Stat 5200: Analysis of Designed Experiments

Jesse Wheeler

Overview

Objectives and overview

- This material is based Chapter 12 on the textbook "A First Course in Design and Analysis of Experiments" (Oehlert, 2000).
- Chapter outcomes include:
  - Students will be able to define both crossed and nested effects, and articulate the differences.
  - Students will be able to fit mixed effects models to data with nested or crossed designs.
  - Students will be able understand Hasse diagrams, and be able to build the basic "skeleton" of these diagrams.

Nesting vs. Crossing

Nesting Versus Crossing

- Up to this point, our multi-factor designs have utilized crossed factors.

Definition: Crossed Factors

Factors A and B are said to be crossed if treatments are formed by all possible combinations of the levels of the factors. That is, the same levels of Factor A are applied across all levels of Factor B.

Example: Replicated Latin Squares (LS)

In a replicated LS design, if we reuse one of the block factors (e.g., row blocks) across the squares, then the **Row** and **Square** factors are crossed, because each row sees each block.

Nesting Versus Crossing II

Definition: Nested Factors

Factor B is *nested* in Factor A if there is a completely different, unique set of levels of B for every level of A. That is, some levels of B only exist *within* certain levels of A.

Example: Replicated Latin Squares (LS)

In a replicated LS design, if we *do not* reuse any blocking factors, (i.e., unique row and column blocks for each square), then the **Row** and **Column** factors are said to be nested in **square**.

Fixed vs Random effects

- Nested factors can be random or fixed.
- A common situation is a random effect, resulting from some kind of subsampling.

Random nested effect

Within brand, we randomly sampled jars. It makes sense for Brand to be fixed, whereas we are not interested in the specific jars we picked, rather we want to use them as a representative of the brand that it is nested in. Thus, A (brand) should be fixed, and the nested factor B should be random.

Fixed vs Random effects II

- Nested factors can also be fixed.

Fixed nested effect

Consider a company with work crews. Each crew consists of four members.

- Members are nested in crews
- We get the same four crew members whenever we look at a given crew.

In this case, crew members would naturally be a fixed effect.

Example: Testing vitamin A content

Vitamin A study

Suppose you are concerned about differences in Vitamin A content within jars of baby food carrots. Our designed study:

- Go to the grocery store, randomly select 4 jars from each of the 3 brands of baby food that are sold in our region.
- Take two samples from each jar and measure the vitamin A in every sample.
- In total, there are 24 responses (2 per jar, 4 jars per brand, 3 brands).

Example: Testing vitamin A content II

- In the study design above, there are 3 sources of variation:
  - Variation between brands.
  - Variation between jars for each brand.
  - Variation between samples for every jar.
- It does *not* make sense to consider the main-effect of jar, OR the brand-by-jar interaction.

Example: Testing vitamin A content III

- Let A be the brand factor, B the jar factor. A nested model for this situation may be:

$$y_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \epsilon_{k(ij)}.$$

  - α_i: Fixed effect of Brand i.
  - $\beta_{j(i)}$: Random effect of Jar j *nested within* Brand i.
  - $\epsilon_{k(ij)}$: Random error of sample k nested in Jar j and Brand i.

  We have the standard model assumptions of normality and independence of random terms.

Nesting in R

The Nesting Coding in R

- How the data are encoded dictates your `lme4` syntax.
- **Implicit IDs:** If you number Jars $1, 2, 3, 4$ for Brand A, and $1, 2, 3, 4$ for Brand B, fitting `(1 | Jar)` is wrong. The software assumes Jar 1 is crossed between brands.
- **Explicit IDs:** If you number Jars uniquely $(1 \ldots 8)$, `(1 | Jar)` works perfectly.

The Nesting Coding in R II

```r
library(lme4)
library(lmerTest)

# Assume 'carrots' data is loaded (Oehlert, Ch 12)
# Brand is Fixed. Jar is Random.

# SCENARIO 1: Implicit IDs (Jars 1-4 repeated in each Brand)
# You MUST use the interaction operator ':' to nest Jar in Brand
mod_implicit <- lmer(VitA ~ Brand + (1 | Brand:Jar), data = carr

# SCENARIO 2: Explicit IDs (Jars uniquely numbered 1-8)
# lme4 automatically recognizes the nesting structure
mod_explicit <- lmer(VitA ~ Brand + (1 | Jar), data = carrots)
```

R formula "shortcuts"

- There are R formula shortcuts. For instance, we have used already:

$$\texttt{A*B} \implies \texttt{A + B + A:B}$$

- Similarly, there is a common nesting shortcut:

$$\texttt{A/B} \implies \texttt{A + A:B}$$

- **We must be careful with fixed vs random effects!** shortcuts will give you either fixed, or random, but not both:

$$\texttt{(1 | A/B)} \implies \texttt{(1 | A) + (1 | A:B)}$$

This may not be what you want if A is fixed and B is a nested, random effect!

