Topics in Causal Inference

STAT41530

Jingshu Wang

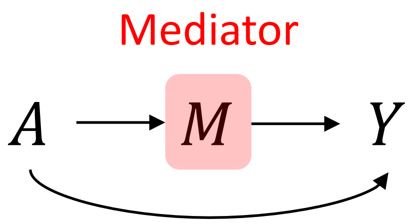
Lecture 8

Topic:
Causal mediation analysis

- Traditional regression approach and examples
- Direct and indirect effects
  - Causal definitions
  - Identifiability conditions
  - Connections with the regression approach

Some slides refer to Tyler's lecture https://www.youtube.com/watch?v=EI5y6pV87-Q

Mediation analysis



Causal effect of A on Y $\begin{cases} \text{Direct effect of } A \\ \text{Indirect effect of } A \text{ through } M \end{cases}$

Example:

A: a genetic variant 15q25.1

M: Smoker or not

Y: lung cancer

We see associations: $\mathrm{Cov}(A, M) \neq 0, \quad \mathrm{Cov}(A, Y) \neq 0$

Would that imply that M is a mediator for the causal effect of A on Y?

Motivations of mediation analysis

- Scientific understanding and explanation
  - E.g. Do genetic variants affect lung cancer through smoking or independently?
- Limiting the effects of exposure by intervening on a mediator
  - E.g. Can we eliminate the effect of antipsychotic medication on mortality by preventing the primary mechanism for mortality?
- Refinement of interventions
  - Improving components of an intervention to target mechanism
  - Eliminating costly ineffective components of an intervention
    - E.g. Does a CBT intervention improve depressive symptoms only through antidepressant use?

Traditional regression approaches

- Estimate direct effect of A

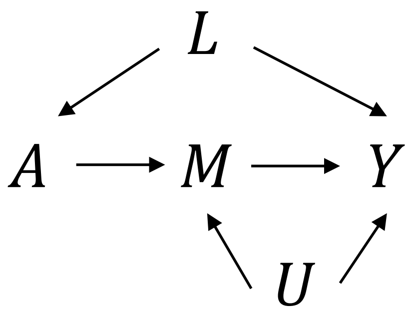
Direct effect

$$Y = \alpha_0 + \alpha_1 A + \alpha_2 M + \alpha_3 L + \varepsilon$$

- Two ways to estimate indirect effect of A
  - $Y = \beta_0 + \beta_1 A + \beta_2 L + \varepsilon$ Indirect effect: $\beta_1 - \alpha_1$
  - $M = \gamma_0 + \gamma_1 A + \gamma_2 L + \varepsilon$ Indirect effect: $\gamma_1 \alpha_2$

- When do these regressions estimate causal effect?
- Which way is correct for estimating the "indirect effect"?
- Which variables need to be included in L?
  - Mediator-outcome confounding: Even if all exposure-outcome confounders are considered, there can still be confounders of mediator-outcome relationship
  - Exposure-mediator interaction: The linear model assumes to interaction between the exposure and mediator on the outcome

Mediator-outcome confounding: example 1



A: Smoking mother

M: infant birthweight

Y: infant mortality

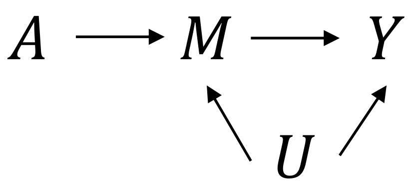
(Yerushalmy 1972, Wilcox 1993, Hernandez-Diaz 2006)

- Partition samples into birthweight stratas
- Within each strata, estimate the effect of mother smoking after controlling for confounders including mother's age, health conditions …
- Result: for the lowest birthweight strata, $\boldsymbol{A}$ **has a negative association with** $\boldsymbol{Y}$ even after controlling for other covariates
- Does that indicate that smoking mother has a beneficial direct effect for some infants?
- No!

  Unmeasured mediator-outcome confounder U: birth defect

  Conditioning on M causes a collider bias

Mediator-outcome confounding: example 2



SMaRT trial (Strong et. al. 2008)

A: Cognitive behavioral therapy intervention (completely randomized)

M: Use of antidepressant

Y: Depression symptoms after 3 months

- Total causal effect of A on Y: $\tau = E[Y(1)] - E[Y(0)]$

  result: $\tau < 0$

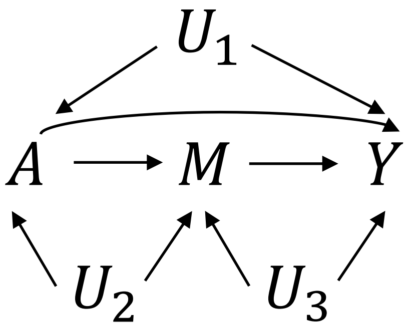
- Effect of A on M: $E[M(1)] - E[M(0)] > 0$
- Regress Y both on treat A and mediator M: $Y = \alpha_0 + \alpha_1 A + \alpha_2 M + \varepsilon$

  result: $\alpha_2 > 0$! !

