Segment 1: Fundamentals of Causal Inference

Section 09: The Assignment Mechanism

Observed-Data Comparison $\stackrel{?}{=}$ Causal Effect

Table: Potential Outcomes

Unit, i	Treatment Z_i	Potential Outcome, Y_i^c	Potential Outcome, Y_i^t
Audrey	0	140	135
Anna	0	140	135
Bob	0	150	140
Bill	0	150	140
Caitlin	1	160	155
Cara	1	160	155
Dave	1	170	160
Doug	1	170	160

Table: Observed Data

Unit, i	Treatment Z_i	Potential Outcome, Y_i^c	Potential Outcome, Y_i^t	Observed Outcome, Y_i
Audrey	0	140	?	140
Anna	0	140	?	140
Bob	0	150	?	150
Bill	0	150	?	150
Caitlin	1	?	155	155
Cara	1	?	155	155
Dave	1	?	160	160
Doug	1	?	160	160

The Assignment Mechanism

- **Key** feature of causal inference methods relates to *known* or *assumed* features of the mechanism that dictates which treatments are assigned
  - $\Rightarrow$ which potential outcomes are revealed
  - This is called the *assignment mechanism*
- That is, the *assignment mechanism* plays a key role in determining which (if any) observed-data comparisons can be interpreted as estimates of causal effects
- Example: Randomized experiments imply certain types of assignment mechanisms

Two Components of the Assignment Mechanism

"Assignment mechanism" is sometimes used to describe the combination of two mechanisms:

1. *Sampling Mechanism* governing which experimental units in the population are sampled for study
2. *Treatment Assignment Mechanism* governing which of the sampled units receive each treatment
   - We will focus mainly the treatment assignment mechanism, even when using the term *assignment mechanism*

Assignment Mechanism

For a population of N units and binary $Z \in \{0,1\}$

A probability function,

$$Pr(\mathbf{Z}|\mathbf{X}, \mathbf{Y}^c, \mathbf{Y}^t)$$

taking values in $[0,1]^N$ satisfying

$$\sum_{\mathbf{Z}\in[0,1]^N} Pr(\mathbf{Z}|\mathbf{X}, \mathbf{Y}^c, \mathbf{Y}^t) = 1$$

for all $\mathbf{X}, \mathbf{Y}^c, \mathbf{Y}^t$

- Mapping of each possible treatment assignment vector to a probability, given "The Science"
    - I.e., can depend on the unobserved potential outcomes
    - Z is the only random variable!
- Describes the reasons for the missing and observed values of Y^t and Y^c

Assignment Mechanism

We will focus on *individualistic* assignment mechanisms

$$Pr(\mathbf{Z}|\mathbf{X}, \mathbf{Y}^c, \mathbf{Y}^t) = c \prod_i p_i(Z_i|X_i, Y_i^c, Y_i^t)$$

For example, if Z_i is binary,

Note: Most key assumptions for causal inference are assumptions about the assignment mechanism!

Known vs. Unknown Assignment Mechanism

1. Experimental Studies
   - In a prospectively-designed experiment, the assignment mechanism is *known*
   - I.e., the experimenter gets to decide the rule for assigning treatments to units
   - These decisions can be encoded in an assignment mechanism
   - Notions of causal inference can be directly linked to the known assignment mechanism

2. Observational Studies
   - With data outside of the conduct of experiments, the assignment mechanism is *unknown*
   - The reasons for certain units receiving certain treatments are not under experimenter control
   - Observational data
   - Notions of causal inference can be linked to *assumptions* about the unknown assignment mechanism

Example: Perfect Doctor

Table: The "Science" Representing the Underlying Values of All Potential Outcomes

	Potential Outcomes		
	Y^c	Y^t	$Y^t - Y^c$
	13	14	1
	6	0	-6
	4	1	-3
	5	2	-3
	6	3	-3
	6	1	-5
	8	10	2
	8	9	1
True Average:	7	5	-2

