IDS 702: Module 6.1

The potential outcomes framework and causal estimands

Dr. Olanrewaju Michael Akande

CAUSALITY

> We do not have knowledge of a thing until we have grasped its why, that is to say, its cause.

-- Aristotle, Physics

- Over the next few modules, we will discuss causal inference, specifically, on measuring the effects of causes.
- For now, we will simply lay the foundations for causal inference.
- We will get more into the actual methods later.

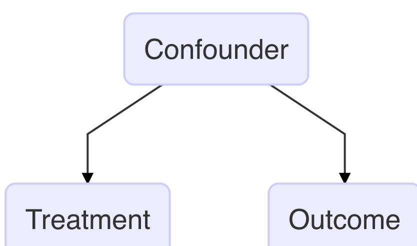
Association vs. causation

- In the models we have covered so far, our focus has been on inferring associations using samples drawn from our population of interest.
- For example, we have been asking questions such as, do people who receive job training tend to earn more wages than people who do not?
- Causal inference goes further as we try to infer aspects of the actual data generating process, that is, causation.
- For example, does receiving job training actually cause one to earn more wage than they would have without the training?
- The additional information needed to move from association to causation is often provided by causal assumptions (often untestable).
- Note: in most cases, association does not imply causation!

Confounding

- Why is it that association does not often imply causation? confounding variables or confounders!
- Causal relationship

- Confounding



Examples of confounding

- Ice cream consumption and number of people who drowned.
  Confounder: temperature; people tend to consume more ice cream and also swim more when it is hot.
- Medical treatment and patient outcome.
  Confounders: age, sex, other complications
- Education and income.
  Confounder: socio-economic status of family
- An extreme example of confounding is Simpson's paradox: where a confounder reverses the sign of the correlation between treatment and outcome

IDS 702

Simpson's paradox

- Example: kidney stone treatment (Charig et al., BMJ, 1986).
    - Compare the success rates of two treatments for kidney stones
    - Treatment A: open surgery. Treatment B: small puncture

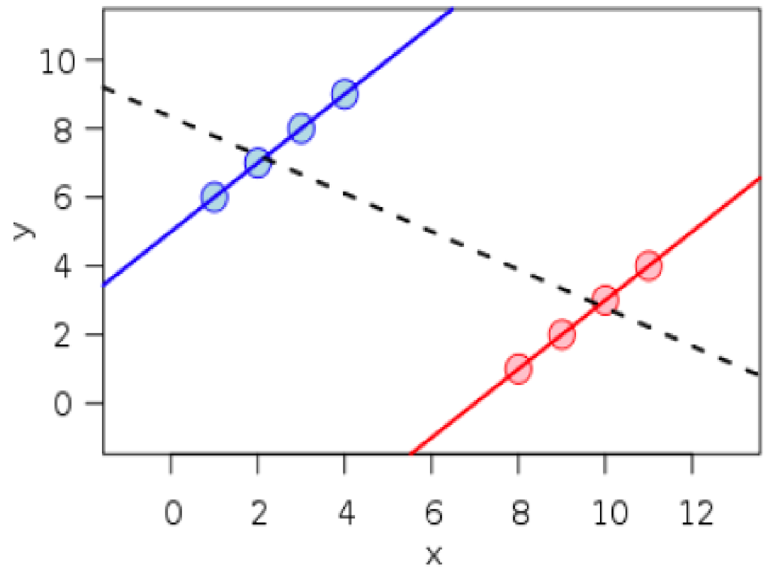
	Treatment A	Treatment B
Small stones	**93%** (81/87)	87% (234/270)
Large stones	**73%** (192/263)	69% (55/80)
Both	78% (273/350)	**83%** (289/350)

- Overall treatmnent B has a higher success rate but treatment A actually has higher success rates given the type of stones.
- What is the confounder here? Severity of the case/type of stones.

IDS 702

Simpson's paradox or Yule-Simpson effect

- Simpson's paradox: a trend appears in different groups of data but disappears or reverses when these groups are combined.



- Mathematically, it is about conditioning.
- Another well-known example is the Berkeley admission gender bias (Bickel et al., Science, 1976).

IDS 702

General notation

- W: Treatment (e.g. job training); we will focus on binary treatments.
- Y: Outcome (e.g. annual wages).
- X: Observed predictors or confounders (e.g. age, education, etc).
- U: Unobserved predictors or confounders.

- Examples of causal questions:
  - Causal effect of exposure to a disease.
  - Comparative effectiveness research such as in clinical trials: whether one drug or medical procedure is better than the other.
  - Program evaluation in economics and policy.

IDS 702

Potential outcomes framework

Potential outcomes framework

- The potential outcomes framework or counterfactual framework or Rubin Causal Model (RCM) is arguably the most widely used framework across many disciplines, e.g., medicine, health care, policy, social sciences.
- Under this framework, causal inference is viewed as a problem of missing data with explicit mathematical modeling of the assignment mechanism as a process for revealing the observed data.
- Rooted in the statistical work on randomized experiments by Fisher (1918, 1925) and Neyman (1923), as extended by Rubin (1974, 1976, 1977, 1978, 1990).

