**Solution to Exercise 3.1** (Version 1, 22/09/14)

**from Statistical Methods in Biology: Design & Analysis of Experiments and Regression (2014) S.J. Welham, S.A. Gezan, S.J. Clark & A. Mead. Chapman & Hall/CRC Press, Boca Raton, Florida. ISBN: 978-1-4398-0878-8**

**Exercise 3.1**

Suppose that you are planning an experiment to investigate the impact of nutrient deprivation on plant metabolites. You have four different nutrient levels to test, obtained by applying appropriate nutrients to four sub-samples from a single bag of base compost. The resulting volume of each nutrient level is sufficient for four seed trays (i.e. 16 seed trays in total), and six plants will be grown in each seed tray. To achieve the required growing conditions, a small controlled environment (CE) cabinet will be used. The cabinet has four shelves, and you can fit four seed trays on each shelf in a $2 \times 2$ arrangement (Figure 3.7). Although the cabinet is supposed to provide a uniform environment, a technician suggests that light levels may vary between the shelves, and that this might affect plant growth. When they reach the required growth stage, the six plants from each seed tray will be bulked and processed together to form a single sample to be read by a machine. Your colleague tells you that the machine shows some drift over time, but that readings should stay stable across a set of up to six samples.

How might you design this experiment to obtain an unbiased assessment of differences between the four nutrient levels? Consider and discuss the different factors which might affect your choice of design and produce a candidate design. You should consider both stages of the experiment and the following issues

- What is the experimental unit for the nutrient treatments?
- What are the sources of heterogeneity in the experimental process?
- How might you deal with this heterogeneity?
- How would you allocate the treatments to the experimental units?
- What replication do you have for each treatment?
- What are the advantages/disadvantages of your design?

How would you modify your design if

a) a temperature gradient was discovered between the front and back of the shelves;
b) you want to include a $CO_2$ treatment that can only be applied to a whole CE cabinet and you obtain sufficient resources for 32 trays (eight for each nutrient level).

**Solution 3.1**

As treatments (the different nutrient levels) are applied to seed trays, the seed tray is the experimental unit. The major sources of heterogeneity are expected to be the shelves in the CE cabinet and drift in the machine readings over time. We will deal with this heterogeneity by using it to define blocks in our design.

As a first step, we consider the allocation of treatments (applied to seed trays) to positions within the cabinet. As we expect differences between the shelves, the shelves are used as blocks, giving four blocks which can each hold four seed trays. If we place one tray of each treatment on each shelf,

then we will have four replicates of each treatment, and each treatment will be tested once in each block. As conditions within each shelf are assumed to be homogeneous, we can randomly allocate treatments to positions within shelves, using a different randomization for each shelf. The design for the first stage of the experiment is therefore a RCBD with shelves as blocks (see Section 3.3.2).

At the second stage of the experiment, we need to specify the order in which the 16 samples will be read by the machine. We have been told that machine readings should be stable for sets of six readings or fewer. This implies that it should be stable over sets of four consecutive readings, which allows us to process samples for our each four treatments under similar (homogeneous) conditions. We will call a set of four machine readings a session, and use these sessions as an additional blocking factor. We will consider two different options for allocating samples to sessions, but there are many other valid schemes.

Option 1: Confound shelves with sessions

This option is simple: all the trays from a cabinet shelf (i.e. one replicate of each treatment) are processed in the same session. The order of processing trays within shelves/sessions is randomly determined and a separate randomization is used for each shelf/session. A block then represents a particular combination of one shelf from the cabinet and one session on the machine.

The only disadvantage of this design is that the effects of shelf and session cannot be separated. This is not important from the point of view of comparing treatments, but information on these sources of variation might be useful for designing further experiments.

Option 2: Use one tray from each shelf within each session

This option is more complicated because, to guard against bias, we want each treatment to occur once in each session and this is unlikely to occur if we simply select one tray at random from each shelf. Ideally, we constrain the randomization so that each treatment appears once on each shelf in the cabinet and once in each session. A possible allocation which fits this criterion is shown in Table S3.1.1 (with the treatments labelled as A to D). This design is a Latin square: a design with two crossed blocking factors (here shelves and sessions), each having number of levels equal to the number of treatments (see Section 3.3.3).

The advantage of a crossed blocking structure is that the contributions of shelves and sessions to variation in the data can be estimated independently. The major disadvantage in such a small experiment is that the separation of these sources of variation means that less information is available on the background variation used to assess treatment differences.

**Table S3.1.1** Allocation of treatments to shelves and sessions using a Latin square design.

| | | Machine session | | | |
| --- | --- | --- | --- | --- | --- |
| | | **1** | **2** | **3** | **4** |
| **CE cabinet shelf** | **1** | D | C | A | B |
| | **2** | A | D | B | C |
| | **3** | B | A | C | D |
| | **4** | C | B | D | A |

Modifications to the design:

a) to account for a temperature gradient between the front and back of the shelves

If there is a systematic difference between the fronts and backs of the shelves, then differences between backs and fronts should be used to define structure as well as the overall differences between the shelves, which we will label as the effect of shelf height. This is a crossed blocking structure (front-back * height). The complication is that now there are only two units within each block (front or back of a shelf), so it is not possible for every treatment to appear within each block.

