12
Experimental Designs

Designing an effective experiment is of primary importance to apply the most proper statistical tool. Experimental design goes through two steps: i) defying the scientific hypothesis; ii) describing the research objective. The experiment will thus be aimed to verify the defined hypothesis measuring the phenomenon of interest in different conditions and then relying on inferential statistics. How can we explain what are experiments?

Experiments are characterized by the: i) manipulation of one or more independent variables; ii) use of controls such as randomly assigning participants or experimental units to one or more independent variables; and iii) careful observation or measurement of one or more dependent variables. The i) and ii) characteristics- and thus manipulation of an independent variable and the use of controls- distinguish experiments from other research strategies, such as observational studies.

An experiment carried out in a controlled environment has the aim of studying the effect of one or more factors (independent variables) affecting the variability of the dependent variable. Experimental designs allow obtaining valid results, albeit on a limited number of statistical units, by controlling the factors of variability and thus minimizing the experimental random error (linked to the nuisance that can occur in the complex reality of a phenomenon in nature). The design of an experiment involves some inter−related activities:

1. Formulation of statistical hypotheses in line with the scientific hypothesis. A statistical hypothesis is a statement about one or more parameters of a population. Statistical hypotheses are testable formulations of scientific hypotheses.
2. Determination of the treatment levels (independent variable) to be manipulated, the measurement to be recorded (dependent variable), and the extraneous conditions (random or nuisance variables) that must be controlled.
3. Specification of the number of experimental units required and the population from which they will be sampled.
4. Specification of the procedure for assigning the experimental units to the treatment levels.

In summary, an effective experimental design must be carried out identifying the independent, dependent, and nuisance variables and must indicate how the statistical aspects of an experiment are to be carried out. The primary objective of an experimental design is thus to establish a causal connection between the independent and dependent variables. A secondary objective is to extract as much information as possible with the minimum expenditure of resources (i.e. plants, animals, reagents, patients;

also because of a more sustainable scientific research and a reduction in animal testing).

Some useful notions to understand the design of an experiment. The experimental unit is the base unit (e.g. a patient, an animal, a plant, a land parcel, a bunch of cultured cells on a Petri dish) that we are assigning to a treatment. Experimental units are completely independent of each other. The effect of a treatment is measured on observational units, which in some cases may be the same as the experimental unit, while in others may be part of the experimental unit (e.g. each experimental unit is made of several observational units). For example, we have pigs in boxes and we are feeding each box with a different diet; in that case, pigs are observational units and each box is an experimental unit. Let's outline an example to better understand those notions: let's assume we want to test three different amounts of a fertilizer compound, and we have 3 plots (land parcels). The hypothesis is that variable amounts of fertilizer may cause differences in the amount of protein in the grass. The objective is to evaluate the effects of three different amounts of fertilizer compound on the grass protein content. Therefore:

- The dependent variable is the protein content in grass
- The independent variable is the fertilizer
- The levels of the independent variable are three and depend on the amount of fertilizer: low, medium and high amounts.
- The experimental unit is the plot (or land parcel) fertilized with a specific amount of fertilizer
- Replicates: are plots treated with the same amount of fertilizer

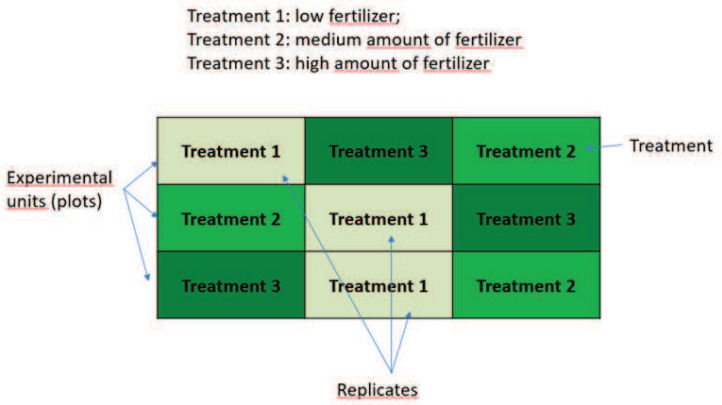

Fig. 12.1 Example of experimental design

The total variability of a biological phenomenon is made of a part of the variability that can be explained by the factors we take into account and by another part that however remains unexplained. the latter is for example the variability we observe when we are dealing with the replicates, and thus we are not able to understand. This unexplained part is commonly indicated as e, error. The error may be: i) systematic