Potential outcomes framework

- For a binary treatment, each individual gets exactly one of the two options, and we observe the corresponding response for that.
- Conceptually, under the potential outcomes framework, we think about what each individual's response should have been had they gotten the other treatment option instead of the one they actually got.
- The individual causal effect then is the difference between the two "potential" outcomes, only one of which is observed.
- Clearly, we never observe the two potential outcomes for any individual, making it natural to think of this as a missing data problem.

IDS 702

Potential outcomes framework

- No causation without manipulation - "cause" must be (hypothetically speaking) something we can manipulate. e.g., intervention, action, treatment.
- That is, gender, time and age are not well defined “causes" under the RCM.
- Three integral components of the potential outcomes framework:
  - potential outcomes corresponding to the various levels of a treatment.
  - assignment mechanisms, that is, the treatment indicator for all observations.
  - a model for the science (the potential outcomes and covariates).

IDS 702

Potential outcomes framework: basic concepts

- Unit: The person, place, or thing upon which a treatment will operate, at a particular time (note: a single person, place, or thing at two different times comprises two different units).

- Treatment: An intervention, the effects of which (on some particular measurement of the units) the investigator wishes to assess relative to no intervention (i.e., the control).

- Potential Outcomes: The values of a unit's measurement of interest after (a) application of the treatment and (b) non-application of the treatment (i.e., under control).

- Causal Effect: For each unit, the comparison of the potential outcome under treatment and the potential outcome under control.

IDS 702

Causal effects

- For a single unit, let $Y(0)$ denote the outcome given the control treatment and $Y(1)$, the outcome given the active treatment.
- For example, suppose Y denotes a score (level of severity) for headache, then for a single unit, we could have

Raw scores

Unit	Initial headache	Potential outcomes		Causal effect
	X	Y(asp)	Y(not)	Y(asp) - Y(not)
you	80	25	75	-50

Gain scores

Unit	Initial headache	Potential outcomes		Causal effect
	X	Y(asp) - X	Y(not) - X	[Y(asp) - X] - [Y(not) - X]
you	80	-55	-5	-50

IDS 702

The fundamental problem of causal inference

As mentioned before,

- The fundamental problem of causal inference: we can observe at most one of the potential outcomes for each unit.
- Causal inference under the potential outcome framework is essentially a missing data problem.
- To identify causal effects from observed data, under any mathematical framework, one must make assumptions (structural or/and stochastic)
- Since we can at most observe a single potential outcome, we must rely on multiple units (and a lot of assumptions) to make causal inferences.

Basic setup

- Target population: a well-defined population of individuals whose outcomes are going to be compared
- Data: a random sample of N units from a target population.
- A treatment with two levels: $w = 0, 1$.
- For each unit i, we observe
  - the binary treatment status $W_i \in \{0, 1\}$,
  - a vector of p predictors/covariates $X_i = (X_{i1}, \ldots, X_{ip})$, and
  - an outcome Y_i^{obs}.

IDS 702

Basic setup

- For each unit i, there are two potential outcomes $(Y_i(0), Y_i(1))$.
- That is, the outcomes under the two values of the treatment, at most one of which is observed.
- Potential outcomes and assignments jointly determine the values of the observed outcomes

$$Y_i^{\text{obs}} \equiv Y_i(W_i) = W_i \cdot Y_i(1) + (1 - W_i) \cdot Y_i(0)$$

and the missing outcomes:

$$Y_i^{\text{mis}} \equiv Y_i(1 - W_i) = (1 - W_i) \cdot Y_i(1) + W_i \cdot Y_i(0)$$

IDS 702

Causal estimands

- The average treatment effect (ATE):

$$\tau = \mathbb{E}[Y_i(1) - Y_i(0)].$$

- The average treatment effect for the treated (ATT):

$$\tau = \mathbb{E}[Y_i(1) - Y_i(0)|W_i = 1].$$

- The average treatment effect for the control (ATC):

$$\tau = \mathbb{E}[Y_i(1) - Y_i(0)|W_i = 0].$$

- For binary outcomes, causal odds ratio (OR) or risk ratio (RR)::

$$\tau = \frac{\mathbb{P}\mathrm{r}[Y_i(1) = 1]/\mathbb{P}\mathrm{r}[Y_i(1) = 0]}{\mathbb{P}\mathrm{r}[Y_i(0) = 1]/\mathbb{P}\mathrm{r}[Y_i(0) = 0]}.$$

- Obviously these estimands are not identifiable without further assumptions.

- We will start to explore those soon.

IDS 702

Example

	Potential Outcomes			Observed Data		
	Y(0)	Y(1)		W	Y(0)	Y(1)
	13	14		1	?	14
	6	0		0	6	?
	4	1		0	4	?
	5	2		0	5	?
	6	3		0	6	?
	6	1		0	6	?
	8	10		1	?	10
	8	9		1	?	9
True averages	7	5	Observed averages		5.4	11

IDS 702

Acknowledgements

These slides contain materials adapted from courses taught by Dr. Fan Li.

IDS 702

What's next?

Move on to the readings for the next module!

IDS 702