EMORY | COLLEGE OF ARTS AND SCIENCES

Potential outcomes

Intro. to causal inference | SPSP 2023 Annual Convention

Wen Wei Loh (wen.wei.loh@emory.edu)
Department of Quantitative Theory and Methods (QTM)

February 25, 2023

What are potential outcomes?

Average causal effect

Conditional exchangeability and no unmeasured confounding

Propensity score

Other applications building on the potential outcomes framework

What are potential outcomes?

Point Treatment

- Focus today: Treatment at a single specific time point.
- *Treatment can be manipulated or naturally observed.*
- E.g., what is the effect of
  - awe on humility;
  - bias habit-breaking intervention on hiring practices;
  - misinformation inoculation training on resistance to falsehoods;
  - social contact frequency on well-being.
- Treatment should be conceptually and practically manipulable.

Observed vs. potential outcomes

- Participants in a post-lunch Saturday workshop each suffered from tiredness.
- A few drank a coffee; others did not.
- One hour later, each participant reported how tired they felt on a scale of 1 - 7.

Observed vs. potential outcomes

	Had a coffee? (1=Yes)	Level of tiredness? (1 - 7)
A	0	6
B	1	3
C	1	2
D	0	4
E	0	1
F	1	5

- C had a coffee and was not tired an hour later.
- Did the coffee cause her to be less tired?

Observed vs. potential outcomes

- We do not know.
- To answer this, we need to know what would have happened had C not taken the coffee.
- In other words, we need to know what each participant's outcome would have been under a condition different from the observed.

Potential outcomes

For each individual:

- X: treatment (1) or control (0)
- Y: observed outcome
- $Y(1)$: potential outcome under treatment
- $Y(0)$: potential outcome under control
- Exactly one of $Y(1)$ or $Y(0)$ can be observed or revealed
  - $X = 1 \quad \Rightarrow \quad Y = Y(1)$;
  - $X = 0 \quad \Rightarrow \quad Y = Y(0)$.
- The other is counterfactual.
- Aside: notation $Y^1, y(1), y^1$, etc. also possible.

Potential outcomes

	Had a coffee? (1=Yes)	Y(1)	Y(0)
A	0	??	6
B	1	3	??
C	1	2	??
D	0	??	4
E	0	??	1
F	1	5	??

Fundamental problem of causal inference

- $Y(1) - Y(0)$: Individual treatment effect.
- Cannot observe both $Y(1)$ and $Y(0)$ for the same individual in reality.
- Recall C had an observed outcome of $Y = Y(1) = 2$.
- Suppose C had a counterfactual outcome of $Y(0) = 4$.
- Did the coffee have an individual causal effect?
- What would the individual causal effect have been if $Y(0) = 2$ instead?

Fundamental problem of causal inference

- Not knowing both $Y(1)$ and $Y(0)$ for the same individual is termed the fundamental problem of causal inference [Holland, 1986].
- Potential outcomes framework also commonly referred to as *Neyman-Rubin Causal Model* [Splawa-Neyman et al., 1990, Rubin, 1990].

Assumptions: Causal consistency and SUTVA

Causal or counterfactual consistency assumption [VanderWeele, 2009]:

- Observed outcome can be determined from the potential outcomes and the observed treatment as:

$$Y = Y(0)(1 - X) + Y(1)X.$$

Assumptions: Causal consistency and SUTVA

SUTVA (stable unit treatment value assumption) [Imbens and Rubin, 2015]:

- *well-defined* treatment and control conditions: "no hidden treatment variations" among individuals experiencing the same condition; and
- each individual's potential outcomes are unaffected by *treatments of other individuals* so that there is "no (treatment) interference."

Activity!

- Together with your neighbor(s), think of an example of a possible treatment and outcome.
- Don't use something familiar from your own work – consider content from elsewhere.
- Discuss how SUTVA may be violated, and suggestions to better justify SUTVA.

