Chapter 9: Repeated measures and multilevel models

1. A researcher conducts an fMRI study to investigate how the brain responds to emotionally charged images. Each participant completes the task inside the scanner and views 60 images (20 negative, 20 neutral, 20 positive). The main dependent variable is brain activation in the amygdala, extracted for each image trial.

    - Consider the structure of this dataset. What variables vary *within* participants, and what varies *between* participants?
    - Are the amygdala responses across trials likely to be independent from one another? Why or why not?
    - Identify two different sources of grouping or clustering in this dataset. How might this influence the similarity between observations?
    - Describe the nested hierarchical structure of this dataset (e.g., levels of variation).
    - Reflect: If you wanted to examine the effect of image valence (negative vs. positive) on brain response, what would be the limitations of using a regular linear regression model?

2. A cognitive psychologist runs an experiment to study the effects of background noise on working memory. Participants complete a memory task under three conditions: silence, café noise, and traffic noise. This is a multi-site study, conducted across different research labs, each based at a different university, using similar but independently run testing procedures.

    - Suppose the research question is whether background noise type affects performance, regardless of where the study was conducted. Should lab be treated as a fixed or random effect?
    - Suppose the researcher suspects that certain labs consistently produce higher or lower scores – perhaps due to differences in equipment, recruitment, or testing environment. Would this change how you model lab? How might you incorporate this information while still accounting for clustering across labs?

3. In this exercise, you will analyse data from a randomised feasibility trial that investigated the effectiveness of a psychological intervention to support young adults with perinatally acquired HIV (PAH) in their HIV disclosure decision-making[1]. The subset of data presented here contains only the subset of participants recruited in Uganda. Participants were randomly assigned to an *intervention* or *standard of care* condition. Psychological wellbeing was measured at two timepoints: before the intervention (baseline) and approximately six months later (follow-up). Wellbeing was assessed using the 6-

---

[1] Evangeli, M., Gnan, G., Musiime, V. et al. (2024) 'The HIV Empowering Adults' Decisions to Share: UK/Uganda (HEADS-UP) Study—A Randomised Feasibility Trial of an HIV Disclosure Intervention for Young Adults with Perinatally Acquired HIV.' *AIDS and Behavior* 28, 1947–1964. https://doi.org/10.1007/s10461-024-04294-2.

item psychological domain of the WHOQOL-BREF scale, a validated measure of psychological quality of life.

The dataset includes the following variables:

- `participant_id`: unique code identifying each participant
- `group`: treatment group (*Intervention* or *Standard_of_care*)
- `baseline`: wellbeing score (WHOQOL psychological domain) at baseline
- `follow_up`: wellbeing score at six-month follow-up
- `age`: participant age (in years)
- `sex`: participant sex

You can load the dataset by downloading it in your working directory and using `read.csv("headsup_qol.csv")`.

- Reshape the dataset from wide to long format so that `baseline` and `follow_up` scores are stored in a single column named `wellbeing`, and create a new column named `timepoint` indicating whether the score comes from the baseline or follow-up session. The resulting dataset should contain one row per participant per timepoint.
- Fit a linear multilevel model predicting `wellbeing` from `timepoint`, `group`, and their interaction, with a random intercept for `participant_id`. Interpret the results:
  - Did wellbeing scores change significantly over time?
  - Was there a difference in the amount of change between the two groups?
  - Do the results suggest that the intervention had a positive impact?
- Extend the model to include participant age and sex as additional fixed effects. Does controlling for these variables change the results or conclusions?
- Create a visualisation that clearly displays the change in wellbeing scores across time for the two groups. You might use a line plot showing the average score at each timepoint for each group, or a plot of individual participant trajectories.

4. This exercise uses the `mathachieve` dataset, based on the 1982 *High School and Beyond* survey, which includes data from 7,185 students across 160 schools in the United States. Assuming the data has been downloaded in your working directory, it can be loaded into R with:

```
mathachieve <- read.csv("mathachieve.csv")
```

The main outcome variable is `mathach`, a score on a math achievement test. Each row corresponds to a student, and schools are identified by the variable `school`.

Socioeconomic status (SES) is measured at two levels:

- `ses`: individual student's socioeconomic status (continuous)

- `meanses`: average SES for the school (continuous)

Additional variables include:

- `sex`: student sex (male/female)
- `minority`: whether the student belongs to an ethnic minority group
- `sector`: school type (Catholic or Public)
- `size`: number of students in the school

- Fit a first model, predicting the math achievement from individual student socioeconomic status (`ses`), including a random intercept for school. Does socioeconomic status predict match achievement? Test this using both p-values and bootstrapped confidence intervals.
- Given that data is 'clustered' or grouped at the level of school, which predictors can have a random slope, and which cannot?
- Fit a multilevel model predicting `mathach` from `ses`, `meanses`, `sector`, and `size`, allowing the intercept and the slope for `ses` to vary by school. What do the fixed effects tell you about the relationship between SES and math achievement? Do school-level SES and sector still matter when controlling for individual SES?
- Add an interaction term between `ses` and `sector`. Use a likelihood ratio test to evaluate whether this improves model fit. Interpret the interaction: does the relationship between SES and math achievement differ by school type?
- Calculate the estimated slope of `ses` separately for Catholic and Public schools using the model output. What might this suggest about the role of SES in each type of school?
- Explore and visualise the distribution of `meanses` by school sector. What patterns do you observe? How might this help interpret the interaction between `ses` and `sector`?