STA2020 ANOVA Notes

Ané Cloete

2024-12-17

Table of contents

Preface

This is a Quarto book.

To learn more about Quarto books visit https://quarto.org/docs/books.

```
1 + 1
```

```
[1] 2
```

Part I

Experimental Design

1 Experiments and experimental design

There are two fundamental ways to obtain information in research: by *observation* or by *experimentation*. In an observational study the observer watches and records information about the subject of interest. In an experiment, the experimenter actively manipulates variables hypothesized to affect the response. Although both are important ways of understanding the world around us, only through experiments can we **infer causality**.

That is, by designing and conducting an experiment properly, if we observe a result such as a change in variable A leads to a change in our response (say variable B), we can conclude that A **caused** this change in B. If we were to merely study variable B and saw that as variable A changes, B also changes without conducting an experiment (i.e. by observation) then we can only say that variable A and B are associated. We could not conclude that any change in B is due to A, it could be some other factor that is correlated with A called a confoudning variable or it could be that B caused the change in A! The key is that a well-designed experiment controls and holds constant (as best we can) all other factors that might affect the response, so we can be sure the result is caused by the variable we manipulated.

Imagine a company wants to determine whether their voluntary employee training program (the explanatory variable) increases productivity (the response). They decide to track the productivity of employees who chose to complete the training and those who did not. They note that, on average, trained employees are more productive. Can we confidently conclude that the training program caused increased productivity?

This is an observational study since no variable was actively manipulated, they merely observed and recorded the productivity of two groups of employees. So, we cannot conclude that completing the training program increases productivity - we cannot infer causality. It could be due to many other factors, either observed or unobserved, such as maybe employees who chose to do the training program are inherently more motivated and thus productive. Can you think of any other factors?

If they actively manipulate the variable, training program, by randomly assigning employees to complete the training program or not and control other factors by ensuring the employees are as similar as possible (i.e. conducted an experiment). Any differences in productivity between the two groups could then be ascribed to the training program. If they happen to find that the employees who were assigned the training program are more productive, they can confidently say that the training program caused increased productivity (and perhaps make it compulsory for all employees!).

Experimental studies are extremely important in research and in practice. They are almost the only way in which one can control all factors to such an extent as to eliminate any other possible explanation for a change in a response other than the variable actively manipulated. In this course, we only consider experimental studies and those which aim to compare the effects of a number of **treatments**.

DEFINE treatments

Here are some other reasons for conducting experiments:

1. They are easy to analyse. A well designed experiment results in independent estimates of treatment effects which allow us to easily interpret the effects. EXPAND - independent treatment effects and/or independent treatment variables?

2. Experiments are frequently used to find optimal levels of settings (treatment variables) which will maximise (or minimise) the response. Such experiments can save enormous amounts of time and money. Imagine trying to find the optimal settings for producing electricity from coal without proper experimentation. Such a trial and error process would be extremely costly, wasteful and time consuming. In a similar vein, what if the fictional company in our previous example decided to invest a bunch of money in fine-tuning their training program based solely on the results of an observational study but in reality it turns out that adjusting their hiring process to identify more keen candidates would have been much more efficient and inexpensive.

3. In an experiment we can choose exactly those settings or **treatment levels** we are interested in, e.g. we can investigate the effect of different shift lengths (6, 8 or 9 hours) on employee productivity or test specific price points (R100, R150, R200) to determine which price maximizes sales or revenue. We can actively manipulate our variable to the levels we are interested in.

Experimental studies and their design are fundamental to science, allowing us to further knowledge and test theories. So lets define them more rigorously. We'll start by introducing some terminology.

2 Terminology

Treatment factors, treatment levels and treatments:

The **treatment factor** is the factor that the experimenter actively manipulates to measure its effect on the response. It becomes the **explanatory variable** (mostly categorical) in the model. For each treatment factor, we actively choose a set of **levels**. For example, the treatment factor "temperature" can have levels 10, 20, and 50°C.If temperature is the only treatment factor in the experiment, the treatments will also be 10, 20, and 50°C.