Average causal effect

Population-level causal effect

- Fundamental problem of causal inference makes calculating individual causal effects challenging.
- Instead, target the population-level or average causal effect:

$$\mathrm{E}[Y(1) - Y(0)] = \mathrm{E}[Y(1)] - \mathrm{E}[Y(0)].$$

- Statistics: $\mathrm{E}[Z]$ denotes the population expectation or average of the random variable Z.

Population-level causal effect

- $\mathrm{E}[Y(1)] - \mathrm{E}[Y(0)]$ is an estimand.
- An estimand is conceptual – reflects scientific knowledge and query.
- E.g., difference in average tiredness had everyone drank coffee vs. had everyone not drank coffee.

- *Can we use the sample average of $Y(1)$ to estimate $\mathrm{E}[Y(1)]$, or the sample average of $Y(0)$ to estimate $\mathrm{E}[Y(0)]$?*

Causation vs. association

- Alternatively, the conditional difference in observed outcomes is:

$$\mathrm{E}[Y|X=1] - \mathrm{E}[Y|X=0]$$

- Statistics: $\mathrm{E}[Z|C=c]$ denotes the population conditional expectation or average of the random variable Z, among those with covariate $C=c$.

Causation vs. association

- $\mathrm{E}[Y|X=1] - \mathrm{E}[Y|X=0]$ is an estimator.
- An estimator is empirical – statistical function of observed data.
- E.g., difference in average tiredness among those who actually drank coffee vs. among those who did not drink coffee.

Causation vs. association

- Potential outcome notation helps distinguish causation

$$E[Y(1)] - E[Y(0)] \qquad \text{(estimand)}$$

from association:

$$E[Y|X = 1] - E[Y|X = 0] \qquad \text{(estimator)}$$

Exchangeability

- How to estimate $\mathrm{E}[Y(1)]$ and $\mathrm{E}[Y(0)]$ given the "fundamental problem of causal inference?"
- Have to use observable quantities from the data.
- When is $\mathrm{E}[Y|X=1]$ a suitable substitute of $\mathrm{E}[Y(1)]$?
- When is $\mathrm{E}[Y|X=0]$ a suitable substitute of $\mathrm{E}[Y(0)]$?

Exchangeability

- When selecting (or receiving) treatment is random, individuals are exchangeable.
- I.e., they are "comparable" in terms at baseline or pre-treatment characteristics.
- Treatment selection is unconfounded or strongly ignorable when:

$$\Pr(X = 1 | Y(1), Y(0)) = \Pr(X = 1).$$

- This is why randomized experiments are the gold standard for inferring causal effects.

Exchangeability

- Under exchangeability or no unmeasured confounding, association = causation. Then:

$$E[Y(1)] - E[Y(0)] = E[Y|X = 1] - E[Y|X = 0].$$

Conditional exchangeability and no unmeasured confounding

Association $\neq$ Causation for non-randomized treatments

- In observational studies, treatment is unlikely to be randomized.
- There are likely pre-treatment (or baseline) covariates affecting treatment selection and the outcome simultaneously.
- E.g., those who did not sleep well the night before are *more likely to have a coffee* and to be *more tired*.
- Such confounding or non-random treatment selection distorts comparisons of outcomes across different treatment groups.

Conditional exchangeability

- To draw causal conclusions in observational studies, we would ideally compare "like with like."
- Suppose each participant reported before deciding to drink or not drink coffee how well they slept (denoted by L).
- How would an ideal analysis look like?

Conditional exchangeability

- We might posit that after taking sleep quality into account, drinking or not drinking a coffee was essentially random.
- Under this assumption, individuals with the same exact value of the baseline characteristics are exchangeable.
- E.g., if L was binary (good sleep vs. poor sleep), then those with poor sleep had the same probability to drink a coffee.