if it is caused by constant effects, which may produce a "bias" in our experiment (for example, not well-calibrated instrumentation); ii) random when it is caused by unpredictable factors. The use of replicates allows to better estimate the variability related to these casual effects. Effective experimental designs should aim to avoid systematic errors, and reduce as much as possible random errors. How can we manage and control the systematic errors and estimate the amount of variability due to random errors? With different experimental designs.

12.1 Completely randomized design

More than an experimental design, randomization is a method that can be used to assign experimental units to treatments. Randomization involves a method for assigning experimental units randomly to the treatment levels. This process helps ensure that the groups are equivalent at the beginning of the study (and thus have the same probability of being assigned to each treatment or to the treatment and control levels). That makes it safer to assume the treatments caused any differences between groups that the experimenters observe at the end of the study. This type of experimental design is proper when we hypothesize that the variability of the dependent variable may be caused by a unique factor, and thus we must assign randomly our experimental units to a few treatments. The null hypothesis (H_0) is that there are no differences among the treatments. The statistical model used in this type of experimental design is a univariate model, with one independent variable (with different levels): one−way Analysis of Variance (ANOVA), which will be detailed in the next chapter.

$$y_{ij} = \mu + \alpha_i + e_{ij} \tag{12.1}$$

where: y_{ij} is the dependent variable measured on the experimental unit j treated with the $i - th$ treatment; μ is the mean of the whole sample; α_i is the effect of the ith treatment; e_{ij} is the vector of random errors. When we apply ANOVA, we, therefore, test whether the null hypothesis should be rejected (with a p−value below the significance threshold) or accepted (with a non−significant p−value). ANOVA may however be applied if the dependent variable is quantitative and normally distributed and the independent variable is made of more than two levels. When we are dealing with a quantitative normally distributed dependent variable and an independent variable with only two levels, we can compare the mean of the two groups using Student's t−test (if we assume that the variance is the same between the two groups) or Welch t−test (designed for comparing groups with unequal variance, but with a normal distribution).

If we are dealing with a dependent variable that is not normally distributed, we can use Mann−Whitney test (if the independent variable has only two levels), or Kruskal−Wallis test (if the independent variable has more than two levels). If both the dependent and independent variables are qualitative, we can use χ-square test (if we have at least 5 subjects in each cell and are not repeated measures on the same subjects), or Fisher test (if we have some cells with less than 5 subjects).

12.2 Randomized block design

When the experimental units are not homogeneous, but there is a certain variability for some relevant characteristics, we will be able to divide the subjects according to this characteristic. For example, let's assume we have a field with an increased gradient of fertility from the left to the right part of the field. If we are assigning randomly the 8 treatments to the replicates, it is possible that by chance the replicates of a treatment fall all on the same part of the field (for example treatment D is on the right part of the field, Figure 12.2).

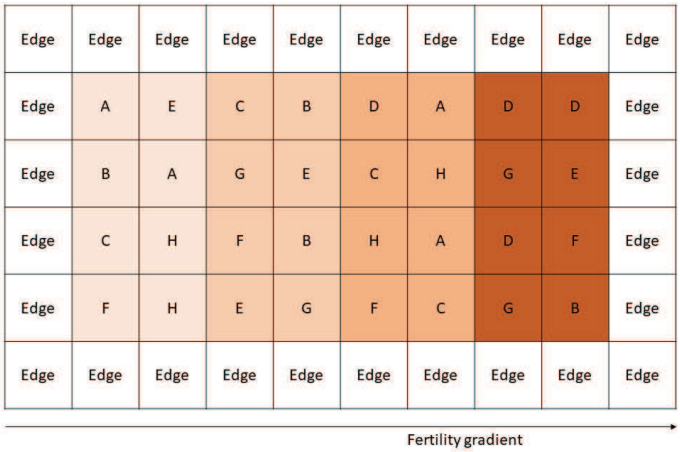

Fig. 12.2 Example of random experimental design

To avoid that, it is essential to use randomized block design, in order to assign randomly the treatments inside each block (Figure 12.2). Note that now the treatment D is represented in each part of the field.