Example: Perfect Doctor

Table: The “Science” Representing the Underlying Values of All Potential Outcomes

	Potential Outcomes		
	Y^c	Y^t	$Y^t - Y^c$
	13	14	1
	6	0	-6
	4	1	-3
	5	2	-3
	6	3	-3
	6	1	-5
	8	10	2
	8	9	1
True Average:	7	5	-2

The Assignment Mechanism will produce a value of $\mathbf{Z}$ that dictates which units receive which treatment and which potential outcomes are “revealed”

Example assignment mechanism: A gifted doctor only treats patients who will definitely benefit

Example: Perfect Doctor

Table: Observed Data After Treatment Assignment

	Potential Outcomes		
Z	Y^c	Y^t	$Y^t - Y^c$
t	?	14	?
c	6	?	?
c	4	?	?
c	5	?	?
c	6	?	?
c	6	?	?
t	?	10	?
t	?	9	?
Observed Average:	5.4	11	

Example assignment mechanism: A gifted doctor only treats patients who will definitely benefit

Example: Perfect Doctor

Table: Implicitly-Assumed Missing Values

	Potentail Outcomes		
Z	Y^c	Y^t	$Y^t - Y^c$
t		14	
c	6		
c	4		
c	5		
c	6		
c	6		
t		10	
t		9	
Observed Average:	5.4	11	5.6

- True effect is -2
- Observed "effect" with this assignment mechanism is 5.6
- What went wrong?

A "Bad" Assignment Mechanism

- Determined by doctor's intuition about which patients would benefit
- That is, the assignment mechanism depended on presumed knowledge of Y_i^c and Y_i^t for each i

$\Rightarrow$ There was some systematic difference between the types of patients who were treated vs. not

$\Rightarrow$ Outcomes in the $Z = c$ group were not a "good substitute" for the unobserved potential outcomes in the $Z = t$ group.

$\Rightarrow$ Comparing outcomes in treatment groups did not reflect the true causal effect

Other "Bad" Assignment Mechanisms

- Doctors treat the sickest patients
- Older people tend to respond to certain advertisements
- People who take supplements are more likely to care about their health generally
- ...

In all cases, certain types of units will tend to have certain values of (Y_i^c, Y_i^t):

$\Rightarrow$ Observed data comparisons will misattribute these differences to a treatment effect

Key Idea: *Balance* Across Treatment Groups

Key Idea: Comparisons between observed outcomes between treated and control must be "balanced" with respect *both* potential outcomes

We can never see these....

1. Ensure balance by specifying a particular type of assignment mechanism
   - E.g., completely randomized study
2. Use pre-treatment covariates as *proxies* for knowledge of both potential outcomes
   - E.g., If women > 55years with no history of disease are expected to respond to treatment a certain way, ensure balance on sex and age

Example: Hypothetical Dietary Experiment

Potential Outcomes

Table: Potential Outcomes from the Hypothetical Dietary Experiment

Unit, i	Female, x_{1i}	Age, x_{2i}	Treatment Z_i	Potential Y_i^c	Potential Y_i^t
Audrey	1	40	0	140	135
Anna	1	40	0	140	135
Bob	0	50	0	150	140
Bill	0	50	0	150	140
Caitlin	1	60	1	160	155
Cara	1	60	1	160	155
Dave	0	70	1	170	160
Doug	0	70	1	170	160

Example: Hypothetical Dietary Experiment

Observed Data

Table: Observed Data from the Hypothetical Dietary Experiment

Unit, i	Female, x_{1i}	Age, x_{2i}	Treatment Z_i	Potential Y_i^c	Potential Y_i^t
Audrey	1	40	0	140	?
Anna	1	40	0	140	?
Bob	0	50	0	150	?
Bill	0	50	0	150	?
Caitlin	1	60	1	?	155
Cara	1	60	1	?	155
Dave	0	70	1	?	160
Doug	0	70	1	?	160

Example: Hypothetical Dietary Experiment

Observed Data

Table: Observed Data from the Hypothetical Dietary Experiment, **Balanced Design**

