Stat 5200: Analysis of Designed Experiments

Jesse Wheeler

Overview

Objectives and overview

- This material is based Chapter 3 on the textbook “A First Course in Design and Analysis of Experiments” (Oehlert2000).
- Chapter outcomes include:
    - TODO

Completely Randomized Designs

Completely Randomized Design

- Completely randomized design (CRD) is the simplest design for comparing treatments and is sometimes the preferred design.
- Setup:
  - There are g treatment groups.
  - N experimental units (EU)s available.
  - You want sample sizes of $n_1, n_2, \ldots, n_g$, with $\sum_{i=1}^{g} n_i = N$.

Completely Randomized Design II

- Idea: All possible sub-group composition are equally as likely.

Performing a Completely Randomized Design

- Randomly choose n_1 EU's for treatment 1.
- Randomly choose n_2 EU's for treatment 2.
- $\ldots$
- Randomly choose n_{g-1} EU's for treatment $g-1$.
- The remaining n_g EU's are assigned to treatment g.

- On a computer, this can be done by permuting $1, 2, \ldots, N$, and assigning the first n_1 numbers to treatment 1, the next n_2 to treatment 2, etc.

Advantages and Disadvantages

- Advantages
    - Simple to do; all you need is one random permutation.
    - Simple to explain and understand.
    - Simple to analyze.
    - No restrictions on the sample sizes (other designs may restrict n_i).
    - Easy with missing values.
- Disadvantages
    - Not always possible or cost-effective.
    - The accuracy of the experiment depends on variability among the EUs, not just the variability among treatment groups. Thus, it may require more EUs ($$$$).

Example: Acid Rain

Wood and Bormann1974) studied the effect of acid rain on trees. One experiment used 240 six-week-old yellow birch seedlings. These seedlings were divided into five groups of 48 *at random*, and the seedlings within each group received an acid mist treatment 6 hours a week for 17 weeks. The five treatments differed by mist pH: 4.7, 4.0, 3.3, 3.0, and 2.3. Othwerwise, the seedlings were treated identically. After the 17 weeks, the seedlings were weighed, and total plant (dry) weight was taken as response.

Questions:

- What type of design is this, and can you describe why?
- What are the EUs, and how many are there?
- What was the response?

Example: Acid Rain II

- What compromises may have been made in this case?

Like many experiments, the researchers necessarily have to narrow down an enormous question to a workable experiment using artificial acid rain under controlled conditions. For any experiment, a considerable amount of non-statistical background work goes into the planning even the simplest experiment.

Models and Parameters

Introduction and definitions

- All formal statistical analysis is based on probability models.
- *EX:* The number of heads in ten tosses of a fair coin would be Binomial$(10, 0.5)$. Here, the model depends on 2 parameters: $N = 10$ the number of trials, $p = 0.5$ is the probability of heads.
- In an experiment, we may posit several different models for the data, all with *unknown* parameters. One of our objectives is often to describe which model is the best description of the data, and make *inference* about parameters.

Definition: Inference

Using data to learn properties of a probability distribution. In our case, using data to estimate parameters of a model.

Introduction and definitions II

- The models we will consider for experimental design will have two basic parts:
  - A description of the expected value, or mean, of the data. Sometimes called "model for the means".
  - A description of how data vary around the treatment means. Sometimes called 'model for the errors".
- For the errors, we will generally assume that they are independent for different data values, have mean zero, and all have a common variance σ^2.
- To perform formal tests and obtain confidence intervals, we will further assume a distribution of the errors. Generally, we assume that the errors follow a Gaussian (Normal) distribution.

Introduction and definitions III

- For a CRD, we start with two basic models.
- Recall: g treatment groups, N EUs.

1. The Full Model (separate group means)

Introduction and definitions IV

2. The Reduced Model (single mean model)

Defining overall mean

- It's sometimes useful to express the group means μ_i as:

$$\mu_i = \mu^* + \alpha_i.$$

  In this case, we call μ^* the *overall mean*, and the α_i's are the ith *treatment effects*.

- This introduces an extra parameter (g+1 total), which we need to address. Generally, we do this by defining μ^* and require that $\alpha_i = \mu_i - \mu^*$.

Defining overall mean II

- One classic choice is to define μ^* as the mean of the treatment means:

Defining overall mean III

- An alternative choice (that can sometimes make hand-work simpler) instead describes μ^* as the weighted average:

Defining overall mean IV

- Another common choice in software is to set $\mu^* = \mu_j$, for some $j \in 1 : g$. Many software just use the last mean: $\mu^* = \mu_g$.

Model Recap

However you choose to define μ^* and α_i's, we end up with the following models for the data:

- *The full model*:

$$y_{ij} = \mu^* + \alpha_i + \epsilon_{ij}$$

- *The reduced model*:

$$y_{ij} = \mu + \epsilon_{ij}$$

Parameter Estimation

Estimating parameters

- We want to use data to estimate μ_i's and the variance σ^2. From these, we'll compute μ^* and α_i's.
- We will focus on computing unbiased estimates of these parameters.
- Estimate for μ_i:
- The estimate for μ^* is fixed, so it depends on how we defined it.

Estimating parameters II

- Unweighted average

- Weighted average

- Reference Group

Estimating parameters III

- The only parameter left to estimate is σ^2:

Estimating parameters IV

- Above, we have described point estimates. In some sense, these are our “best guess” as to the value of a particular parameter.
- We also want some measure of uncertainty of this estimate. The standard frequentist way of doing this is a confidence interval.
- Confidence intervals for α are not very useful, because it depends on how we define them. Confidence intervals for group means are typically defined as:

References and Acknowledgements

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

Wood T, Bormann F (1974). "The effects of an artificial acid mist upon the growth of Betula alleghaniensis Britt." *Environmental Pollution (1970)*, **7**(4), 259–268.

- Compiled on August 27, 2026 using R version 4.5.0.
-

References and Acknowledgements II

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