Why do we use Nesting?

- We don't just "choose" to use a nested model. We use it because the experimental logistics or physical constraints dictate it.

Example: subsampling

A common example is **subsampling:** Taking smaller units from a larger selected "bundle" (Carrot example: Brands $\rightarrow$ Jars $\rightarrow$ samples).

Why do we use Nesting? II

Example: Logistical constraints

Sometimes nested designs are necessary due to logistical constraints:

Suppose there are 3 specialized clinics in the country. We can't realistically send the same patient to all 3 clinics to "cross" patients with clinics. Patients are nested in clinics.

Nesting vs Averaging MUs

- Distinctions between experimental units and measurements units can sometimes be explained by nested factors.

Nesting vs Averaging MUs II

Testing Feed on Pig Growth

Suppose we have multiple pens of pigs, and g different feeds we would like to test. We randomly select pens to apply different feed, and observe the growth of individual pigs.

- The EUs are pig pens, the MUs are pigs.
- Previously: To test differences in treatment, we aggregated results over the measurement unit.
- Blindly treating the pigs as EUs would lead to "pseudo-replication", because randomization was not done to the MUs, which is why we aggregated.

Nesting vs Averaging MUs III

- This same design that we discussed really has a nested structure.

Nested Design: pig pens

In the same design, there are 3 sources of variation: Treatment effect (A), pen effect (B), and individual pig effect (error).

Rather than aggregating MUs, we could instead use a nested model: pen (B) is nested in feed type (A) because Pen 1 only gets one level of the treatment.

Nesting vs Averaging MUs IV

- Which is the correct approach? **They Both are**!
- If the design is balanced, the F statistic and P value for the main treatment effect (feed, factor A) will be the same regardless of the approach we pick.
- This new approach (nested instead of averaging) does not change the treatment effect estimates, but now allows us to compare / estimate pen-to-pen variance σ^2_{pen} against the pig-to-pig variance σ^2.

Crossed and Nested Factors: Examples

Example: Cheese tasting (Ex 12.2)

Cheese tasting experiment

- 4 different types of 50-lb **blocks of cheese** selected.
- The **raters** are students from a large, introductory food-science class.
- Student **background** (urban vs rural) identified; 10 students from each background randomly selected.
- Each rater will taste 8 **bites** of cheese in random order; there are 2 "bites" of each of the 4 different types (raters don't know this).

Example: Cheese tasting (Ex 12.2) II

What are the sources of variability?

- Variability due to different cheese types.
- Variability due to background.
- Variability due to different raters.
- Random error across bites.

Example: Cheese tasting (Ex 12.2) III

- The Raters are nested in Backgrounds (i.e., rater 1 has only one of the backgrounds, rural vs urban).
- Samples (or bites) are nested in Rater (they don't share bites).
- Each rater sees each cheese type, so they are crossed.
- Similarly, each background sees each cheese type, so they are also crossed.
- The random error across bites is nested in the Rater-by-Type interaction (and Rater itself is nested in Background).

Example: Cheese tasting (Ex 12.2) IV

- Finally, the model. The factors are Cheese type (C), Background (B), and Rater (R). The Bites serve as our measurement units (error). An appropriate model would be:

$$y_{ijkl} = \mu + \alpha_i + \beta_{j(i)} + \gamma_k + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk(i)} + \epsilon_{l(ijk)},$$

- α_i (background) and γ_k (cheese type) are fixed effects, and are fully crossed.
- $\beta_{j(i)}$ is the rater (random) effect, and is nested in background. It is crossed with cheese type, giving an interaction (random) effect: $(\beta\gamma)_{jk(i)}$.

Example: Cheese tasting (Ex 12.2) V

```
# Assuming 'cheese' data is loaded
# Type (C) and Background (B) are fixed and crossed.
# Rater (R) is random, nested in Background (B).

mod_cheese <- lmer(score ~ Type * Background +
                     (1 | Background:Rater) +
                     (1 | Background:Rater:Type),
                   data = cheese)

# (1 | Background:Rater) accounts for the random variance of the
# (1 | Background:Rater:Type) accounts for the random interactio

#   how individual raters react differently to different cheese
#   The Background is included because of the nesting.
#   Because we have 2 bites per combination, this can be separat
```

Designs with Crossed and Nested Factors

- Designs can have both crossed an nested factors.
- In addition to a nesting structure, there can be treatment factors that can be combined and assigned at random to units such that all combinations of nesting factors and treatment factors get an equal number of units.
- **Figuring out which factors are crossed and which are nested, and the appropriate model, can be challenging.** The best way to figure it out is practice!