  Looks like Antidepressant use increases depression

- Unmeasured mediator-outcome confounder U: patient experiencing a more difficult situation

Definitions of direct / indirect effects



$U = (U_1, U_2, U_3)$

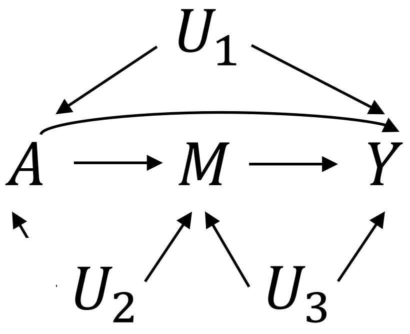
Structural equation

$Y = f(A, M, U, E_Y)$

- Potential outcomes:
  - $Y(a, m) = f(a, m, U, E_Y)$
  - $Y(0), Y(1)$: $Y(A = 0)$ and $Y(A = 1)$

    Total effect of A: $Y(1) - Y(0)$
  - $M(0), M(1)$: potential outcomes of M at different levels of A

Definitions of direct / indirect effects



- Direct effect of A
  - Controlled direct effect at m

$$\begin{aligned}\mathrm{CDE}(m) &\triangleq Y(1,m) - Y(0,m)\\ &= Y \mid do(A)=1, do(M)=m - Y \mid do(A)=0, do(M)=m\end{aligned}$$

Not $Y|do(A), M$!! (should not condition on a post-treatment variable)

  - Natural direct effect

$$\mathrm{NDE} = Y(1, M(0)) - Y(0, M(0))$$

$M(0)$ can be different on different individuals

- Indirect effect of A
  - Natural indirect effect

$$\mathrm{NIE} = Y(1, M(1)) - Y(1, M(0))$$

- Total effect decomposition: $Y(1) - Y(0) = Y(1, M(1)) - Y(0, M(0)) = NIE + NDE$

Example of the definition

	$M(0)$	$M(1)$	$Y(0,0)$	$Y(1,0)$	$Y(0,1)$	$Y(1,1)$	CDE $m = 0$	CDE $m = 1$	NDE	NIE	TE
1	0	1	0	1	0	1	1	1	1	0	1
2	1	1	0	1	0	0	1	0	0	0	0
3	0	1	0	0	0	1	0	1	0	1	1

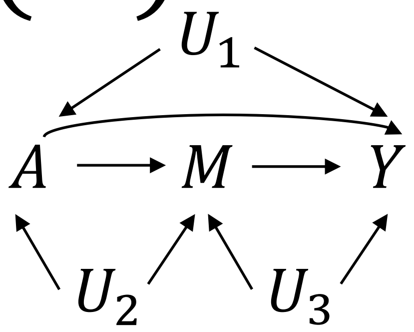
Average NDE: $\frac{1}{3}$

Average NIE: $\frac{1}{3}$

Average TE: $\frac{2}{3} = \frac{1}{3} + \frac{1}{3}$

Average CDE(m): $\frac{2}{3}$ for m = 0 or 1

Identification of average $\mathbf{CDE}(m)$



Need to identify $\mathrm{E}(Y(a,m))$

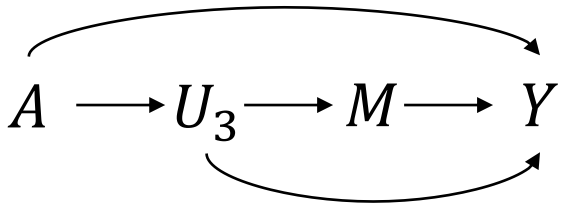
[or $P(Y|do(A,M))$]

- Need to include confounders L that block all backdoor paths between (A, M) and Y
- L does not need to block the backdoor paths between A and M

- Specifically, we need three assumptions for identifying $E(Y(a,m))$
  1. $Y(a,m) \perp A \mid L$
  2. $Y(a,m) \perp M \mid A, L$
  3. Positivity: $P(M = m, A = a \mid L) > 0$

$$\begin{aligned} \mathbb{E}\left[Y(a,m)\right] &= \mathbb{E}\left[\mathbb{E}\left[Y(a,m)|L\right]\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[Y(a,m)|L, A = a\right]\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[Y(a,m)|L, A = a, M = m\right]\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[Y|L, A = a, M = m\right]\right] \end{aligned}$$

- The first two assumptions exclude the following DAG

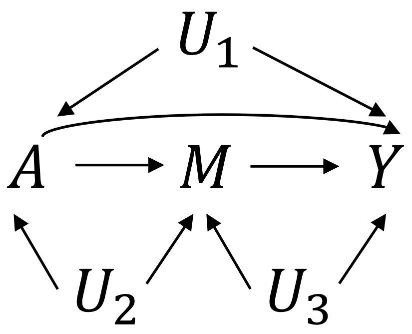


Identification of average NDE and IDE

$$\text{NDE} = Y(1, M(0)) - Y(0, M(0))$$

$$\text{NIE} = Y(1, M(1)) - Y(1, M(0))$$

Need to identify $E(Y(a, M(a')))$



Need three assumptions to identify $E(Y(a, M(a')))$ (Imai et. al. 2010, Stat. Sci.)

