PSY 503 Problem Set 1

Dr. Jason Geller

September 13, 2022
Due: September 23, 2022

Question 1

A parody appearing in the British Medical Journal questioned whether parachutes are in fact effective in preventing death when skydivers are presented with severe "gravitational challenge." The authors point out that no randomized trials have assigned parachutes to skydivers. Why is it reasonable to believe that parachutes are effective even in the absence of randomized experiments that establish their efficacy?

Question 2

In experiments, participants are randomly assigned to an experimental condition. Explain in your own words why random assignment to experimental conditions allows to obtain unbiased estimates of average treatment effects.

Question 3

A. Explain the notations $Y_i(0)$ and $Y_i(1)$.

B. How do you understand the notation $Y_i(0)$ in an experiment in which $d_i = 1$

C. Describe the difference between potential outcomes and observed outcomes?

Question 4

Create a simple Treatment and Control experiment with fake data. Use the values to illustrate that

$$\mu_{Y_i(0)} - \mu_{Y_i(1)} = \mu_{(Y_i(0) - Y_i(1))}$$

Question 5

1. Open a new R Script in R Studio. Save it as "pset_1_yourlastname.R"
2. Create an object called `subject_id`. This object should return a vector of 100 randomly selected numbers between 1,001 and 9,999. Make sure that these numbers cannot appear more than once in this vector.
3. Imagine that each subject id refers to a participant in your experiment. In this study, you measured participants' ability to memorize a series of words in two different environments: a noisy vs. silent environment. Specifically, participants were presented with a list of 30 words for 30 seconds and had 2 minutes to write down as many of these words as possible.

(a) Simulate the schedule of potential outcomes these 100 subjects, assuming that the silent condition has a **constant** effect of size 3 words. (Hint: You will need to create two objects: `Y_0` and `Y_1`. For `Y_0`, draw random numbers from the range of possible values that `Y_0` can take.)

What is the mean of `Y_1`? What is the mean of `Y_0`.

(b) Create an object called `condition`, which randomly assigns each subject to a condition.

(c) Create a tibble called `dat`, which includes `subject_id`, `Y_0`, `Y_1`, and `condition`.

4. Imagine that you just collected data based on this specific random assignment.

(a) Calculate and report the estimated ATE.

(b) Assume that the 100 participants constitute your population of interest. Is the estimated ATE biased?

5. Save your tibble as an RData file and submit it, along with your code.

Question 6

Name that Threat!

1. As part of a course, an instructor has his students take the GREs. The students do not do well. The instructor decides to offer a one-hour session to help his students prepare for the GRE. Afterwards, he has the students take the GRE a second time and finds that their scores have improved. Can we conclude that the one hour prep-session was effective?

2. Researchers want to evaluate a pre-school readiness program for 3- to 4-year olds. To determine the effectiveness of the program, the researchers randomly assign children either to participate in the readiness program or to participate in a daycare program. During the course of the study several parents notice that their children are performing poorly in the readiness program. These parents decide that their children are not developmentally ready to participate and take them out of the readiness program. Can the researchers conclude that the program is effective if they find more positive results among those in the readiness program versus those in the daycare program?

3. Researchers were interested in examining the effectiveness of a cognitive-behavioral treatment on students' stress levels. The researchers made contact with several university counseling centers and recruited participants with extremely high stress levels. Any post-treatment improvement could have been due to the treatment, but it could also have been due to what?