Conditional exchangeability

- Treatment selection is strongly ignorable conditional on baseline covariates L when:

$$\Pr(X = 1 | Y(1), Y(0), L) = \Pr(X = 1 | L).$$

- This is conditional exchangeability given baseline covariates.
- When this assumption holds, there is *no confounding after adjusting for L*, or that *L is sufficient to adjust for all confounding*.

Conditional exchangeability

- This assumption cannot be empirically validated using observed data; we need to justify it using prior knowledge [Steiner et al., 2010, VanderWeele, 2019].
- It is a causal concept that does not apply when simply assessing associations [Hernán and Robins, 2020].

Conditional exchangeability

- Under conditional exchangeability, among individuals with the *same unique value(s) of* L, association = causation:

$$\mathrm{E}[Y(1) - Y(0)|L] = \mathrm{E}[Y|X = 1, L] - \mathrm{E}[Y|X = 0, L].$$

Conditional exchangeability

- In practice, L is often continuous or multivariate.
- E.g., there may be multiple measures of sleep or rest, and they may be continuous variables.
- It is often unfeasible in practice to stratify the sample based on unique value(s) of L due to the curse of dimensionality.
- How to parsimoniously condition or statistically control for L?

Propensity score

Propensity score

- We have to account for the non-random treatment selection.
- Probability of selecting treatment given baseline covariates is termed the propensity score [Rosenbaum and Rubin, 1983].

$$e(L) = p(L) = \Pr(X = 1|L).$$

Propensity score

- Suppose that for all unique values of L:

$$0 < \Pr(X = 1|L) < 1. \tag{1}$$

- Under no unmeasured confounding and (1):

$$\Pr(X = 1|Y(1), Y(0), L) = \Pr(X = 1|e(L)).$$

- Adjusting for $e(L)$ – in place of L – suffices for conditional exchangeability.
- Aside: (1) is termed the positivity assumption; concerns covariate overlap (Part 3).

Propensity score methods

- Propensity scores are typically unknown in practice.
- They have to be estimated using e.g., logistic regression.
- The estimated propensity scores are then used to construct effect estimators. (Part 3)
- How to choose L to defensibly justify conditional exchangeability? (Part 2)

Other applications building on the potential outcomes framework

Other applications using potential outcomes

Principal strata

- Each individual has a pair of potential outcomes $\{Y(1), Y(0)\}$.
- For binary Y, four possible values for the pair:
  - $\{0, 0\}$ or $\{1, 1\}$: No treatment effect
  - $\{1, 0\}$: Protected or is helped by treatment
  - $\{0, 1\}$: Harmed or is hurt by treatment
- These are termed principal strata [Frangakis and Rubin, 2002].
- Fundamental problem of causal inference rules out knowing which stratum an individual belongs to.
- Permits testing no individual treatment effects vs. no average treatment effect [Loh et al., 2017, Rigdon et al., 2017].

Other applications using potential outcomes

Causal inference in the presence of interference

- SUTVA assumes no treatment interference; may be unfeasible under certain settings.
- E.g., social influence: my tiredness may depend not only on whether I drink coffee but whether you drink coffee.
- Need to expand notation and inferential methods to reflect the dependence of the outcomes on different possible treatments of others [Loh and Ren, 2022a, Loh et al., 2020]

Other applications using potential outcomes

Mediation analysis

- Investigate causal mechanisms via intermediate variables
- Nested potential outcomes that depend on treatment and mediator (possibly under a different treatment level); e.g.,
  - $M(1)$: counterfactual mediator under treatment;
  - $Y(0, M(1))$: potential outcome under control but mediator set to its counterfactual level under treatment.
- Single mediator [Nguyen et al., 2021, Tingley et al., 2014].
- Multiple mediators using interventional effects [Loh and Ren, 2022b, Loh et al., 2021].
- Confounding adjustment in mediation analysis [Loh and Ren, 2023].

