Quantitative Research Methods IV - 17.806

Recitation, Week 12.
Topic: Causal Mediation.

Benjamín Muñoz

May 5, 2023

MIT

Table of contents

1/ Mediation

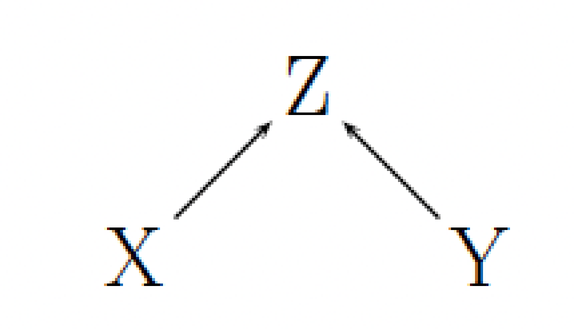
Directed Acyclic Graphs (DAGs)

- Explicit representation of causal system. Describe all causal relationships relevant (key assumptions: absence of edges and arrows).
- Prohibition of cycles (simultaneous/reverse causation and feedback loops).

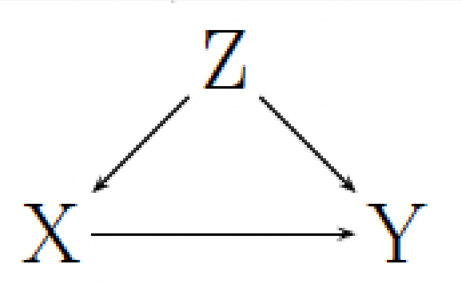
  1. **Nodes** (letter): random variables. Solid circles for observed RV and hollow circles for unobserved/unmeasured RV.
  2. **Directed Edges** ($\longrightarrow$) causal effect from on node (cause) to another (effect).

- Draw DAGs: use R packages $\rightsquigarrow$ ggdag and `dagitty`.

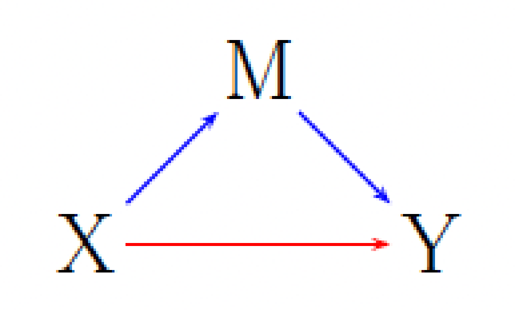
Directed Acyclic Graphs (DAGs)



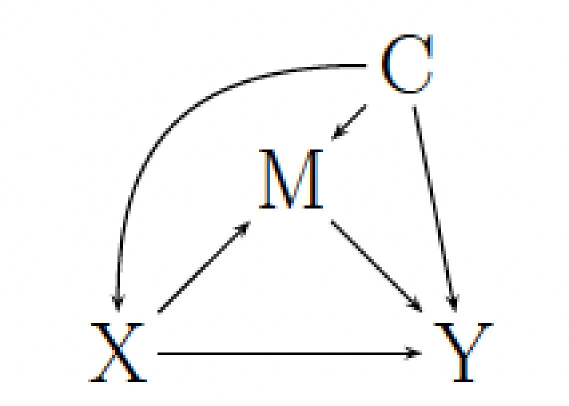
Directed Acyclic Graphs (DAGs)



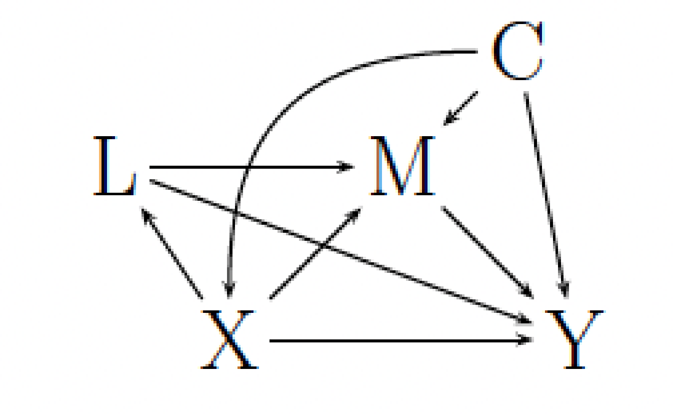
Directed Acyclic Graphs (DAGs)



Directed Acyclic Graphs (DAGs)



Directed Acyclic Graphs (DAGs)



2/ Theory

Causal Mediation

- **Potential mediators:** $M_i(t)$, where $M_i = M_i(T_i)$ observed.
- **Potential outcomes:** $Y_i(t, m)$, where $Y_i = Y_i(T_i, M_i(T_i))$ observed.
- **Causal Quantities:**
  1. Total Causal Effect: $\tau_i \equiv Y_i(1, M_i(1)) - Y_i(0, M_i(0))$
  2. Causal Mediation Effects (CME): $\delta_i(t) \equiv Y_i(t, M_i(1)) - Y_i(t, M_i(0))$
  3. (Causal) Direct Effects (NDE): $\zeta_i(t) \equiv Y_i(1, M_i(t)) - Y_i(0, M_i(t))$
- CME are also known as **Natural Indirect Effects**. Causal effect of the change in M_i on Y_i that would be induced by T_i.
- The NDE is the causal effect of T_i on Y_i , holding the mediator constant at its potential value that would realize when $T_i = t$.

