Stat 5200: Analysis of Designed Experiments

Jesse Wheeler

Contents

1 Course Introduction

Course Overview

- This class is all about developing techniques for answering scientific questions.
  - Is a newly proposed drug a safe and effective cure for a disease?
  - Which combination of protein and carbohydrate sources provide the best nutrition for growing lambs?
  - How will long-distance telephone usage change if our company offers a different rate structure to our customers?
- Our goal is to help decision makers and researchers design good experiments, analyze them properly, and answer their research questions.
- The course is based on the textbook "A First Course in Design and Analysis of Experiments" (Oehlert, 2000).

2 Course Notation

Useful course notation
Basic notations we will use throughout the class:

- y_{ij}
  - A general observation, with two subscripts. For instance, this might be the jth observation for the ith treatment group. We will often use the letter y for data, but in principle we could use any other letter, and we may do so when convinient. A comma is sometimes used if we use actual numbers; for instance, if $i = 1$ and $j = 2$, then we would write $y_{1,2}$ and not y_{12}, to avoid confusion.
- $j : k$, $y_{j:k}$, etc.

  - This is a convinient shorthand for a range of numbers. for instance, we may say $j : k$ is the set of numbers $j, j+1, \ldots, k-1, k$, with the understanding that if $k = j$, then $j : k = j$, and if $k < j$, then $j : k$ is the empty set. This will be used to say something like, "for all i in $1 : k$." When used as a subsecript, this implies the collection of all numbers with the defined subscripts. For instance, $y_{2,4:8}$ would be $y_{2,4}$, $y_{2,5}$, ..., $y_{2,8}$, meaning the fourth through eighth observation of the second (treatment) group.

- $y_{\cdot j}$ and $y_{i\cdot}$
  - This is a special shorthand to denote "all" of the indices in a particular group. For instance, $y_{\cdot j}$ might indicate "all jth observations from each group", and $y_{i\cdot}$ is "all observations from the ith group". Note that if there are g groups, then the following would be considered equivalent: $y_{1:g,j} = y_{\cdot j}$.

- $\bar{y}$, $\bar{y}_{i\cdot}$, etc.
  - The bar above a number is intended to be an average. For instance, $\bar{y}_{i,\cdot}$ is the average value of y in group i, and $\bar{y}_{\cdot,\cdot}$ would be the average of all y. Sometimes due to pure laziness on my part, I may indicate $\bar{y} = \bar{y}_{\cdot,\cdot}$ If there is any confusion, don't be afraid to ask.

- y^*_{ij}
  - Personal favorite, though probbaly won't be used much in this class. y_{ij} might represent any arbitrary value, whereas the "star" version y^*_{ij} is used to emphasize that it's not just an arbitrary value, but a real data point. This is particularly relevant if we are using function notation; $f(y_{ij})$ could be a function for any arbitrary input y_{ij}, whereas $f(y^*_{ij})$ is used to emphasize that it's not just any value, but the data plugged into that function.

- $\mathbb{R}$
  - All real numbers.

3 Experiments

What is an "Experiment" and why it matters

- Experiments are the gold-standard for scientific inference.
- Origins of the design a study of experiments starting with R. A. Fisher in the 1920s–1940s at UK agricultural center.
- His work developed the field of experimental design, which is used now in all forms of scientific investigations.
- Large advances in medicine, science, and product quality improvements can be attributed to proper experimental design and analysis.

Fisher's early examples
Fisher found decades worth of fertilizer treatment data:

- Fertilizer had been applied to a field one year, and not the next, in order to compare yield of grain produced in the two different years.
- Fertilizer was applied to one field in one year, and not to a nearby field in the same year, again to compare the yield of grain produced in the same year.

From a statistical perspective, there are three components that define an experiment: *Treatments*, *Experimental Units*, and *Assignment Method*.

Experiment
An *experiment* is specified by the combination of treatments, experimental units, and assignment methods.

- *Note:* There is no mention of a *method* for analyzing the results of an experiment. Strictly speaking, the analysis is not part of the design, though a wise experimenter will consider the analysis when planning an experiment.

Treatments

- *Treatments* are the different procedures we want to compare.
  - There can be more than one treatment.
  - Each individual treatment is possibly determined by several *factors*.
  - Treatments can be either categorical or numerical.
- **EX:** In a fertilizer formula, one gets to choose whether or not to include one of 20 chemical additives. How many possible treatments are there?

- **EX:** The baking treatment for a cake involves a given time at a given temperature. Considering 5 baking times and 5 temperatures, how many possible treatments can one include in an experiment?

