Lecture 14: Causality

Prof. Esther Duflo

14.310x

Game plan

1. What is causality?
2. Different ways to model causality
   - Structural equation modelling
   - The Rubin Causal Model
3. Causality and RCT

What is Causality?

Causal statements we make in everyday life:

- Her headache got better because she took a pill
- She got a good job because she went to MIT
- She cannot get a job interview because she is African American.

What do we mean by these statements?

- There is a counterfactual world in which she does not take a pill
- Instead of going to MIT, she could have done something else (What? not entirely clear from the statement)
- Not entirely clear what we mean here? change her race? change the way people think about race when they make hiring decisions?

In general, when we think of causality we think of the possible effect of *manipulating* a cause, and what would happened if we had nor had not manipulating this cause.

Do we only care about causal statements?

- Many of the questions we want to answer in economics/social science are causal questions: Does immigration lower the wages of the native workers? Does trade increase inequality? Would a wall between Mexico and the US stop immigration?
- So a lot of data science in social sciences aims to answer question of causes and effects.
- But it is worth nothing that there are some questions that are also important and are *not* causal questions. For example, we may be interested in identifying early warning sign of children at risk in school, so we can focus our effort on them. Google wants to predict what someone may be interested in based on their search patterns to serve them the add that is most likely to be of interest.

Regression equations

- When we start looking at regressions, we will typically aim to represent the data with a simple equation.
- Bivariate: $y = \beta x + \epsilon$
- Typically (though not often), we think this model as a causal interpretation: In an ideal experiment where we control a random variable X to x, and leave the rest of the world unchanged the value y of Y is given by $\beta x + \epsilon$

Regression equation

- Model with control
- $$Y_i = \alpha + \beta s_i + \gamma X_i + \epsilon_i$$
- Y_i is the outcome (wet floor), s_i is the treatment (sprinkler) and X_i are control variables (season, rain).
- There is no formal distinction between the "treatment" and the control, just a conceptual one
- ϵ_i is the residual, or error term: the difference between the true value of Y_i and the fitted value.
- More on this later!

The potential outcome framework

- This model is due to the Harvard Statistician Donald Rubin
- I find it very useful to think about randomized controlled trials (which will occupy it for a couple of lectures) and about causality more generally.
- It is not the only (or perhaps most common way) to think about causality in social science (SEM is much more common), but it is spreading, and it is very useful to be conversant with it and to be able to toggle from one to other
- Given a unit, and a set of actions, we associate each action-unit pair with a potential outcome.

Three examples

Thinking in terms of potential outcomes forces us to think about the counterfactual, and help us define well posed causal questions.

1. Her headache got better because she took a pill
2. She got a good job because she went to MIT
3. She cannot get a job interview because she is African American.

1. headache (yes; No) –Pill (yes, no). Some times we may refer to the no pill as the "control" and pill as the "treatment"
2. Our second example is a bit less clear: what was the alternative if she did not go to college?
3. And our third example even less clear: what do we mean by what would happen if she was from another race? what are various ways

Definition of causal effect

For any unit, the causal effect of a treatment is the difference between the potential outcome with and without the treatment.

- Headache example. Four possibilities
  1. Y(aspirin)=No headache; Y(no aspirin)=Headache
  2. Y(aspirin)=Headache; Y(no aspirin)=Headache
  3. Y(aspirin)=No headache; Y(no aspirin)=No headache
  4. Y(aspirin)=Headache; Y(no aspirin)=No headache
- Treatment effect
  1. Make headache go away
  2. No effect
  3. No effect
  4. Make headache appear

Discussion: the problem of causal inference

- The definition of treatment effect depend on potential outcomes but not on what is actually observed
- The causal effect is the comparison of the same unit, at the same time (post treatment). "fundamental problem of causal inference" (Holland, 1986) is at that at most one of the potential outcomes can be realized, and thus observed.
- For the estimation of treatment effect, we will need to make comparison for what we observe.
- Thus we will need many units (for this discussion, two different measurements of the same person over time is two different units.
- It will be critical to know (or make assumption about) the way that some potential outcomes got realized, and not others: This is the discussion of assignment mechanisms, which we will have in a bit.

The problem with many units

- When think about more than one unit, things can quickly become complicated.
- Imagine I am with Sara in her office and we are both preparing class notes.
- We can may be both have a headache, and both of us has the option to take aspirin
- Now each of us as 4 potential outcomes: Y(EA, SN), Y(EA, SA), Y(EN, SN), Y(EN, SA).
- In this situation, there are $\binom{4}{2} = 6$ different comparisons depending on which of the potential outcomes are compared
- As we are adding more units we are adding more potential comparison: we will never get enough data to estimate what we want.