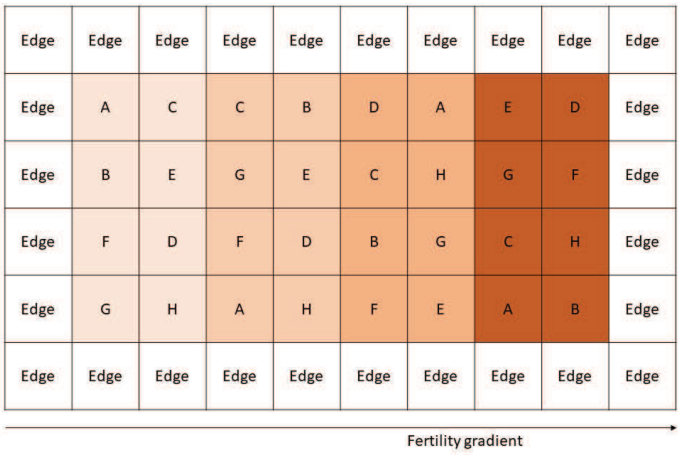

Fig. 12.3 Example of randomized block design

In statistics, the term "block" is used to identify a known source of variability which is not the variable objective of the experiment but may have a strong effect on the dependent variable. The experimental design must account also for this "secondary" variable, thus controlling possible sources of variability and limiting the experimental error. Blocks are groups of subjects homogeneous for that particular characteristic (for example the gender, the genotype for a genetic locus, the breed or variety, the soil characteristic). With a randomized block design, the first step we must perform is to divide subjects into groups called blocks. Then, subjects within each block are randomly assigned to treatment conditions. Compared to a completely randomized design, this design reduces variability within treatment conditions and potential confounding, producing a better estimate of treatment effects. This type of experimental design is the most effective to test the effects of two different drugs on male and female individuals (we divide our sample into males and females and then randomly assign them to the two treatments). The null hypothesis is that there are no differences among the treatments. The type of analysis we can use is the two−way ANOVA:

$$y_{ijk} = \mu + \alpha_i + \beta_j + e_{ijk} \tag{12.2}$$

where: y_{ij} is the dependent variable measured on the experimental unit k treated with the ith treatment in the block j; μ is the mean of the whole sample; α_i is the effect of the ith treatment; β_j is the effect of the block j; e_{ijk} is the vector of random errors. The limitations of this type of experimental design are we are not able to add more than two independent variables in the model, and it is not possible to compare variables with lots of levels.

12.3 Factorial experimental design

Factorial experimental designs are applied when we are interested in comparing the means of the levels of two or more independent variables and their interaction (Figure 12.3). In the factorial design we: i) test the effects of the independent variables taken individually; ii) test the combined effects of the independent variables (interaction).

Variable A	Variable B	
	B1	B2
A1	A1B1	A1B2
A2	A2B1	A2B2
A3	A3B1	A3B2

Fig. 12.4 Example of factorial experimental design

The statistical model applied to a bi-factorial design (2 variables) is:

$$y_{ijk} = \mu + \alpha_i + \beta_j + \alpha_i\beta_j + e_{ijk} \tag{12.3}$$

where: yij is the dependent variable measured on the experimental unit k subjected to the ith level of the α variable and the jth level of the variable β; μ is the mean of the whole sample; α_i is the effect of the ith level of the variable α; β_j is the effect of the jth level of the variable β; $\alpha_i\beta_j$ is the interaction effect between the ith level of the variable α and the jth level of the variable β; e_{ijk} is the vector of random errors. This type of experimental design has three null hypotheses: there is no significant interaction between the variables; the independent variable A has no significant effect on the dependent variable; the independent variable B has no significant effect on the dependent variable.

