Part #1: Thorny assumptions, tricky estimands! Causal inference is hard, especially with mediators and post-randomization exposures

session: Trade-offs & tensions in mediation analyses of experimental data

Trang Quynh Nguyen
tnguye28@jhu.edu | trang-q-nguyen.weebly.com

Johns Hopkins Bloomberg School of Public Health

SREE, March 9, 2019

Litwok talk. Non-experimental vs. experimental

- Context: HPOG intervention
- Target estimand: effect of using an enhanced program feature on those who use the feature
- Experimental study: randomize feature offering, estimate average effect of offering on the compliers (up-takers)

$$\begin{tikzcd} C \arrow[rr, bend left] \arrow[rrr, bend left] & \text{offer} \arrow[r] & \text{use} \arrow[r] & Y \end{tikzcd}$$

- Non-experimental study: use data from the offering arm, estimate average effect of use on the users (up-takers)

$$\begin{tikzcd} C \arrow[rr, bend left] \arrow[rrr, bend left] & & \text{use} \arrow[r] & Y \end{tikzcd}$$

- Agree on one feature (peer support) but not the other two (emergency assistance, non-cash incentives)

Litwok talk. Non-experimental vs. experimental

Reminds of the impacts of ASSUMPTIONS!

- Different assumptions give different results
  - one (non-experimental) vs. two types (experimental) of people
  - if exclusion restriction doesn't hold, IV result is biased
  - if strong ignorability doesn't hold, ATT is biased
- If assumptions untestable, need to be very careful
  - consider its plausibility
  - consider the plausibility of its violation
  - conduct sensitivity analyses
- Also hidden/implicit assumptions

Litwok talk. Non-experimental vs. experimental

Ignorability likely does not hold – yet we invoke it a lot!

- Are some exposures more prone to ignorability violation?
  - i.e., despite a rich set of covariates
  - *non-ignorable* exposure or exposure *not at random*?
- Often the decision on which variables to include as covariates in a propensity score analysis or regression analysis is ad hoc
  - perhaps we would benefit from some shared lists of usual confounders (or at least causes) of types of exposures
  - could use as an ideal list or a starting point

Litwok talk. Non-experimental vs. experimental

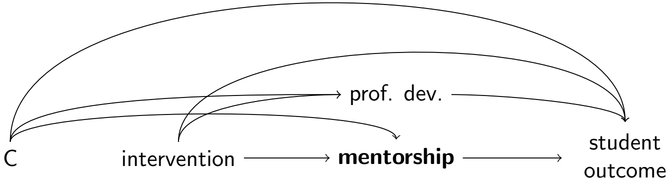
Another angle: different ESTIMANDS

What is the counterfactual?

- The experimental study
  - contrasts those that use the emergency assistance when offered to those with same need but not given the opportunity
  - effect of such assistance on people who need it and would use it
- The non-experimental study – if exclusion restriction holds
  - contrasts those who use the assistance to those similar to them in observed characteristics who do not use the assistance (and may not need it)
  - effect of having the need (or circumstances that give rise to a need) for the assistance – in the context where such need is met
- More transparent about what we estimate – the *effective estimand*

Harvill talk. A combination of questions

- Context: Comprehensive Teacher Induction



- Broad question: how does mentorship (in the context of the intervention) influence student outcome?
- Specific questions/estimands of interest
  - effect of mentorship on student outcome
  - intervention effect mediated and not mediated by mentorship
  - intervention effect modification by potential/expected mentorship-under-intervention (a baseline variable)

Harvill talk. A combination of questions

Issue raised: assumptions required by common methods often don't hold

- Effect of mentorship on student outcome
  - IV method requires exclusion restriction (violated)
  - methods that adjust for confounders require no unobserved mentorship-outcome confounding (likely violated)
- Intervention effect mediated by mentorship
  - also requires no unobserved mentorship-outcome confounding
- Intervention effect modification by potential/expected mentorship-under-intervention (ASPES)
  - fit outcome model incl. interaction of intervention with predicted mentorship-under-intervention $\hat{M}(1)$ (based on observed baseline X) – or generally, any meaningful function $g(X)$
  - interpretation as effect modification by $M(1)$ (or $\mathrm{E}[M(1)|C]$) requires exclusion restriction (w.r.t. effect modification):
    - X only modifies intervention effect through $M(1)$ (may be violated given multiple mediators)
    - and perhaps either X captures all effect modification or other effect modification is separate from effect modification by X (hard to judge)
  - not sure easier than the unobserved confounding assumption, since need good predictors of $M(1)$

Harvill talk. A combination of questions

What can be done about these violated assumptions?