References

C. E. Frangakis and D. B. Rubin. Principal stratification in causal inference. *Biometrics*, 58(1):21–29, 2002.

M. A. Hernán and J. M. Robins. *Causal Inference: What If*. Chapman & Hall/CRC, Boca Raton, FL, 2020.

P. W. Holland. Statistics and causal inference. *Journal of the American Statistical Association*, 81(396):945–960, 1986.

G. W. Imbens and D. B. Rubin. *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press, 2015.

W. W. Loh and D. Ren. Estimating social influence in a social network using potential outcomes. *Psychological Methods*, 27(5):841–855, 2022a. doi: 10.1037/met0000356.

W. W. Loh and D. Ren. Improving causal inference of mediation analysis with multiple mediators using interventional indirect effects. *Social and Personality Psychology Compass*, e12708, 2022b. doi: 10.1111/spc3.12708.

W. W. Loh and D. Ren. Adjusting for baseline measurements of the mediators and outcome as a first step toward eliminating confounding biases in mediation analysis. *Perspectives on Psychological Science*, Ahead of print, 2023. doi: 10.31234/osf.io/s387m. URL psyarxiv.com/s387m.

W. W. Loh, T. S. Richardson, and J. M. Robins. An apparent paradox explained. *Statistical Science*, 32(3): 356–361, 2017. doi: 10.1214/17-STS610. URL https://doi.org/10.1214/17-STS610.

W. W. Loh, M. G. Hudgens, J. D. Clemens, M. Ali, and M. E. Emch. Randomization inference with general interference and censoring. *Biometrics*, 76(1):235–245, 2020. doi: 10.1111/biom.13125.

W. W. Loh, B. Moerkerke, T. Loeys, and S. Vansteelandt. Disentangling indirect effects through multiple mediators without assuming any causal structure among the mediators. *Psychological Methods*, 2021. doi: 10.1037/met0000314. URL psyarxiv.com/3q4hg.

T. Q. Nguyen, I. Schmid, and E. A. Stuart. Clarifying causal mediation analysis for the applied researcher: Defining effects based on what we want to learn. *Psychological Methods*, 26(2):255–271, 2021. doi: 10.1037/met0000299.

J. Rigdon, W. W. Loh, and M. G. Hudgens. Response to comment on 'randomization inference for treatment effects on a binary outcome'. *Statistics in Medicine*, 36(5):876–880, 2017. doi: https://doi.org/10.1002/sim.7192. URL https://doi.org/10.1002/sim.7192.

P. R. Rosenbaum and D. B. Rubin. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55, 1983.

D. B. Rubin. [On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9.] Comment: Neyman (1923) and Causal Inference in Experiments and Observational Studies. *Statistical Science*, 5(4):472 – 480, 1990. doi: 10.1214/ss/1177012032. URL https://doi.org/10.1214/ss/1177012032.

J. Splawa-Neyman, D. M. Dabrowska, and T. P. Speed. On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9. *Statistical Science*, 5(4):465 – 472, 1990. doi: 10.1214/ss/1177012031. URL https://doi.org/10.1214/ss/1177012031.

P. M. Steiner, T. D. Cook, W. R. Shadish, and M. H. Clark. The importance of covariate selection in controlling for selection bias in observational studies. *Psychological Methods*, 15(3):250–267, 2010. doi: 10.1037/a0018719.

D. Tingley, T. Yamamoto, K. Hirose, L. Keele, and K. Imai. mediation: R package for causal mediation analysis. *Journal of Statistical Software*, 59(5):1–38, 2014. doi: 10.18637/jss.v059.i05. URL http://www.jstatsoft.org/v59/i05/.

T. J. VanderWeele. Brief report: Concerning the consistency assumption in causal inference. *Epidemiology*, 20(6): 880–883, 2009. URL http://www.jstor.org/stable/25662778.

T. J. VanderWeele. Principles of confounder selection. *European journal of epidemiology*, 34(3):211–219, 2019. doi: 10.1007/s10654-019-00494-6.