12.3.1 Split−plot

The split−plot design is an experimental design that is used when a factorial treatment structure has two levels of experimental units. In the case of the split−plot design, two levels of randomization are applied to assign experimental units to treatments1. The first level of randomization is applied to the whole plot and is used to assign experimental units to levels of treatment factor A. The whole plot is split into subplots, and the second level of randomization is used to assign the subplot experimental units to levels of treatment factor B. Since the split−plot design has two levels of experimental units, the whole plot and subplot portions have separate experimental errors. In the example below (Figure 12.3.1) eight plots each of which is split into three subplots receiving three levels of the treatment (G1, G2 and G3).

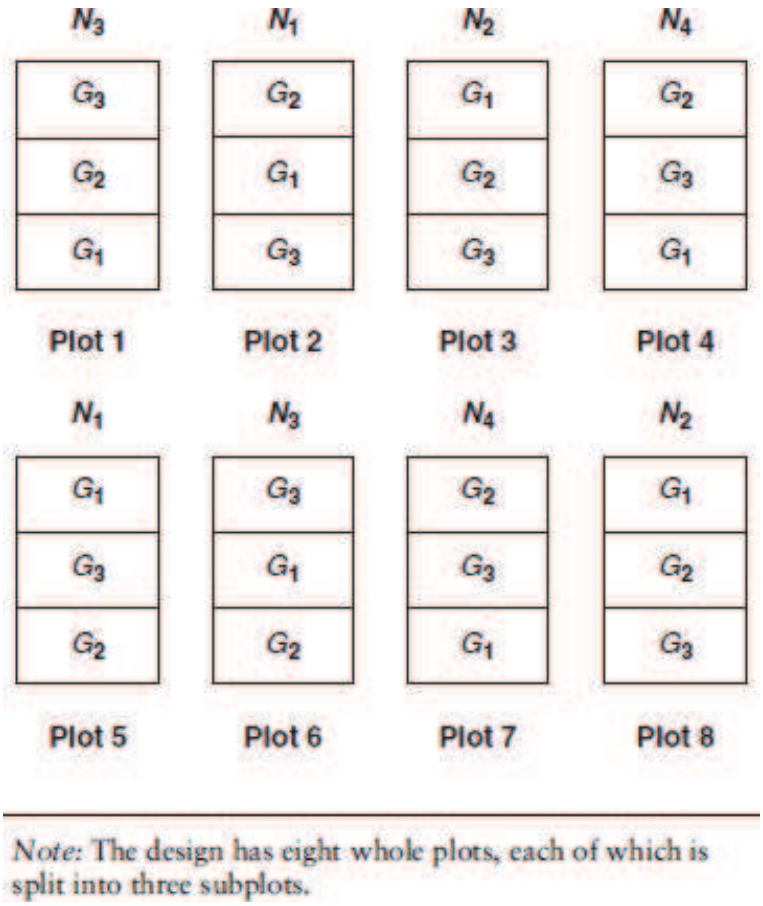

Fig. 12.5 Example of split−plot factorial experimental design

12.4 Latin square design

In some experimental situations, it is not possible to have a high number of experimental units. A good way to overcome those problems without losing statistical power may be to use a Latin Square experimental design. This type of experimental design is highly efficient and allows for two blocking factors. In other words, this design is used to simultaneously control (or eliminate) two sources of nuisance variability. Latin square design represents a particular case of randomized block design where treatments are assigned to each experimental unit following a square matrix scheme: each treatment (a level of the main factor) is present in each row and in each column. This type of experimental design requires a lower number of experimental units compared with randomized block design (and thus it's less expensive), it's more efficient than other types of experimental design. Latin square design does not permit however to estimate the interaction between the independent variables.