- Unobserved confounding
  - there are sensitivity analyses for both non-mediation and mediation settings
  - need tutorials and tools to make these easier to understand and use
- ASPES's exclusion restriction w.r.t. effect modification
  - assuming a good set of X
  - in this multiple mediator case, can this be tackled more directly?
  - using info from all $\{\hat{M}(1), \hat{M}(0), \hat{N}(1), \hat{N}(0)\}$

Unlu talk. Treatment as high dose received

- Context: Leadership Training Program

$$\begin{tikzcd}
C \arrow[rr, bend left] \arrow[rrr, bend left] & \text{intervention} \arrow[r] \arrow[rr, bend left] & \text{dose} \arrow[r] & \text{student outcome}
\end{tikzcd}$$

- Target estimand: effect of HIGH dose on students who receive it
  - relevant generally – dichotomizing in defining exposure

	complier	non-complier
intervention	HIGH dose	some/no
control	no/little	no/some

- Not conventional IV, as exclusion restriction is violated

Unlu talk. Treatment as high dose received

Consider alternative strategies

- Subgroup regression analyses: combine compliers (non-compliers) with all controls, adjust for observed baseline X
- Principal score weighting: weight controls by their principal score (probability of being in principal stratum) given baseline X
- Multi-site instrumental variables

I'll comment on the first two

- commonality: adjust for X
- difference analogous to difference b/w regression adjustment and propensity score weighting adjustment

Unlu talk. Treatment as high dose received

ASSUMPTIONS – subgroup regression and principal score weighting

- Weak principal ignorability

  $$\mathrm{E}[Y(0) \mid \text{complier}, X] = \mathrm{E}[Y(0) \mid \text{noncomplier}, X] = \mathrm{E}[Y(0) \mid X]$$

  - allows all controls, given X values, to serve as controls for both compliers (estimating CACE) and noncompliers (estimating NACE)
  - violated because dose in control condition depends on principal stratum – same X different doses for compliers and noncompliers

- I propose extension: Weak PI for *outcome function of dose*

  $$\mathrm{E}[Y(0,d)|\text{complier}, X] = \mathrm{E}[Y(0,d)|\text{noncomplier}, X] = \mathrm{E}[Y(0,d)|X]$$

  where d (indexing dose) is in the support of dose under control given X

  - within X values, compliers and noncompliers in the control condition are exchangeable in the sense that given the same dose they share the same expected outcome
  - licenses all persons ($\neq$ outcomes) in control condition to serve as controls for both compliers and noncompliers

Unlu talk. Treatment as high dose received

But what to do with dose variation? ... Let's zoom in to CACE...

- This is an ESTIMAND question: what is the counterfactual? effect of high dose compared to what?
- If want the natural complier dose variation under control condition
  - problem: dose may depend on principal stratum conditional on X, so the dose distribution is generally not identified
  - identification requires predictors of dose that render stratum independent of control dose
- If want the zero-dose counterfactual
  - need to zero out the control dose
  - under *weak principal ignorability for outcome function of dose*
    - weighting estimation: discard controls with positive doses and weight those with zero dose up to each X stratum
    - regression estimation: adjust additionally for dose in controls, or discard controls with positive doses
  - discarding not desirable if lose a lot of controls

Part #2: Mediator of post-randomization stochastic exposure with exclusion restriction (Yang talk)

