.125

Term                 Coef    SE Coef          T          P
Constant          16.6750     0.3972      41.98      0.000
BLOCK
1                  0.0417     0.6880       0.06      0.953
2                  2.3917     0.6880       3.48      0.003
3                 -1.4750     0.6880      -2.14      0.049
BEAN
1                  5.0750     0.8882       5.71      0.000
2                  5.7000     0.8882       6.42      0.000
3                 -0.6000     0.8882      -0.68      0.510
4                 -0.2500     0.8882      -0.28      0.782
5                 -3.7000     0.8882      -4.17      0.001

Unusual Observations for YIELD

Obs         YIELD          Fit       SE Fit        Residual      St Resid
24        18.3000      21.4167       1.1916         -3.1167        -2.03R

R denotes an observation with a large standardized residual
```

There are four blocks and five varieties of bean, yet coefficients are given for only three levels of block and five levels of bean. This is because the final level of each categorical variable is aliased (as first discussed in Chapter 3). The aliasing convention followed in the main text is the same as that followed by Minitab, and so the full set of fitted values may be calculated as described in the main text.

5.3 The concept of orthogonality

Loss of orthogonality leads to differences in adjusted and sequential SS for a variable. This may be illustrated by the *Beans* dataset in which the variables

MYIELD, MBLOCK, and MBEAN are shortened versions of the original variables, due to loss of orthogonality.

MINITAB COMMANDS FOR BOX 5.5 Loss of orthogonality for two categorical variables

Commands

```
glm MYIELD = MBLOCK + MBEAN;
brief 3.
```

Menu route

Stat > ANOVA > General Linear Model
MYIELD → Response
MBLOCK + MBEAN → Model
Results...
⊙ In addition, coefficients for all termsz

MINITAB OUTPUT FOR BOX 5.5 Slight loss of orthogonality in a randomised block design

General Linear Model: MYIELD versus MBLOCK, MBEAN

```
Factor         Type Levels Values
MBLOCK        fixed      4 1 2 3 4
MBEAN         fixed      6 1 2 3 4 5 6

Analysis of Variance for MYIELD, using Adjusted SS for Tests

Source          DF      Seq SS      Adj SS      Adj MS          F         P
MBLOCK           3      49.291      49.690      16.563       6.47     0.006
MBEAN            5     449.413     449.413      89.883      35.09     0.000
Error           13      33.296      33.296       2.561
Total           21     532.000

Term                Coef       SE Coef        T           P
Constant         16.7007        0.3529     47.33       0.000
MBLOCK
1                 0.0160        0.5813      0.03       0.978
2                 2.3660        0.5813      4.07       0.001
3                -1.5007        0.5813     -2.58       0.023
MBEAN
1                 5.0493        0.7425      6.80       0.000
2                 6.7389        0.8435      7.99       0.000
3                -0.6257        0.7425     -0.84       0.415
4                -0.2757        0.7425     -0.37       0.716
5                -3.7257        0.7425     -5.02       0.000
```

5.5 Exercises

Growing carnations

The first exercise uses the *blooms* data set

MINITAB COMMANDS FOR BOX 5.6 Analysis of the number of tulip blooms with bed, water, and shade

Commands

```
glm SQBLOOMS = BED + WATER + SHADE;
brief 3.
```

Menu route

Stat > ANOVA > General Linear Model
SQBLOOMS → Response
BED + WATER + SHADE → Model
Results...
⊙ In addition, coefficients for all terms

MINITAB OUTPUT FOR BOX 5.6 Analysis of the number of tulip blooms with bed, water, and shade

General Linear Model: SQBLOOMS versus BED, WATER, SHADE

```
Factor         Type Levels Values
BED           fixed      3 1 2 3
WATER         fixed      3 1 2 3
SHADE         fixed      4 1 2 3 4

Analysis of Variance for SQBLOOMS, using Adjusted SS for Tests

Source        DF         Seq SS      Adj SS      Adj MS        F         P
BED            2         4.1323      4.1323      2.0661     9.46     0.001
WATER          2         3.7153      3.7153      1.8577     8.50     0.001
SHADE          3         1.6465      1.6465      0.5488     2.51     0.079
Error         28         6.1173      6.1173      0.2185
Total         35        15.6114

