Statistical Modeling, Winter 2025

Homework 2

Please submit your answers to these questions via Canvas prior to class next Tuesday. I prefer a pdf, or a .Rmd / .Qmd / .R file. Please name the file with your last name and hw number (e.g., elahi_hw1).

Complete the following readings:

- chapter 4 in *Statistical Rethinking* (McElreath 2020).

Make sure you can answer Practice problems 4E1-4E5, 4M1-4M3 (don't submit these!). The answers to all practice problems are here.

Questions

From McElreath

1. From the Howell1 dataset, consider only the people younger than 13 years old. Estimate the causal association between age and weight. Assume that age influences weight through two paths. First, age influences height, and height influences weight. Second, age directly influences weight through age-related changes in muscle growth and body proportions. Draw the DAG that represents these causal relationships. And then write a generative simulation that takes age as an input and simulates height and weight, obeying the relationships in the DAG.
2. Estimate the total causal effect of each year of growth on weight.

Project

3. Read the 'project_overview.pdf' and choose your own adventure below.

Do you have your own dataset in hand, with at least 3 predictors and 1 response variable, AND you are willing to share your data and code with me and your fellow students? If yes to both, answer question A. If not, go to question B.

A. Begin by describing your dataset:

- When was it collected?
- Where was it collected?
- Who collected the data?
- What data were collected?
- Why was it collected?
- How was it collected?

Also, summarize your initial thoughts about the questions you wish to answer with these data. Note that your questions may differ from the original reasons the data were collected. As you work with the dataset in the future, these questions can be refined as necessary.

B. Given that you do not have a dataset in hand ready for analysis, you will instead look for a publication with an appropriate dataset. Do a literature search on a topic relevant to your interests, and find a paper that meets the criteria outlined in the project_overview pdf. For the paper, submit the following:

- citation
- link to paper (e.g., DOI)
- link to the data
- a few sentences about why it is a good fit for this assignment
- comments on limitations or concerns

References

McElreath, Richard. 2020. *Statistical Rethinking: A Bayesian Course with Examples in R and Stan.* CRC Press.