1. $(Y(a,m), M(a')) \perp A \mid L$ ⟵ Much stronger than $Y(a,m) \perp A \mid L$
2. $Y(a,m) \perp M \mid A, L$
3. Positivity: $P(M = m, A = a \mid L) > 0$

Lemma: $(X, Y) \perp Z \implies X \perp Z | Y$

Proof:

$$P(X, Z|Y) = \frac{P(X,Y,Z)}{P(Y)}$$

$$= \frac{P(X,Y)P(Z)}{P(Y)} = P(X|Y)P(Z)$$

$$\mathbb{E}\left[Y(a, M(a'))\right] = \mathbb{E}\left[\mathbb{E}\left[Y(a, M(a')) \mid L, M(a')\right]\right]$$

$$\begin{aligned}
&\mathbb{E}\left[Y(a, M(a')) \mid L, M(a') = m\right] \\
=&\mathbb{E}\left[Y(a, m) \mid L, M(a') = m\right] = \mathbb{E}\left[Y(a, m) \mid L, M(a') = m, A = a'\right] \\
=&\mathbb{E}\left[Y(a, m) \mid L, M = m, A = a'\right] = \mathbb{E}\left[Y(a, m) \mid L, A = a'\right] \\
=&\mathbb{E}\left[Y(a, m) \mid L, A = a'\right] = \mathbb{E}\left[Y(a, m) \mid L\right] \\
=&\mathbb{E}\left[Y \mid L, A = a, M = m\right]
\end{aligned}$$

Formulas for identifying direct and indirect effects

$$\mathbb{E}\left[Y(a, M(a')) \mid L\right] = \sum_m \mathbb{E}\left[Y(a, m) \mid L, M(a') = m\right] \mathbb{P}\left[M(a') = m \mid L\right]$$
$$= \sum_m \mathbb{E}\left[Y \mid L, M = m, A = a\right] \mathbb{P}\left[M = m \mid L, A = a'\right]$$

- $E[\mathrm{CDE}(m)|L] = E[Y|A = 1, M = m, L] - E[Y| A = 0, M = m, L]$
- $E[\mathrm{NDE}|L] = \sum_m P[M = m|L, A = 0]\,\{E[Y|A = 1, M = m, L] - E[Y| A = 0, M = m, L]\}$
- $E[\mathrm{NIE}|L] = \sum_m P[Y|A = 1, M = m, L]\,\{P[M = m|A = 1, L] - E[M = m| A = 0, L]\}$

Connection with linear regression models

- If $E[Y|A, M, L] = \alpha_0 + \alpha_1 A + \alpha_2 M + \alpha_3 L$
- Then $\alpha_1 = E[\text{CDE}(m)|L] = E[\text{NDE}|L]$ direct effect

- What about indirect effect $E[\text{NIE}|L]$?
  - If we assume $E[Y|A, L] = \beta_0 + \beta_1 A + \beta_2 L$

    Then $E[Y(1) - Y(0)|L] = E[Y|A = 1, L] - E[Y|A = 0, L] = \beta_1$

    $E[\text{NIE}|L] = \beta_1 - \alpha_1$ (total effect decomposition)

  - If we assume $E[M|A, L] = \gamma_0 + \gamma_1 A + \gamma_2 L$

$$\begin{aligned}
\mathbb{E}[\text{NIE} \mid L] &= \sum_m \mathbb{E}[Y \mid A = 1, M = m, L] \left(\mathbb{P}[M = m \mid A = 1, L] - \mathbb{P}[M = m \mid A = 0, L]\right) \\
&= \sum_m (\alpha_0 + \alpha_1 + \alpha_2 m + \alpha_3 L) \left(\mathbb{P}[M = m \mid A = 1, L] - \mathbb{P}[M = m \mid A = 0, L]\right) \\
&= \alpha_2 \sum_m m \left(\mathbb{P}[M = m \mid A = 1, L] - \mathbb{P}[M = m \mid A = 0, L]\right) \\
&= \alpha_2 \left(\mathbb{E}[M \mid A = 1, L] - \mathbb{E}[M \mid A = 0, L]\right) = \alpha_2 \gamma_1
\end{aligned}$$