Making Potential Outcomes Concrete

In this exercise, I will describe a simple data science project, and your job is to map the elements of that project onto the Potential Outcomes Framework.

The Study

You have been approached by the Duke Office of Student Wellbeing. They are worried that due to the COVID pandemic, students are spending too much time sitting in front of their computers and not being active, leading to an increase in student Body Mass Index (BMI) scores (a measure of whether someone's weight is more or less than one might expect given their height) which might, in extreme cases, have negative health consequences.

One policy they're considering implementing to address this issue is to replace all sodas in the dining halls with diet (sugar-free) sodas. But they aren't sure if this will actually make a difference.

To help them decide, they wanted to measure the effect of diet soda on BMI, so they collected data on all Duke students, and compared student BMI scores for students who drink diet soda with those who drink regular soda.

For simplicity (and because these are students) assume all students drink Coke, the only question here is whether they drink Diet Coke or Regular Coke.

They found that students who drink diet soda actually have higher BMIs (suggesting they may actually be *less* healthy). They interpreted this as evidence that diet soda is making students *less* healthy, so they changed their plans, and instead of removing sugary soda, they've decided to remove all *diet* soda from the campus.

Mapping to the Potential Outcomes Framework

In words, describe exactly what quantities in this study context would correspond to the different components of the Potential Outcomes Framework we've been studying.

When defining these, remember to define the outcome being measured and the population for whom the outcome is being measured. Every answer should include both these things.

As you do so, avoid using terms like "treatment", "control group", or "potential outcome". The goal of this exercise is to move from the abstract conceptual derivations we've read to the specifics of this study. For example, I've put in an answer to Question 1 below:

1: $E(Y_{T=0})$

> The average BMI of all Duke students if no one drank diet soda.

2: $E(Y_{T=1})$

. .

3: $E(Y_{T=1}) - E(Y_{T=0})$

. .

4: $E(Y_{T=1}|D=1)$

. .

5: $E(Y_{T=0}|D=0)$

. .

6: $E(Y_{T=1}|D=0)$

. .

7: $E(Y_{T=0}|D=1)$

. .

8: $E(Y_{T=1}|D=0) - E(Y_{T=0}|D=0)$

. .

9: $E(Y_{T=0}|D=1) - E(Y_{T=0}|D=0)$

. .

Observability

10: Now, which of the quantities above can be directly observed?

. .

Causal Inference

In order for the difference in BMIs found in the report—that those who drink diet soda have higher BMIs—to be a true estimate of the average effect of drinking diet soda, we know that it must be the case that:

$$E(Y_{T=0}|D=1) - E(Y_{T=0}|D=0) = 0$$

and

$$E(Y_{T=1}|D=0) - E(Y_{T=0}|D=0) = E(Y_{T=1}|D=1) - E(Y_{T=0}|D=1).$$

In the context of this study, what do those two conditions mean *in plain English*?

As above, avoid using abstract terms ("treatment", "baseline", etc.) and try and be as concrete as possible.

11 $E(Y_{T=0}|D=1) - E(Y_{T=0}|D=0) = 0$:

. .

12 $E(Y_{T=1}|D=0) - E(Y_{T=0}|D=0) = E(Y_{T=1}|D=1) - E(Y_{T=0}|D=1)$:

. .

Now, for each of the conditions above, please give one reason—*in plain English*—why those conditions may **not** be met in the context of this study?

As you do so, be specific! Tell me a story about why *in the case of this study* you think one of these conditions may hold. One can always say things like "people in the two groups may have been different", but I want a specific, *affirmative* reason you think they might have been different in a way that meets the conditions.

13 It may be the case that $E(Y_{T=0}|D=1) - E(Y_{T=0}|D=0) \neq 0$ because...:

. .

14 It may be the case that
$E(Y_{T=1}|D=0) - E(Y_{T=0}|D=0) \neq E(Y_{T=1}|D=1) - E(Y_{T=0}|D=1)$
because...:

. .