Unit, i	Female, x_{1i}	Age, x_{2i}	Treatment Z_i	Potential Y_i^c	Potential Y_i^t
Audrey	1	40	0	140	?
Anna	1	40	1	?	135
Bob	0	50	0	150	?
Bill	0	50	1	?	140
Caitlin	1	60	0	160	?
Cara	1	60	1	?	155
Dave	0	70	0	170	?
Doug	0	70	1	?	160

Balance and the Assignment Mechanism

In general, it is impossible to ensure that treatment groups are comparable based on all relevant potential outcomes or pre-treatment unit characteristics

- We may balance on many *observed* pre-treatment characteristics, but how do we know we *observed* everything that matters?

Knowledge of (or assumptions about) **the assignment mechanism** can serve as a guide about whether we believe units will be (approximately) balanced

Unconfounded Assignment Mechanisms

Recall that the *assignment mechanism*

$$Pr(\mathbf{Z}|\mathbf{X}, \mathbf{Y}^c, \mathbf{Y}^t) = c \prod_i p_i(Z_i|X_i, Y_i^c, Y_i^t)$$

is the probability rule dictating how the treatments are assigned and observed data are "revealed" by the assignment to treatments

- "Bad" assignment mechanisms produce $\mathbf{Z}$ that lead to imbalance in (Y_i^c, Y_i^t)
  - That is, treatments are assigned with (proxy) knowledge of (Y_i^c, Y_i^t)
  - E.g., the perfect doctor
- But what if the assignment mechanism *didn't* depend on (Y_i^c, Y_i^t)?
  - That is, what if $\mathbf{Z}$ were assigned to intentionally *not* relate to the potential outcomes?

Unconfounded Assignment Mechanisms

An assignment mechanism is *unconfounded* (conditional on $\mathbf{X}$) if it does not depend on *any* potential outcomes

$$Pr(\mathbf{Z}|\mathbf{X}, \mathbf{Y}^c, \mathbf{Y}^t) = Pr(\mathbf{Z}|\mathbf{X}) = c \prod_i p_i(Z_i|X_i)$$

- Now, we would not expect any systematic imbalance in (Y^c, Y^t) between those who receive $Z = 1$ or 0
- Observed-data comparisons would be "balanced" and thus more capable of estimating causal effects
  - At least on average...

Note: For the purposes of this course, *unconfounded* will sometimes be described as *ignorable*

The Importance of Recording the Assignment Mechanism

If the assignment mechanism depends on [certain] values, then those values must be recorded...if the assignment mechanism is to be ignorable. For example...if a doctor assigns patients to treatments so as to balance the distribution of background variables such as age and sex, these variables must be recorded if the assignment mechanism is to be ignorable. If a doctor assigns treatments according to his unrecorded judgments about the health of patients, the assignment mechanism is not ignorable. Or if patients select the treatments themselves on the basis of their unrecorded opinions of their health, the assignment mechanism is not ignorable....The more involved the assignment mechanism (in the sense of depending on more values), the more complete must be the recording mechanism if the assignment mechanism is to be ignorable. - Rubin (1978)

Unconfounded (Ignorable) Study Designs

- If we *prospectively design a study* to estimate causal effects, the rule for assigning units to treatments should be unconfounded
  - However we make the decisions to assign treatments to units, we have to ensure that it cannot depend on (Y_i^c, Y_i^t)
  - You probably already know one way to do this....

Unconfounded (Ignorable) Study Designs

- If we *prospectively design a study* to estimate causal effects, the rule for assigning units to treatments should be unconfounded
  - However we make the decisions to assign treatments to units, we have to ensure that it cannot depend on (Y_i^c, Y_i^t)
  - You probably already know one way to do this....

- If we have data that has already been collected and want to *retrospectively* analyze it for causal effects, we have to know whether rule for assigning units to treatments was unconfounded
  - This will generally be unknown
  - Will have to assume whether the observed data can approximately represent an unconfounded assignment mechanism