Causal Mediation

Total Effect = Natural Direct Effect + Natural Indirect Effect

$$\tau_i \;=\; \delta_i(t) + \zeta_i(1-t) \;=\; \frac{1}{2}\{\delta_i(0) + \delta_i(1) + \zeta_i(0) + \zeta_i(1)\}$$

- Under sequential ignorability, ACME and ADE are nonparametrically identified.

Sequential Ignorability

$$\begin{aligned} \{Y_i(t', m), M_i(t)\} \;&\perp\!\!\!\perp\; T_i \mid X_i \\ Y_i(t', m) \;&\perp\!\!\!\perp\; M_i(t) \mid T_i = t, X_i \end{aligned}$$

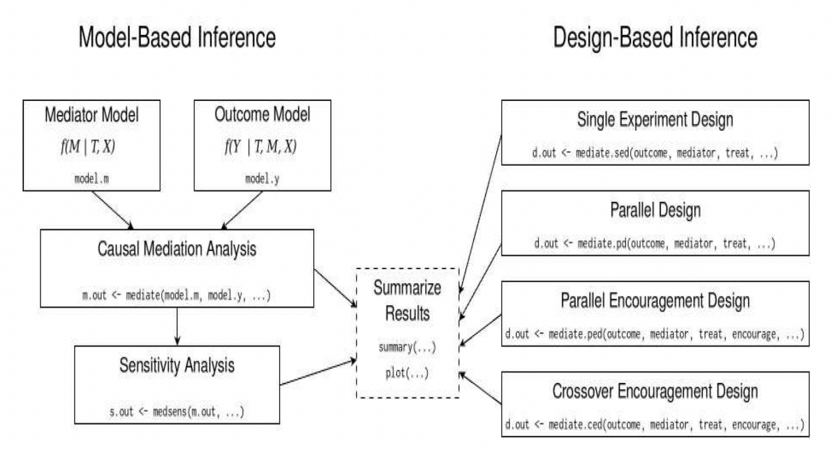
where $0 < P(T_i = t|X_i = x)$ and $0 < P(M_i = m|T_i = t, X_i = x)$ for $t = 0, 1$, and all x and m in the support of X_i and M_i, respectively.

Causal Mediation

- **Strong Assumption:** SI excludes the existence of (measured or unmeasured) post-treatment confounders as well as that of unmeasured pre-treatment confounders.
- Therefore, rules out the possibility of multiple mediators that are causally related to each other (specific techniques for this case).
- Randomization of Treatment is not enough.

3/ Implementation

mediation Package Workflow



Model-Based Approach

1. Estimate models for the outcome $\mathbb{E}(Y_i \mid T_i, M_i, X_i)$ and mediator $p(M_i \mid T_i, X_i)$.
2. Predict mediator for both treatment values $(M_i(1), M_i(0))$.
3. Predict outcome by first setting $T_i = 1$ and $M_i = M_i(0)$, and then $T_i = 1$ and $M_i = M_i(1)$.
4. Compute the average difference between two outcomes to obtain a consistent estimate of ACME

R Code

```
### Load packages
library(mediation))          # Causal Mediation

### Run the Models
med <- glm(mediator ~ 1 + treatment + covariates, family = binomial(link = "logit"), data = df)
out  <- glm(outcome ~ 1 + treatment + covariates + mediator, family = gaussian, data = df)

### Implement Model-Based Mediation Analysis
set.seed(17806)
mediation1 <- mediate(model.m = med, model.y = out, sims = 1000, boot = FALSE,
treat = "treatment", mediator = "mediator", robustSE = FALSE, conf.level = 0.95)
```

Model-Based Approach

	Outcome model types						
Mediator model types	Linear	GLM	Ordered	Censored	Quantile	GAM	Survival
Linear (`lm`/`lmer`)	✓	✓	✓*	✓	✓	✓*	✓
GLM (`glm`/`bayesglm`/`glmer`)	✓	✓	✓*	✓	✓	✓*	✓
Ordered (`polr`/`bayespolr`)	✓	✓	✓*	✓	✓	✓*	✓
Censored (tobit via `vglm`)	–	–	–	–	–	–	–
Quantile (`rq`)	✓*	✓*	✓*	✓*	✓*	✓*	✓
GAM (`gam`)	✓*	✓*	✓*	✓*	✓*	✓*	✓*
Survival (`survreg`)	✓	✓	✓*	✓	✓	✓*	✓

Model-Based Approach

```
R> summary(med.out)

Causal Mediation Analysis

Nonparametric Bootstrap Confidence Intervals with the Percentile Method

                         Estimate 95% CI Lower 95% CI Upper p-value
ACME (control)             0.0848       0.0392         0.13  <2e-16 ***
ACME (treated)             0.0858       0.0405         0.13  <2e-16 ***
ADE (control)              0.0117      -0.0727         0.13    0.58
ADE (treated)              0.0127      -0.0785         0.14    0.58
Total Effect               0.0975      -0.0091         0.23    0.06 .
Prop. Mediated (control)   0.8698      -0.7110         3.65    0.06 .
Prop. Mediated (treated)   0.8804      -0.5530         3.48    0.06 .
ACME (average)             0.0853       0.0398         0.13  <2e-16 ***
ADE (average)              0.0122      -0.0756         0.13    0.58
Prop. Mediated (average)   0.8751      -0.6320         3.57    0.06 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Sample Size Used: 265
```

Design-Based Approach

- The function depends on the type of design and the assumptions involved (interactions, etc.).

	Mediator Manipulation	
Type	Direct	Indirect
Parallel	Parallel Design	Parallel Encouragement Design
Crossover	Crossover Design	Crossover Encouragement Design

- For the PSet use `multimed`.