In this case, for application to field crops, the experimental units have two "gradients": in addition to the experimental treatment, there are differences linked to two important elements (for example fertility gradient). Figure 12.4 shows an experiment with four treatments and as many replicates, in which each treatment is found in all rows and all columns, in order to consider any fertility gradients from the right to the left and from the top to the bottom of the field. The figure also shows why we speak of Latin square; the number of rows is equal to the number of columns, according to a square grid.

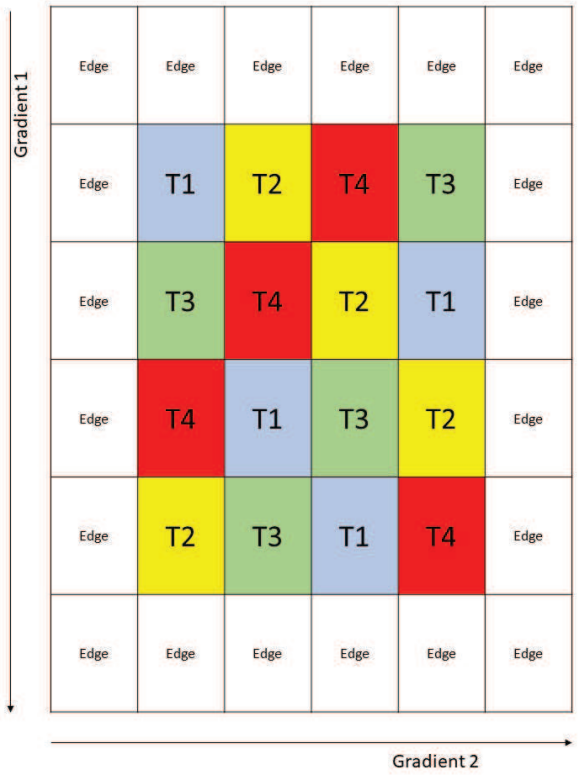

Fig. 12.6 Example of latin square experimental design

12.5 Nested design

In some particular experimental situations, a variable B may be nested in a variable A. In that case, factor A is the main factor (main independent variable), and the levels of factor B depend in part on the levels of factor A. Thus, if factor B is nested in levels of factor A, the levels of the nested factor B don't have exactly the same meaning under each level of the main factor A. In a nested design, the levels of factor (B) are not identical to each other at different levels of factor (A). For example, if we are dealing with parity orders (number of pregnancies) and the age of cows, lower parity orders are related to younger cows, while higher parity orders are associated with older cows. In that case, if we want to test whether milk production changes between different parity orders and between different ages in each parity order, we must deal with a nested design.

Parity 1			Parity 2		
Age 1	Age 2	Age 3	Age 4	Age 5	Age 6
Cow1	Cow4	Cow7	Cow10	Cow13	Cow16
Cow2	Cow5	Cow8	Cow11	Cow14	Cow17
Cow3	Cow6	Cow9	Cow12	Cow15	Cow18

Fig. 12.7 Example of nested experimental design

The statistical model applied to a nested design is:

$$y_{ijk} = \mu + \alpha_i + \beta_j(\alpha_i) + e_{ijk} \tag{12.4}$$

where: y_{ij} is the dependent variable measured on the experimental unit k subjected to the *i-th* level of the α variable and the *j-th* level of the variable β; μ is the mean of the whole sample; α_i is the effect of the *i-th* level of the variable α;$\beta_j(\alpha_i)$ is the effect of the *j-th* level of the variable β in the *i-th* level of the variable α; e_{ijk} is the vector of random errors. In R environment, a variable nested in another factor is indicated with %in% and the variable Error() with between brackets the indication of the nested variable %in% the main variable. The function we must use is the `aov()`.

12.6 Fixed or random variables: which model should I use?

When we are performing an analysis of variance we can identify fixed or random effects. Fixed effects:

- are the factors we are interested in;
- have few levels;
- we are interested in the tested levels and in estimating their effects;

Random effects:

- are factors we are not interested in estimating the effects;
- often have many levels;

- the levels of the random effect are a casual sample of the possible levels found in the population for that random variable.