Recall that our aim is to get precise unbiased comparisons of treatments. Previously, we have got unbiased estimates by randomization and by balancing treatments across blocks, and we have got precise estimates by comparing treatments within blocks. We will apply these same principles here: we would like each treatment to occur at the fronts and the backs of shelves equally often, and to be present on each shelf. For all comparisons to be estimated with equal precision, we require each treatment pair to appear together within blocks equally often. However, it is not possible to achieve both aims in the current situation.

Consider the two arrangements in Table S3.1.2. In the first arrangement, each treatment occurs twice at the front and twice at the back of the shelves. But not all treatment pairs occur together within blocks: AD and BC do not occur together and estimates of these treatment comparisons (A vs D and B vs C) are largely made across shelves and are therefore expected to be less precise than within-block comparisons. In the second arrangement, all treatment comparisons occur at least once within blocks (combinations AD and BC occur twice) but the treatments are not balanced across the front and back of the shelf.

The best design will depend on the amount of variation within the different levels of the blocking structure, but this is often not known beforehand. We usually expect less variation (more homogeneity) at lower levels of the structure and so tend to favour designs that balance treatment comparisons at these lower levels.

This sort of ad-hoc design can be produced for many situations where standard designs are not available. It is often required for laboratory testing (e.g. gels) where a large number of treatments are to be tested, but only a small number of gels can be run per day, and systematic differences between measurements from different days are observed. However, the popularity of standard designs is justified: they allow for straight-forward analysis and their properties are well-understood. Analysis of non-standard designs may be complicated, and produce treatment comparisons of different precisions (see Chapters 11 and 16).

**Table S3.1.2** Possible arrangements balancing treatments across front/back and within front or back of the shelves.

| | | Arrangement 1: Balanced between front and back | | | | Arrangement 2: Aiming for balance in within-block treatment comparisons | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Front | | Back | | Front | | Back | |
| Shelf | 1 | A | B | C | D | A | B | C | D |
| | 2 | A | C | B | D | A | C | B | D |
| | 3 | B | D | A | C | B | C | A | D |
| | 4 | C | D | A | B | A | D | B | C |

b) to include a $CO_2$ treatment that can only be applied to a whole CE cabinet with sufficient resources for 32 trays (eight for each nutrient level).

One possibility you might consider is to run the first stage of the experiment as a RCBD (as described above) with the cabinet set at one $CO_2$ level and then to run it again using the same cabinet (with different randomizations within each shelf) at a second $CO_2$ level, as illustrated in Figure S3.1.1. Or you might have considered using two cabinets at the same time (assuming two are available), one set at each $CO_2$ level. However, as the $CO_2$ treatment is applied to the whole cabinet in each case, neither option gives any real replication of the $CO_2$ treatments. It follows that we have no estimate of background variation due to the experimental runs (or cabinet) and so differences between the $CO_2$ levels cannot be statistically evaluated. If we could obtain more resources we might repeat either of these scenario several times to obtain real replication of the run (or cabinet) variation that we would use to test for $CO_2$ effects.

Working within the specified level of resources (maximum of 32 trays), a better solution from a statistical point of view would be to use four runs but use only two shelves in the cabinet (i.e. eight trays in the cabinet, set up as a RCBD with two shelves as blocks and four trays per shelf) and to repeat the experiment twice with each $CO_2$ treatment, with $CO_2$ treatments allocated to runs at random. This scenario is illustrated in Figure S3.1.2. This design provides some replication of the $CO_2$ treatments, albeit minimal, and valid statistical conclusions about the effects of this treatment can then potentially be drawn. From a statistical point of view it would be even better to use one shelf in the cabinet and repeat the experiment eight times (using a new randomization each time), i.e. four times at each $CO_2$ level. One criticism of these designs is that they are wasteful of resources: the cabinet is not full and is required for a longer period of time. If further material is available, then using all the shelves in the cabinet might give useful additional precision. However, repeating the experiment is not wasteful as it provides real replication of the $CO_2$ treatment that is essential to the statistical validity of the trial.

|       |   |       |   |
|-------|---|-------|---|
| Run 1: Cabinet at control $CO_2$ | | Run 2: Cabinet at treatment $CO_2$ | |

| C | B |
|---|---|
| A | D |
| C | A |
| D | B |
| A | B |
| D | C |
| C | D |
| B | A |

| C | A |
|---|---|
| B | D |
| C | B |
| A | D |
| D | A |
| C | B |
| B | D |
| C | A |

**Figure S3.1.1**. Design using two runs of 16 trays with cabinet at control level in first run and treatment level in second run (allocated at random): no true replication of $CO_2$ treatments.

**Figure S3.1.2**. Design using four runs of eight trays with cabinet set at control level of $CO_2$ in two runs and treatment level in two runs, using a random allocation of $CO_2$ levels to runs. This design has two true replicates of each $CO_2$ treatment.