If we manipulate more than one factor (e.g., temperature and pressure), we have two treatment factors. When several treatment factors are manipulated, the experiment is called factorial and the **treatments** are all possible combinations of the factor levels. If we have pressure levels "low" and "high," there are 6 treatments in total:

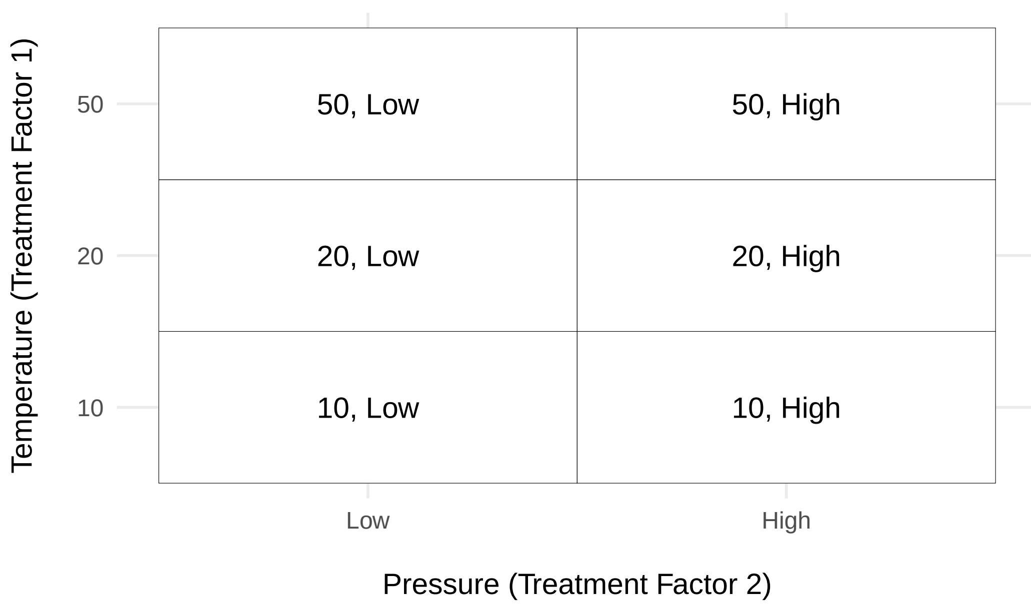

Figure 2.1: Visualization of how treatments are formed as combinations of treatment levels.

In the figure above, there are two treatment factors: Temperature (on the y-axis) and Pressure (on the x-axis). The axis ticks represent the levels of each treatment factor, and the blocks within the grid represent the treatments, which are specific combinations of the levels of Temperature and Pressure. Each treatment is labeled with the corresponding combination of levels (e.g., '50, Low' or '10, High').

Example 1

Three groups of students, 5 in each group, were receivng therapy for severe test anxiety. Group 1 recieved 5 hours, group 2 received 10 hours and group 3 received 15 hours. At the end of therapy each subject completed an evaulation of test anxiety. Did the amount of therapy have an effect on the level of test anxiety?
When faced with a text like this, it is useful to identify the treatment factors, their levels and the treatments, as well the response. Clearly, from the question, we are interested in the effect of therapy on test anxiety. A statement like this can generally be read as the effect of the treatment factor on the response. Nowhere is another treatment factor mentioned, so we only have one in this example. What are the levels of therapy we set? The levels are 5, 10 and 15 hours of therapy and since we only have one factor these are also the treatments. Let's summarise this as follows:

- **Response:** Test Anxiety
- **Treatment Factor:** Therapy
- **Treatment Levels:** 5, 10, and 15 hours of therapy
- **Treatments:** 5, 10, and 15

Experimental unit

The **experimental unit** is the entity (e.g., material, object, or individual) to which a treatment is assigned or that receives the treatment. The experimental unit may differ from the **observational unit**, which is the entity from which a measurement is taken. This distinction is very important because it is the experimental units which determine how often the treatment has been replicated and therefore the precision with which we can measure the treatment effect. In the methods that we cover in this course, we require that in the end there is only one 'observation' (response value) per experimental. If several measurements have been taken on an experimental unit, we will combine these into one observation by taking the mean. Very often, the experimental unit is also the observational unit.

Example 2.1 continued...