Experimental Units

- *Experimental Units* (EU) are the entities which a treatment will be applied. Examples:
  - Plots of land receiving fertilizers.
  - Pens containing many pigs or other animals all getting the same diet (*Individual animal is NOT an EU*).
  - Classroom containing many students, all being taught the same way (Individual student is *NOT* an EU).
  - Individual patients in a clinic getting personalized treatments.

Assignment Methods

- *Assignment Method* defines how treatments are assigned to units (or units assigned to treatments). Examples (some good, some bad):
  - Assign fertilizer one year, not the next, to the same plot of land.
  - Split a single field into several sub-plots, and randomly pick subplots to apply fertilizer. Repeat this over multiple years.

Measurement Units vs Experimental Units

- A common source of difficulty and confusion is the distinction between *experimental units* and *measurement units*.

Definition: Measurement unit
A measurement unit is the individual or piece of experiment material you actually measure.

- Consider an educational study, where 6 classrooms of first graders are assigned at random to two different reading programs, and all first graders are later evaluated via a common reading exam at the end of the year.

- In an animal nutrition experiment, different types of diets are given to pens that each contain many individual animals.

High-level and Low-level EU's

- The distinction between experimental units and measurement units can become complicated.
- For example, in split-plot experiments (discussed later), there can be multiple layers of comparisons, with two or more types of experimental units with *low-level EUs*, serving as measurement units for the *high-level EUs*.

Baking experiment
A baking experiment is used to study the effect of high and low levels of three ingredients in a recipe ($2^3 = 8$ possible combinations), and the effect of baking temperature (6 different levels). You have an oven that can bake 8 batches at one time, at one temperature.
In the experiment, you run the oven 6 times, each with 8 batches made up of the 8 combinations of ingredients randomly placed in the oven.

What are the treatments, experimental units, and measurement units?

- The Treatments are temperature (6 levels, we call it a high-level treatment), and recipe (combinations of ingredients, a low-level treatment).
- The *high-level EUs* are the "runs" of the oven at a given temperature.
- The *low-level EUs* are the individual "batches" with unique ingredients that make up the 8 combinations.
- The *measurement units* are the 8 batches in an oven (the thing we measure), which is the same as the low-level EUs.

4 Observational studies vs experiments

Observational studies vs experiments

- Experiments help answer questions, but there are also non-experimental techniques. What makes experiments special?
- The most important feature is that an experiment gives researchers control: we get to pick the treatment and assignments.
- This is different from *observational studies*, where there are also treatments, units, and responses, but the researcher does not get to control treatment assignment.
- If an experiment is designed well, we can
  - Directly compare impact of various treatments, minimizing *confounding*.
  - Design experiments to minimize any form of *bias*.
  - Design experiments so that the *error* in comparison is small.

5 Design considerations

Design issues and considerations

- An important feature of an experiment is that the units are *independent*, meaning that the response of one unit is *not* influenced by treatments or responses of other units.
- We should also consider *edge effects*, which can occur in both space and time:
  - **EX (Space):** Plants at the edge of a plot or a greenhouse may respond differently to treatments.
  - **EX (Time):** Flushing out old treatments, or applying treatments at certain time of day.
  - We should account for these possibilities by ensuring sufficient separation of units, and provisioning a "buffer" in both space and time.

Fixed resource problem

- An often-encountered design problem is when there is a fixed amount of resource available for experimenting.
- Examples:
  - Fixed sized field
  - Fixed amount of time
  - Fixed amount of money to spend on experiment.
- In these situations, a key part of the problem include: Size / shape of plots, number of samples per session, etc.
- *Useful rule of thumb:* Do *more*, less well. That is, have more experimental units instead of more measurement units.

EX: In a comparison of teaching methods, it may be better to study 1 classroom in each of 20 schools (20 EUs, 20 measurement units), instead of 5 classrooms in 4 schools (4 EUs, still 20 measurement units). The trade-off here is that the first approach has more replicates, but higher-variance amongst EUs; the second approach has fewer replicates, but also lower variance in each replicate. Often the first approach will yield better final results.

Other design considerations

Open to the class: **What issues may arise that you should consider in an experimental design?**

Acknowledgments

- Compiled on August 19, 2026 using R version 4.5.0.

- We acknowledge students and instructors for previous versions of this course / slides.

References

Oehlert GW (2000). *A First Course in Design and Analysis of Experiments*. W. H. Freeman. ISBN 0-7167-3510-5. 1