Term                Coef       SE Coef         T         P
Constant         4.02903       0.07790     51.72     0.000
BED
1                 0.0620        0.1102      0.56     0.578
2                 0.3805        0.1102      3.45     0.002
WATER
1                -0.4110        0.1102     -3.73     0.001
2                 0.3731        0.1102      3.39     0.002
SHADE
1                 0.0965        0.1349      0.72     0.480
2                 0.2934        0.1349      2.17     0.038
3                -0.1191        0.1349     -0.88     0.385
```

```
Unusual Observations for SQBLOOMS

Obs         SQBLOOMs          Fit        SE Fit       Residual      St Resid
 32          2.44900      3.68892       0.22034       -1.23992        -3.01R

R denotes an observation with a large standardized residual
```

The second analysis does not include BED as a block.

MINITAB COMMANDS FOR BOX 5.7 The carnation bloom analysis without bed used as a block

Commands

```
glm SQBLOOMS = WATER + SHADE;
brief 3.
```

Menu route

Stat > ANOVA > General Linear Model
SQBLOOMS → Response
BED + WATER + SHADE → Model
Results...
⊙ In addition, coefficients for all terms

MINITAB OUTPUT FOR BOX 5.7 The carnation bloom analysis without bed used as a block

General Linear Model: SQBLOOMS versus WATER, SHADE

```
Factor          Type Levels Values
WATER          fixed      3 1 2 3
SHADE          fixed      4 1 2 3 4

Analysis of Variance for SQBLOOMS, using Adjusted SS for Tests

Source            DF        Seq SS        Adj SS       Adj MS         F         P
WATER              2        3.7153        3.7153       1.8577      5.44     0.010
SHADE              3        1.6465        1.6465       0.5488      1.61     0.209
Error             30       10.2496       10.2496       0.3417
Total             35       15.6114

Term                 Coef       SE Coef          T          P
Constant          4.02903       0.09742      41.36      0.000
WATER
1                 -0.4110        0.1378      -2.98      0.006
2                  0.3731        0.1378       2.71      0.011
SHADE
1                  0.0965        0.1687       0.57      0.572
2                  0.2934        0.1687       1.74      0.092
3                 -0.1191        0.1687      -0.71      0.486
```

```
Unusual Observation for SQBLOOMS

Obs        SQBLOOMS        Fit     SE Fit    Residual    St Resid
 19         5.38500    4.28303    0.23863     1.10197       2.07R
 32         2.44900    4.13136    0.23863    -1.68236      -3.15R

R denotes an observation with a large standardized residual
```

In the third analysis, three plots have been removed.

MINITAB COMMANDS FOR BOX 5.8 Analysis of the carnation blooms with three plot values removed

Commands

```
glm SQ2 = B2 + W2 + S2;
brief 1.
```

Menu route

Stat > ANOVA > General Linear Model
`SQ2` → Response
`B2 + W2 + S2` → Model
[Results]
⊙ Analysis of variance table

MINITAB OUTPUT FOR BOX 5.8 Analysis of the carnation blooms with three plot values removed

General Linear Model: SQ2 versus B2, W2, S2

```
Factor          Type Levels Values
B2             fixed      3 1 2 3
W2             fixed      3 1 2 3
S2             fixed      4 1 2 3 4

Analysis of Variance for SQ2, using Adjusted SS for Tests

Source          DF      Seq SS      Adj SS     Adj MS       F       P
B2               2      2.7626      2.6490     1.3245    8.03   0.002
W2               2      5.0793      4.6764     2.3382   14.18   0.000
S2               3      0.8072      0.8072     0.2691    1.63   0.207
Error           25      4.1213      4.1213     0.1649
Total           32     12.7704
```

The dorsal crest of the male smooth newt

See Minitab output for this exercise in the answers to exercises in Chapter 14.