Potential Outcomes

Kosuke Imai

Harvard University

STAT186/GOV2002 CAUSAL INFERENCE

Fall 2018

Announcements

- Things to do asap if you are taking this course for credit:
  1. Please sign up for Piazza
  2. Take Google form survey
  3. Those who are submitting a petition to enroll in this class should state a reason using the “Comment” functionality
- Use of electronic devices is allowed only in a designated section

Defining Causal Effects

- Units: $i = 1, \ldots, n$
- Treatment: $T_i = 1$ if treated, $T_i = 0$ otherwise
- Observed outcomes: Y_i
- Pre-treatment covariates: $\mathbf{X}_i$
- Potential outcomes: $Y_i(1)$ and $Y_i(0)$ where $Y_i = Y_i(T_i)$

Voters	Contact	Turnout		Age	Gender
i	T_i	$Y_i(1)$	$Y_i(0)$	X_{i1}	X_{i2}
1	1	1	?	20	M
2	0	?	0	55	F
3	0	?	1	40	F
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	1	0	?	62	M

- Causal effect for unit i:

$$\tau_i = Y_i(1) - Y_i(0)$$

Fundamental Problem of Causal Inference

- Any causal quantity is a function of potential outcomes:

$$\log Y_i(1) - \log Y_i(0), \quad \frac{Y_i(1)}{Y_i(0)}, \quad \frac{Y_i(1) - Y_i(0)}{Y_i(0)} \times 100, \quad etc.$$

- Causal inference as a missing data problem:
  - potential outcomes are thought to be fixed for each unit
  - potential outcomes as "attributes"
  - potential outcomes do have a distribution across units
  - treatment variable determines which potential outcome is observed
  - observed outcomes are random because the treatment is random
- Non-binary treatment:
  - categorical: $Y_i(0), Y_i(1), \ldots, Y_i(K-1)$
  - continuous: $Y_i(t)$ for any $t \in \mathbb{R}$
- Only one potential outcome is observed for each unit

The Key Assumptions

- The notation implies three assumptions:
  1. Causal ordering: Y_i cannot causally affect T_i
  2. No interference between units:

$$Y_i(T_1, T_2, \ldots, T_n) \;=\; Y_i(T_i)$$

  3. Same version of the treatment
- Stable Unit Treatment Value Assumption (SUTVA)

- Potential violations:
  1. feedback effects $\rightsquigarrow$ need more frequent observations
  2. spillover effects $= Y_i(T_i = t, \mathbf{T}_{-i} = \mathbf{t}) - Y_i(T_i = t, \mathbf{T}_{-i} = \mathbf{t}')$
     vs. direct effects $= Y_i(T_i = 1, \mathbf{T}_{-i} = \mathbf{t}) - Y_i(T_i = 0, \mathbf{T}_{-i} = \mathbf{t})$
  3. multiple versions of treatment $\rightsquigarrow$ redefine them as separate treatments, assume treatment-then-version sequence

Causal Effects of Immutable Characteristics

- "No causation without manipulation" (Holland, 1986. *J. Amer. Stat. Assoc*)
- Immutable characteristics or attributes: gender, race, age, etc.
- Can immutable characteristics have meaningful causal effects?
- Strategies:
  1. Causal effects of perceived characteristics:
     - Causal effect of a job applicant's gender/race on call-back rates (Bertrand and Mullainathan, 2004. *Am. Econ. Rev*)
  2. Reinterpretation:
     - Causal effect of having a female politician on policy outcomes (Chattopadhyay and Duflo, 2004. *Q. J. Econ*)
  3. Redefinition:
     - Race as a "bundle of sticks": skin color, genes, neighborhood, socio economic status, etc. (Sen and Wasow, 2016. *Annu. Rev. Polit. Sci*)
  4. Group-level inference:
     - Would racial disparity go away if we equalize socio-economic status of blacks and whites? (VanderWeele and Robinson, 2014. *Epidemiology*)

Average Treatment Effects

- Unit causal effects are difficult to estimate
- We can average them over a sample of units
  1. sample average treatment effect:

$$\text{SATE} = \frac{1}{n}\sum_{i=1}^{n} Y_i(1) - Y_i(0)$$

  2. sample average treatment effect for the treated

$$\text{SATT} = \frac{1}{n_1}\sum_{i=1}^{n} T_i(Y_i(1) - Y_i(0)) \quad where \quad n_1 = \sum_{i=1}^{n} T_i$$

- Population average treatment effects:

$$\begin{aligned} \text{PATE} &= \mathbb{E}(Y_i(1) - Y_i(0)) \\ \text{PATT} &= \mathbb{E}(Y_i(1) - Y_i(0) \mid T_i = 1) \end{aligned}$$

Other Causal Quantities of Interest

- Heterogenous effects:
  - Conditional average treatment effect (CATE)

$$\tau(\mathbf{x}) \; = \; \mathbb{E}(Y_i(1) - Y_i(0) \mid \mathbf{X}_i = \mathbf{x})$$

