ECON 20: Lecture #2
Practice Problems (Answers)

1 Rubin causal model

In the Rubin causal model, Y_i represents the outcome of interest and D_i is the treatment status indicator ($D_i = 1$ if treated; $D_i = 0$ if not treated) for person i. For concreteness, suppose the outcome Y_i is "earnings at age 30" and the treatment is "attend Dartmouth." Answer the following questions.

1. Suppose you are person i. In this framework, what do Y_{i1} and Y_{i0} repersent in words? What does D_i represent in words? What value does it take for you?

   Answer: You (person i) are a Dartmouth student. Therefore, $D_i = 1$ because you are attending Dartmouth. Y_{i1} is your actual earnings at age 30. Y_{i0} is what your earnings at age 30 *would have been* had you not attended Dartmouth.

2. Which of the following are *counterfactual*?

   (a) $E[Y_{i1}|D_i = 0]$

   (b) $E[Y_{i1}|D_i = 1]$

   (c) $E[Y_{i0}|D_i = 0]$

   (d) $E[Y_{i0}|D_i = 1]$

   Answer: (a) and (d) are counterfactual because the version of i's earnings we are examining do not match their actual treatment status. Specifically, $E[Y_{i1}|D_i = 0]$ is the expected earnings had person i attended Dartmouth for someone who did not attend Dartmouth and $E[Y_{i0}|D_i = 1]$ is the expected earnings had person i not attended Dartmouth for someone who did attend Dartmouth.

3. For Dartmouth students, what is the causal effect of attending Dartmouth on earnings at age 30 using the notation of the Rubin model?

   Answer: $E[Y_{i1}|D_i = 1] - E[Y_{i0}|D_i = 1]$ is average earnings for Dartmouth students ($D = 1$) minus what Dartmouth students would have earned on average had they

not gone to Dartmouth. This is a theoretical concept because $E[Y_{i0}|D_i = 1]$ is not observable.

4. Now, suppose that you want to figure out the causal effect of attending Dartmouth on earnings. To do so, you compare the average earnings of people who did and did not attend Dartmouth. How would you describe this calculation in the notation of the Rubin model?

   Answer: $E[Y_{i1}|D_i = 1] - E[Y_{i0}|D_i = 0]$

5. In the context of the research approach described in part (4), you conduct a *balance test*. What would that consist of? Please provide a specific example. And what is the purpose of the balance test?

   Answer: A balance test would compare the average characteristics of those who attended Dartmouth and those who did not. We should be especially interested in things that might predict long-run earnings separately from attending Dartmouth. For example, we might want to compare family background or test scores across the two groups of individuals. The purpose of the balance test is to assess whether there is selection bias as we discussed in lecture.

6. Under what conditions does the approach in part (4) deliver the causal effect of attending Dartmouth? I recommend trying to write the answer both mathematically and in words. Ultimately, being able to answer this type of question cleanly and clearly in words is one of the goals of this class.

   Answer: Mathematically, we need $E[Y_{i0}|D_i = 1] = E[Y_{i0}|D_i = 0]$. What does this mean in words? It says that the average earnings of those who did not attend Dartmouth, $E[Y_{i0}|D_i = 0]$, is the same as the average earnings of those who did attend Dartmouth *would have been had they not attended Dartmouth.* That is, absent attending Dartmouth, the two groups you are comparing would have had the same earnings.

7. Suppose you had some data that included a random sample of 50 Dartmouth graduates and a random sample of 50 non-Dartmouth grads. You calculate the mean earnings of the two groups and take the difference. Name two reasons why this difference may not necessarily represent the causal effect of attending Dartmouth?

   Answer: (1) Dartmouth graduates may differ from non-Dartmouth graduates in other ways that affect their earnings. in other words, the condition in part (6) is violated

and there is selection bias. (2) Pure chance in a random sample. This is an issue of sample variation that we will cover in the coming weeks.

8. Is the question "does attending Dartmouth raise your earnings at age 30" well-posed? Choose yes or no and briefly justify your answer.

   Answer: I think so. Attending Dartmouth is a meaningful treatment which could be manipulated and randomized. Generally, things that can be direclty randomized correspond to well-posed questions.