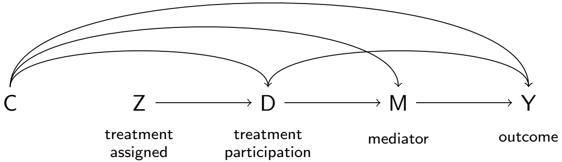
session: Trade-offs & tensions in mediation analyses of experimental data

Trang Quynh Nguyen
tnguye28@jhu.edu | trang-q-nguyen.weebly.com

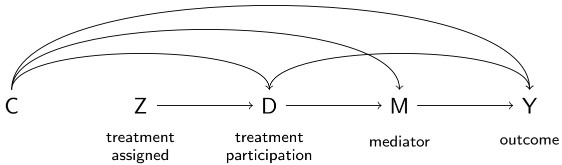
Johns Hopkins Bloomberg School of Public Health

SREE, March 9, 2019

Stochastic framework

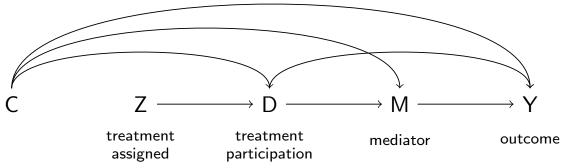


Stochastic framework



- Stochastic potential responses $D_i(z), M_i(d), Y_i(d, m)$ are nice!
  - lets reality be random
- Stochastic $D_i(z)$ seems to be more fundamental reality, and principal strata a coarsened version of it
  - a perception of reality based on what we observe that serves as a nice tool for describing reality as we perceive it
  - it's easier to process discrete categories and yes/no states than probabilities and uncertainties

Stochastic framework



- Stochastic $M_i(d)$
  - is helpful for conceptualizing effects at the individual level: $Y_i(D_i = 1, M_i = M_i(0))$ hard to conceive of if $M_i(1)$ does not have a chance to take on the value that $M_i(0)$ manifests
  - also helps disentangle two pairs of concepts:
    - deterministic vs. stochastic assignment
    - natural vs. interventional effects

    disentangled, they can be crossed:
    - deterministic natural effects and stochastic natural effects (both descriptive) are defined differently but are equal (in expectation)
    - stochastic natural and stochastic interventional effects (one descriptive, one prescriptive) are generally not equal

Assumptions

D1 no interference

D3 exclusion restriction

D5 non-zero average effect of Z on D

S2 stochastic monotonic effect of Z on D within levels of C

S1,3,4 sequential ignorability, of $D(z)$ and $M(d)$

Effect definition strategy and its scope

- WATE generalizes CACE, but no parallel generalization of NACE
  - because WATE is effect of D, not of Z
    - in the deterministic framework, CACE and NACE refer to effects of Z on compliers and noncompliers
    - with exclusion restriction, monotonicity, and binary treatment dose,
      - NACE $= 0$
      - CACE $=$ average effect of D (treatment participation) on compliers
      - average effect of treatment participation on noncompliers undefined
    - WATE generalizes the average effect of D on compliers, hence no NACE counterpart

- Can this effect definition strategy be stretched to cover situations with fewer assumptions?
  - if remove exclusion restriction, the answer seems to be no
  - if monotonicity, so compliance score may be negative
    - dichotomize at zero (partitioning covariate space) & have 2 WATEs?
    - perhaps more principled to let effect vary with compliance score?
  - how about if treatment dose is not truly binary, as in Unlu's case?

Unobserved U

- It is helpful to separate the different types of (unobserved) confounders: U_{DY}, U_{DM}, U_{MY}
  - if believe IV assumptions, can test for the existence of U_{DY} and U_{DM} (Litwok's case)
  - if U_{DY} separate from U_{DM}, U_{MY}, confounding of direct and indirect effects are separate

M-Y unobserved confounding and post-treatment confounding

- Bounds for NIE_Z and NDE_Z that decompose the ITT
  - relevant if interest is in effect of policy
  - a solution for the problem of intermediate confounding for this case
- An assumption-lean sensitivity analysis for unobserved pre-treatment M-Y confounders

Thank you to the speakers for your illuminating papers

Thank you the organizer for the opportunity

Thank you all for bearing with me :-)