Stable Unit Treatment Value Assumption

- (perhaps) Natural assumption in the headache example: Sara's headache does not influence mine.
- Ways in which it will fail?

Assumption

(SUTVA) The potential outcome for any unit do not vary with the treatments assigned to other units and , for each unit, there are no different forms or versions of each treatment unit leading to different outcomes.

- No interference:
- What this means ?
- Examples where it will fail
  - Immunization
  - Help with job applications for unemployed.
- Ways to solve these problem?

The importance of the Assignment Mechanism

From now on assume SUTVA holds. Then the aspirin example simplifies to two situations for Sara and I: each of us can either take or not take aspirin, and what the other does is not relevant. This extends to many units.

Notation: Let's assume we have a population of size N, indexed by i taking values $1 \dots N$

Define :

$Y_i^{obs} = Y_i(W_i) = Y_i(0)$ if $W_i = 0$ and $Y_i(1)$ if $W_i = 1$

$Y_i^{miss} = Y_i(1 - W_i) = Y_i(1)$ if $W_i = 0$ and $Y_i(0)$ if $W_i = 1$

Causal effect for person i: $Y_i(1) - Y_i(0)$

Missing data problem: we only see Y_i^{obs} so we cannot calculate the treatment effect for each person.

We will try to infer something about Y_i^{miss} from the data we do observe, but in doing that, knowing *the assignment mechanism* will be essential: why did some people end up treated and other not?

The selection problem

Imagine we have a larger group of people who took aspirin, and a group who did not, and we decide to take the sample mean of people of headache for people who got or did not get the pill.
By property of the sample mean, we know that this is a good estimator for: $E[Y_i|W_i = 1] - E[Y_i|W_i = 0]$
Let's think about what this is:

$$E[Y_i^{obs}|W_i = 1] - E[Y_i^{obs}|W_i = 0] = E[Y_i(1)|W_i = 1] - E[Y_i(0)|W_i = 0]$$

$$E[Y_i(1)|W_i = 1] - E[Y_i(0)|W_i = 1] + E[Y_i(0)|W_i = 1] - E[Y_i(0)|W_i = 0]$$

$E[Y_i(1)|W_i = 1] - E[Y_i(0)|W_i = 1]$ treatment effect on the treated
$E[Y_i(0)|W_i = 1] - E[Y_i(0)|W_i = 0]$ selection bias

The selection problem

$E[Y_i(0)|W_i = 1] - E[Y_i(0)|W_i = 0]$ is the selection effect

- People take headache pills because their headache is pretty bad.
- People who go to college have all sorts of attributes that are different than people who do not.
- When does the bias disappear?

Randomization solve the selection problem

In a completely randomized experiment, N_t unit are randomly drawn to be in the treatment group, and N_c units are drawn to be in the control group.
Then, the probability of assignment does not depend on potential outcomes: $E[Y_i(0)|W_i = 1] - E[Y_i(0)|W_i = 1] = 0$ and

$$E[Y_i^{obs}|W_i = 1] - E[Y_i^{obs}|W_i = 0] = E[Y_i(1)|W_i = 1] - E[Y_i(0)|W_i = 1]$$

$$E[Y_i(1) - Y_i(0)|W_i = 1] = E[Y_i(1) - Y_i(0)]$$

Types of RCT

1. Completely randomized
2. Stratified randomization : Blocks of come covariate X are created, randomization is done within each block
3. Pairwise randomization: Pairs are created, randomization is done within each pair.
4. Clustered randomization: Units are not individual, but groups of individuals (e.g. classrooms)
5. The result above hold whatever the type of randomization. why?
6. Why would we prefer a type of design rather than another?

Back to our examples

Let's think of an RCT that would make sense for each of these questions:

- Aspirin
- College
- Race and jobs

Aspirin

This is the traditional medical RCT. Completely randomized experiment, individual level randomization.

College

- We may not be able to randomly assign college going: some people may not want to go to college no matter what, and some people will want to go no matter what.
- What could we do?
- A scholarship program would be an encouragement, and ensure that some people are *more likely* than other to go to college.
- The scholarship program itself can be analyzed as a simple randomized experiment.
- But this is not the impact of going to college!
- This is called an *encouragement design*. The analysis of these types of experiment is left for later in the semester.

Race

- We need to start by refining the question.
- Are we meaning to say that African American are in general disadvantaged?
- Or specifically that they are discriminated against by potential employers, at given characteristics?
- If that is the case we can manipulate race perception without changing anything else
  - Audit studies : perception of race is manipulated by sending pairs of actors with similar characteristics except race
  - Resume studies: perception of race is manipulated by sending resumes with identical characteristics except the extent to which the name sounds African American: this can be analyzed as stratified randomized experiments.

References

- Imbens and Rubin *Causal Inference for Statistics Social and biomedical Sciences*