# Linear mediation and moderation

**Session 12**

MATH 80667A: Experimental Design and Statistical Methods
for Quantitative Research in Management
HEC Montréal

# Outline

## Linear mediation model

--

## Interactions and moderation

# Linear mediation

# Three types of associations

## Confounding

Common cause

Causal forks $X \leftarrow Z \rightarrow Y$

## Causation

Mediation

Causal chain $X \rightarrow Z \rightarrow Y$

## Collision

Selection / endogeneity

inverted fork $X \rightarrow Z \leftarrow Y$

# Key references

- Imai, Keele and Tingley (2010), A General Approach to Causal Mediation Analysis, *Psychological Methods*.
- Pearl (2014), Interpretation and Identification of Causal Mediation, *Psychological Methods*.
- Baron and Kenny (1986), The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations, *Journal of Personality and Social Psychology*

Limitations:

- Bullock, Green, and Ha (2010), Yes, but what's the mechanism? (don't expect an easy answer)
- Uri Simonsohn (2022) Mediation Analysis is Counterintuitively Invalid

# Sequential ignorability assumption

Define

- treatment of individual $i$ as $X_i$,
- potential mediation given treatment $x$ as $M_i(x)$ and
- potential outcome for treatment $x$ and mediator $m$ as $Y_i(x, m)$.

Given pre-treatment covariates $W$, potential outcomes for mediation and treatment are conditionally independent of treatment assignment.

$$Y_i(x', m), M_i(x) \perp\!\!\!\perp X_i \mid W_i = w$$

Given pre-treatment covariates and observed treatment, potential outcomes are independent of mediation.

# Total effect

**Total effect**: overall impact of $X$ (both through $M$ and directly)

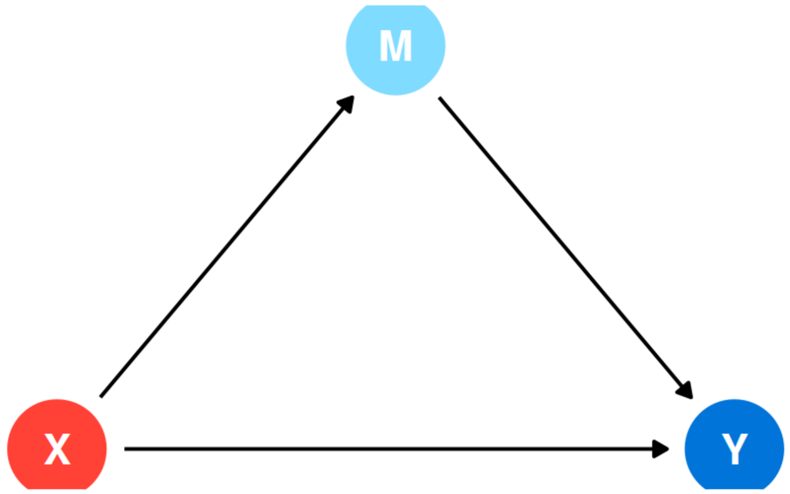
$$\text{TE}(x, x^*) = \textsf{E}[Y \mid \mathrm{do}(X = x)] - \textsf{E}[Y \mid \mathrm{do}(X = x^*)]$$

This can be generalized for continuous $X$ to any pair of values $(x_1, x_2)$.

**X → M → Y**

**plus**

**X → Y**



# Average controlled direct effect

$$
\begin{aligned}
\mathsf{CDE}(m, x, x^*) &= \mathsf{E}[Y \mid \mathrm{do}(X = x, m = m)] - \mathsf{E}[Y \mid \mathrm{do}(X = x^*, m = m)] \\
&= \mathsf{E}\{Y(x, m) - Y(x^*, m)\}
\end{aligned}
$$

Expected population change in response when the experimental factor changes from $x$ to $x^*$ and the mediator is set to a fixed value $m$.

# Direct and indirect effects

**Natural direct effect:** $\text{NDE}(x, x^*) = \text{E}[Y\{x, M(x^*)\} - Y\{x^*, M(x^*)\}]$

- expected change in $Y$ under treatment $x$ if $M$ is set to whatever value it would take under control $x^*$

**Natural indirect effect:** $\text{NIE}(x, x^*) = \text{E}[Y\{x^*, M(x)\} - Y\{x^*, M(x^*)\}]$

- expected change in $Y$ if we set $X$ to its control value and change the mediator value which it would attain under $x$

Counterfactual conditioning reflects a physical intervention, not mere (probabilistic) conditioning.

Total effect is $\text{TE}(x, x^*) = \text{NDE}(x, x^*) - \text{NIE}(x^*, x)$

# Linear structural equation modelling and mediation

# The Baron–Kenny model

Given **uncorrelated** unobserved noise variables $U_M$ and $U_Y$, consider linear regression models

$$
\begin{aligned}
M &= c_M + \alpha x + U_M \\
Y &= c_Y + \beta x + \gamma m + U_Y
\end{aligned}
$$

Plugging the first equation in the second, we get the marginal model for $Y$ given treatment $X$,

$$
\begin{aligned}
\mathsf{E}_{U_M}(Y \mid x) &= \underbrace{(c_Y + \gamma c_M)}_{\text{intercept}} + \underbrace{(\beta + \alpha\gamma)}_{\text{total effect}} \cdot x + \underbrace{(\gamma U_M + U_Y)}_{\text{error}} \\
&= c_Y' + \tau X + U_Y'
\end{aligned}
$$

# The old method

Baron and Kenny recommended running regressions and estimating the three models with

1. whether $\mathscr{H}_0 : \alpha = 0$
2. whether $\mathscr{H}_0 : \tau = 0$ (total effect)
3. whether $\mathscr{H}_0 : \gamma = 0$

The conditional direct effect $\alpha\gamma$ and we can check whether it's zero using Sobel's test statistic.