Example: Gum arabic (Example 12.1)

- Gum arabic is used to lengthen the shelf life of food items.
- We wish to see how different gums and gum preparation effect shelf-life.

Gum arabic design

- We obtain 4 raw gum samples from each of 2 varieties of trees (8 total samples).
- Each sample is split into 2 subsamples (16 total).
- One of the subsamples (randomly selected) will be demineralized, the other will not.
- All 16 subsamples are dried, and we make 5 measurements from each subsample.

Example: Gum arabic (Example 12.1) II

First, what are the sources of variability?

- Variance from Tree variety (2-levels)
- Variance from random samples from each tree (4-levels for each variety)
- Variance from the treatment effects (2 levels)
- Variance of the sub-sample*

Example: Gum arabic (Example 12.1) III

- The larger samples are nested in tree variety;
- The subsamples are nested in the samples.
- Subsamples get applied one of two treatments (demineralization)
  - Because each subsample gets exactly one of each treatment, each sample gets both levels of the treatment, and sample and treatment are crossed.
  - Similarly, each variety of tree gets both levels of treatment, so variety and treatment are crossed.
- The five individual measurements within a single subsample are nested in subsample, or the variety-sample-treatment combinations.

Example: Gum arabic (Example 12.1) IV

- Finally, the model. Let Variety, Sample, and demineralization (treatment) be factors A, B, and C. An appropriate model would be:

$$y_{ijkl} = \mu + \alpha_i + \beta_{j(i)} + \gamma_k + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk(i)} + \epsilon_{l(ijk)},$$

- α_i and γ_k (and their interaction) are fixed effects, and are fully crossed.
- $\beta_{j(i)}$ is the sample (random) effect, and is nested in variety. It is crossed with treatment, giving an interaction (random) effect: $(\beta\gamma)_{jk(i)}$.

Gum Example Code in R

```
# Assuming 'gum' data is loaded
# Variety (A) and Demin (C) are fixed and crossed.
# Sample (B) is random, nested in Variety.

mod_gum <- lmer(ShelfLife ~ Variety * Demin +
                  (1 | Variety:Sample) +
                  (1 | Variety:Sample:Demin),
                data = gum)

# (1 | Variety:Sample) accounts for the random variance of the s
# (1 | Variety:Sample:Demin) accounts for the random interaction

# how individual samples react differently to demineralization.
```

Building models

Recommendations for building models

- Building designs and models for nested, mixed-effects models is not easy, and takes some practice. Some recommended things to keep in mind:
- **Identify sources of variability**. These are things that we think might be important to account for, both intentional treatments and also natural structure of the data. These become the Factors in our model.
- **Determine Nested vs Crossed**. Identify which (if any) of the factors are nested, and which are crossed.
- **Identify fixed vs Random Effects**. Are the specific sources of variance of interest, or are they more of a nuisance / sample of larger population?

Hasse Diagrams

Visualizing Design: Hasse Diagrams

- As designs get complex, keeping track of what is nested inside what, and what crosses with what, becomes incredibly difficult.
- **Hasse Diagrams** (Lohr, 1995) are graphical representations of a model showing the exact nesting/crossing and random/fixed structure.
- While historically used to manually derive F tests and random effects estimates, they are useful for building `lmer` formulas and understanding designs.

Rules for Drawing a Hasse Diagram

1. Start at the top with a node **M** for the Grand Mean.
2. Add nodes below M for each factor that is not nested. Draw lines connecting them to M.
3. **Random factors go in parentheses:** e.g., (Jar)
4. For nested factors, draw them *below* the factor they nest in, connected by a line (e.g., Brand $\rightarrow$ (Jar)).
5. For crossed interactions, draw a node below the interacting main effects, with lines connecting up to both parents.
6. **Infection Rule:** Any term located below a random term in the diagram becomes random itself, and must be in parentheses!
7. End at the bottom with a node for **(E)** Error.

References and Acknowledgements

Lohr SL (1995). "Hasse Diagrams in Statistical Consulting and Teaching." *The American Statistician*, **49**(4), 376–381. doi: 10.1080/00031305.1995.10476188. URL https://www.tandfonline.com/doi/abs/10.1080/00031305.1995.10476188.

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

- Compiled on November 24, 2026 using R version 4.6.2.
- Licensed under the Creative Commons Attribution-NonCommercial license. Please share and remix non-commercially, mentioning its origin.

References and Acknowledgements II

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