  - Applications to precision medicine and micro targeting
- Non-additive effects:
  - Quantile treatment effects

$$Q_{Y_i(1)}(\alpha) - Q_{Y_i(0)}(\alpha) \quad \text{where } Q_{Y_i(t)}(\alpha) = \inf\{y \in \mathbb{R} : \alpha \leq F_{Y_i(t)}(y)\}$$

    This is different from $Q_{Y_i(1)-Y_i(0)}(\alpha)$
  - Odds ratio

$$\frac{\Pr(Y_i(1) = 1)/\Pr(Y_i(1) = 0)}{\Pr(Y_i(0) = 1)/\Pr(Y_i(0) = 0)}$$

Truncation by Death

- Setup
  1. Units: patients
  2. Treatment: new medicine
  3. Outcome: cholesterol level
  4. Truncation: patient death
- Truncation by death problem (Zhang and Rubin, 2003. *J. Educ. Behav. Stat.*):
  - cholesterol level (test score) undefined for the dead
  - survivors in the treament group are not comparable to those in the control group
  - Post-treatment bias: treatment may affect survival
- In general, one should not adjust for post-treatment variables
- Other examples:
  - drop-out in program evaluation
  - utilization when evaluating customer satisfaction
  - registration, turnout, and vote choice in get-out-the-vote studies

Principal Stratification (Frangakis and Rubin, 2002. *Biometrics*)

- Potential truncation variable: $W_i(1), W_i(0)$
- Observed truncation variable: $W_i = W_i(T_i)$
- Potential outcomes: $Y_i(t, w) \rightsquigarrow Y_i(0,1)$ and $Y_i(1,1)$ do not exist
- Observed outcome: $Y_i = Y_i(T_i, W_i(T_i))$ for $W_i = 0$
- Four principal strata defined by $(W_i(0), W_i(1))$

	$W_i = 1$	$W_i = 0$
$T_i = 1$	(0,1) and (1,1)	(0,0) and (1,0)
$T_i = 0$	(1,0) and (1,1)	(0,0) and (0,1)

- ATT for "always-survivors": $\mathbb{E}(Y(1) - Y(0) \mid W(1) = W(0) = 0)$

Necessary and Sufficient Causes

- Effects of causes $\leftrightsquigarrow$ Causes of effects
- Prospective causal effects $\leftrightsquigarrow$ Retrospective causal effects
- Latter is more difficult than former

- Notion of necessary and sufficient causes
- Example (Geddes, 1990, *Political Anal*): Is village autonomy necessary and sufficient for revolution?

	Revolution	No Revolution
Village Autonomy	Russia France China, in area controlled by Communists	
Village Dependent		Britain, 1640–60 Germany, 1848 China, before Communists

Probabilistic Causal Necessity and Sufficiency

(Pearl, 2000. *Causality*, Cambridge UP)

- Probability of causal necessity: village autonomy must be present in order for revolution to occur

$$\text{PN} \;=\; \Pr(Y_i(0) = 0 \mid T_i = 1, Y_i = 1)$$

- Probability of causal sufficiency: the presence of village autonomy gurantees the occurance of revolution

$$\text{PS} \;=\; \Pr(Y_i(1) = 1 \mid T_i = 0, Y_i = 0)$$

- Probability of causal necessity and sufficiency

$$\text{PNS} \;=\; \Pr(Y_i(1) = 1, Y_i(0) = 0)$$

- Relationship:

$$\text{PNS} \;=\; \Pr(Y_i = 1, T_i = 1) \times \text{PN} + \Pr(Y_i = 0, T_i = 0) \times \text{PS}$$

Experimental and Observational Studies

- Two types of studies:
  1. Randomized experiments: randomized & controlled intervension
     - Laboratory experiments
     - Survey experiments
     - Field experiments
  2. Observational studies: no intervension

- Tradeoff between internal and external validity
  - Confounding: endogeneity, omitted variables, selection bias
  - Generalizability: sample selection, Hawthorne effects, realism

- "Designing" observational studies
  - Natural experiments $\rightsquigarrow$ haphazard treatment assignment
  - Examples: birthdays, weather, close elections, arbitrary administrative rules, etc.

- Generalizing experimental results $\rightsquigarrow$ possible extrapolation

Summary

- Causal effects $\rightsquigarrow$ function of potential outcomes
  - Stable Unit Treatment Value Assumption
  - feedback effects, spillover effects, multiple versions of treatment
- Fundamental problem of causal inference: only one potential outcome is observed
- Causal inference as a missing data problem
- Examples of potential outcomes:
  - principal stratification
  - necessary and sufficient causes
- Suggested readings:
  1. IMBENS AND RUBIN, Chapter 1
  2. HOLLAND, P. (1986). "STATISTICS AND CAUSAL INFERENCE." *J. Am. Stat. Assoc*