Problems?

# Sobel's test

Based on estimators $\widehat{\alpha}$ and $\widehat{\gamma}$, construct a Wald-test

$$S = \frac{\widehat{\alpha}\widehat{\gamma} - 0}{\sqrt{\widehat{\gamma}^2 \mathsf{Va}(\widehat{\alpha}) + \widehat{\alpha}^2 \mathsf{Va}(\widehat{\gamma}) + \mathsf{Va}(\widehat{\gamma})\mathsf{Va}(\widehat{\alpha})}} \stackrel{\cdot}{\sim} \mathsf{No}(0,1)$$

where the point estimate $\widehat{\alpha}$ and its variance $\mathsf{Va}(\widehat{\alpha})$ can be estimated via SEM, or more typically linear regression (ordinary least squares).

# Null distribution for the test

The large-sample normal approximation is poor in small samples.

The popular way to estimate the *p*-value and the confidence interval is through the nonparametric **bootstrap** with the percentile method.

Repeat $B$ times, say $B = 10\,000$

1. sample **with replacement** $n$ observations from the database
   - tuples $(Y_i, X_i, M_i)$
2. recalculate estimates $\widehat{\alpha}^{(b)}\widehat{\gamma}^{(b)}$

# Boostrap *p*-values and confidence intervals

## Confidence interval

Percentile-based method: for a equi-tailed $1-\alpha$ interval and the collection

$$\{\widehat{\alpha}^{(b)}\widehat{\gamma}^{(b)}\}_{b=1}^{B},$$

compute the $\alpha/2$ and $1-\alpha/2$ empirical quantiles.

## Two-sided *p*-value

Compute the sample proportion of bootstrap statistics $S^{(1)}, \ldots, S^{(B)}$ that are larger/smaller than zero.

If $S^{(M)} < 0 \leq S^{(M+1)}$ for $1 \leq M \leq B$.

$$p = 2\min\{M/B, 1-M/B\}$$

and zero otherwise

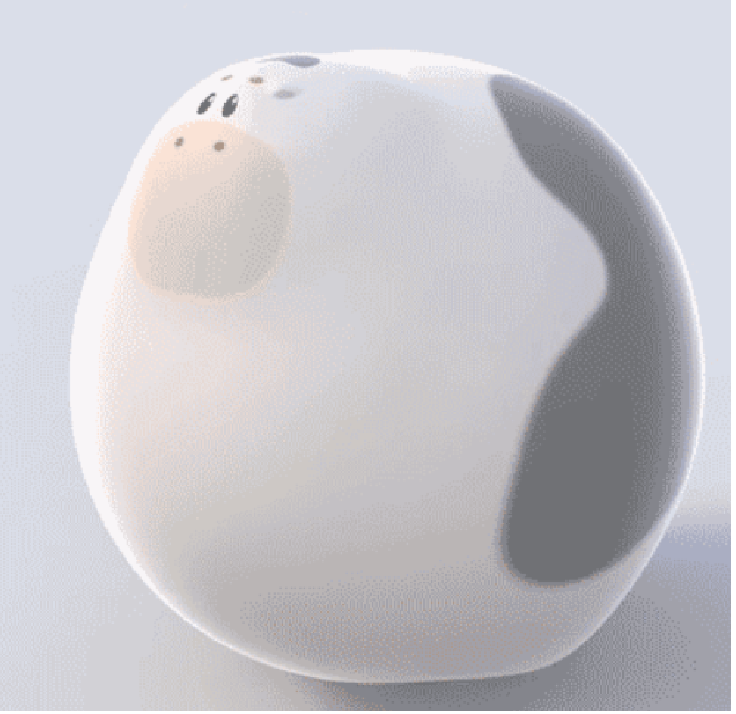
# Example from Preacher and Hayes (2004)

> Suppose an investigator is interested in the effects of a new cognitive therapy on life satisfaction after retirement.
>
> Residents of a retirement home diagnosed as clinically depressed are randomly assigned to receive 10 sessions of a new cognitive therapy $(X = 1)$ or 10 sessions of an alternative (standard) therapeutic method $(X = 0)$.
>
> After Session 8, the positivity of the attributions the residents make for a recent failure experience is assessed $(M)$.
>
> Finally, at the end of Session 10, the residents are given a measure of life satisfaction $(Y)$. The question is whether the cognitive therapy's effect on life satisfaction is mediated by the positivity of their causal attributions of negative experiences. "

# Defaults of linear SEM

- Definitions contingent on model
  - (causal quantities have a meaning regardless of estimation method)
- Linearity assumption not generalizable.
  - effect constant over individuals/levels

Additional untestable assumption of uncorrelated disturbances (no unmeasured confounders).



Keenan Crane

# Assumptions of causal mediation

Need assumptions to hold (and correct model!) to derive causal statements

- Potential confounding can be accounted for with explanatories.
- Careful with what is included (colliders)!
  - *as-if* randomization assumption
- Generalizations to interactions, multiple mediators, etc. should require careful acknowledgement of confounding.