What are the experiemental units? To determine this, revisit the text of Example 1 and ask yourself: what entity received the treatments or to what were treatments applied? Most of you, will probably answer the students and this is correct. Each student received the respective treatment assigned to their group and so there are $5 \times 3 = 15$ experimental units. :
There is an argument to be made that it is not clear whether the students received therapy on their own or that the groups of students received therapy together. In that case, treatments were applied to groups of students and so there would be three experimental units.
We also need to know what the observational units are. The text states that at the end of therapy, each student completed an evaluation to determine their level of test anxiety. So the response, test anxiety, was measured on the student level which means students are the observational units. In the first case we looked at above, the students are both the experimental units and observtional units. But this would not be the case if groups are the experimental unit.
We also require that there is only one observation per experimental unit, the first scenario meets this requirement. For the second scenario, we have 5 observations per group and so we would have to take the mean of these values to end up wth one response value per group.
Let's add to the summary assuming students are the experimental units:

- **Experimental unit (no):** Student (15)
- **Observational unit (no):** Student (15)

Homogeneity of experimental units

When the set of experimental units are as similar as possible such that there are no distinguishable differences between then, they are said to be **homogeneous** (a fancy word for saying they are of the same kind). The more homogeneous the units are the smaller the experimental error variance (natural variation between between observations of the same treatments) will be. It is super important to have fairly homogeneous units because it allows us to detect differences between treatments more easily.

Blocking

If the experimental units are not fairly similar but are heterogeneous (the opposite of homogeneous), we can group them into sets of similar units. This process is called **blocking** and the groups are called blocks. We then compare the treatments within each block as if each block is its own mini-experiment. This way we account for the differences between blocks and can better isolate the effect of the treatments. More on this later.

Example 2.2 EDIT THIS STILL

Imagine you're testing the effectiveness of two marketing strategies (A and B) to increase sales at a chain of coffee shops. However, the coffee shops are located in different neighborhoods, and factors like neighborhood income level might influence sales. To ensure these differences don't skew the results, you group the coffee shops into "blocks" based on neighborhood income level (e.g., low, medium, high).
Within each block, you randomly assign coffee shops to either Strategy A or Strategy B. This way, you can compare the strategies while accounting for the variability caused by neighborhood income levels.

Replication

If a treatment is applied independently to more than one experimental unit it is said to be replicated. Treatments must be replicated! Making more than one observation on the same experimental unit is not replication, but *pseudoreplication*. Pseudoreplication is a common fallacy (REF?). The problem is that without true replication, we don't have an estimate of uncertainty, of how repeatable, or how variable the result is if the same treatment were to be applied repeatedly.

Example 1 continued one more time…

If we didn't take the average of the observations per group, we would have pseudoreplication as each student would not be an independent replicate of a treatment - effectively, we have only applied the each treatment once. You might notice that we then only have one true replicate per treatment group and this is problematic. To get an estimate of uncertainty, we would have to repeat this experiment a few more times to get more than one true replicate.
The first scenario, however, did not have this problem and each treatment was replicated five times. After going through all this, we have the following summary:

- **Response:** Test Anxiety

- **Treatment Factor:** Therapy
- **Treatment Levels:** 5, 10, and 15 hours of therapy
- **Treatments:** 5, 10, and 15
- **Experimental unit (no):** Student (15)
- **Observational unit (no):** Student (15)
- **Replicates:** 5

Creating a sumary like this, is a handy exercise for any experiment you come across, and we'll keep doing it for every example in this book. As we go along, we'll also add information about the type of experiment that was conducted. Next, we talk a bit more about experimental design.

3 The three R's of experimental design

Experimental Design is a detailed procedure for grouping, if blocking is necessary, experimental units and for how treatments are assigned to the experimental units. There are three fundamental principles, known as the 'three R's of experimental design' which are at the core of a good experiment.

3.1 Replication

We spoke about replication previously, here we will expand on why replication is critical to experimentation. Replication is when each treatment is applied to several experimental units. This ensures that the variation between two or more units receiving the same treatment can be estimated and valid comparisons can be made between treatments. In other words, replication allows us to separate variation due to differences between treatments from variation within treatments. for true replication, each treatment should be independently applied to several experimental units. If this is not the case, treatment effects become confounded with other factors. Confounding means that is not possible to separate the effects of two (or more) factors on the response,i.e. it is not possible to say which of the two factors is responsible for any changes in the response. This is what happened in the Example 1 when groups are the experimental units. With only one observation per experimental unit, the effect of therapy is confounded with the experimental unit or the effect of group on text anxiety.

Part II

Completely Randomised Designs

Part III

Randomised Block Designs

Part IV

Factorial Experiments