Bayesian Methods for Inference on the Causal Effects of Mediation

Chanmin Kim

Department of Statistics
Univsesity of Florida

Final Defense
June 26 2013

Outline

What is the Causal Effect of Mediation?

- Investigators interest in Causal Mechanism as well as causal effects.
- Randomized experiments often only determine the effect of the treatment on the outcome.
- Causal mediation analysis:

$$
\begin{array}{ccccc}
 & & \textit{Mediator}\,(M) & & \\
 & \nearrow & & \searrow & \\
\textit{Treatment}\,(Z) & & \longrightarrow & & \textit{Outcome}\,(Y)
\end{array}
$$

- Question : How can we learn about the causal mechanism from observational or experimental studies?

Examples

- Family and social support mediating the effect of care managed-based interventions on depression.
- The effect of vaccines can be from stimulation of subject's immune system or reducing risk of infection (herd effect) among population.
- Motivation and self-efficacy mediating the effect of perceived person-job fit on job performance.
- Compliance level (measured by self-monitoring records) mediating the effect of face-to-face counseling treatment on weight loss.

Potential Outcomes Framework

Also called 'Rubin Causal Model' (RCM) (Holland, 1986)

- Binary treatment : Z (Control=0 vs. Intervention=1).
- Mediator : M
- Outcome : Y
- Baseline covariates : X
- Potential mediators : M_z where $M_{obs} = ZM_1 + (1-Z)M_0$ observed
- Potential outcomes : $Y_{z,m}$ where $Y_{obs} = ZY_{1,M_1} + (1-Z)Y_{0,M_0}$ observed

Only one of $\{Y_{1,M_1}, Y_{0,M_0}\}$ can be observed for each i. And $\{Y_{1,M_0}, Y_{0,M_1}\}$ are never observed!

Causal Mediation Effects

Then,

- Total Effect (TE) :

$$Y_{1,M_1} - Y_{0,M_0}$$

- Indirect Effect (IE) (Mediation Effect) :

$$Y_{1,M_1} - Y_{1,M_0}$$

- Causal effect of the change in M on Y while holding the treatment constant.
- Represents the mechanism through M.
- Contains unobservable Y_{1,M_0}!

- Direct Effect (DE) :

$$
\begin{aligned}
& TE - IE \\
&= Y_{1,M_1} - Y_{0,M_0} - \{Y_{1,M_1} - Y_{1,M_0}\} \\
&= Y_{1,M_0} - Y_{0,M_0}
\end{aligned}
$$

- Causal effect of treatment z on outcome Y while holding mediator constant at its natural value under control.
- Represents causal mechanism around M.
- Also, contains unobservable Y_{1,M_0}!

Controlled Effect vs Natural Effect (Pearl, 2001)

Controlled Direct Effect

- $Y_{1,m} - Y_{0,m}$ for a certain level $m \in M$.
- Causal effect if one can directly manipulate the mediator.

Natural Direct Effect

- $Y_{1,M_0} - Y_{0,M_0}$.
- Mediator is set to have its natural value.

Average Causal Effect of Intervention

- Since we observe either Y_{1,M_1} and Y_{0,M_0} for an individual, compute $E\{Y_{1,M_1} - Y_{0,M_0}\}$.

Assumptions for Identification

Several combinations of assumptions are typically required for inference on causal effects.

- Sequential ignorability

$$\{Y_{z',m}, M_z\} \perp Z \mid X$$

$$Y_{z',m} \perp M_z \mid Z, X$$

- No-interaction assumption (about controlled direct effect)

$$Y_{1,m} - Y_{0,m} = B$$

- Exclusion restriction

$$Y_{z,m} = Y_{z',m} \quad \text{for all } z, z' \text{ and for all } m$$

states treatments affect potential outcomes only through mediators. So, no D.E.

Literature

1. Structural Equation Models (SEM) (Baron & Kenny, 1986; MacKinnon et al, 2002; Sobel, 1982; Preacher & Hayes, 2008) : highly parametric. assume sequential igonorability and no interaction.
2. Principal Stratification (PS) (Frangakis & Rubin, 2002, Gallop et al, 2009)
   : assume SUTVA. problems with drawing the population effects. (VanderWeele, 2011)
3. Instrumental Variables (IV) (Angrist et al., 1996; Albert, 2008; Sobel, 2008)
   : assume exclusion restriction
4. Structural Mean Models (SMM) (Ten Have et al., 2004)
   : assume no-interaction.
5. Marginal Structural Models (MSM) (Robins et al., 2000)
   : assume sequential ignorability.

Goals of the Dissertation

Present a Bayesian framework for inference on causal effects of mediation

1. Use different and intuitive assumptions with sensitivity parameters.
2. Examine sensitivity to the assumptions.
3. Develop for several common scenarios (e.g.,single mediator, longitudinal mediation, multiple mediators).

Single Mediator Case

assume only one mediator on the causal pathway

Review of Notation

From the potential outcome framework, define

- Binary Treatment : Z (Control=0 vs. Intervention=1).
- Potential Mediator : M_z under treatment z.
- Observed Mediator : $M_{obs} = ZM_1 + (1 - Z)M_0$
- Potential Outcome : $Y_{z,M_{z'}}$ for $z, z' \in \{0, 1\}$
- Observed Outcome : $Y_{obs} = ZY_{1,M_1} + (1 - Z)Y_{0,M_0}$

Note that Y_{1,M_0} and Y_{0,M_1} are not observable.

Assumptions

Assumption 1 (Randomization)

$$\{Y_{z',m}, M_z\} \perp Z.$$

Assumption 2a

Let $D = (M_1 - M_0)$. For a fixed ϵ,

$$P(Y_{1,M_0} = 1 | M_0 = m, |D| < \epsilon) = P(Y_{1,M_1} = 1 | M_1 = m, |D| < \epsilon).$$

Assumption 2b

For a fixed ϵ, and χ, we assume

$$P(Y_{1,M_0} = 1 | M_0 = m, |D| \geq \epsilon) = \\ \chi^{\mathrm{sgn}(d)} P(Y_{1,M_1} = 1 | M_1 = m, |D| \geq \epsilon).$$

Assumption 3

$$f_{M_{z'}}(m_{z'} | m_z, Y_{z,M_z}) = f_{M_{z'}}(m_{z'} | m_z)$$

Assumption 4 (Joint Distribution of Mediators)

This follows a Gaussian copula model (Nelsen, 1999),

$$F_{M_0,M_1}(m_0, m_1) = \Phi_2\left[\Phi_1^{-1}\{F_{M_0}(m_0)\}, \Phi_1^{-1}\{F_{M_1}(m_1)\}\right]$$

where Φ_1 is the univariate standard normal CDF and Φ_2 is the bivariate normal CDF with mean $(0,0)^T$, variance $(1,1)^T$ and correlation $\rho \in (-1, 1)$.

Assumption 5 (Conditional Independence)

$$f_{(1,M_1),(1,M_0),(0,M_0)}(y_{11}, y_{10}, y_{00} | m_0, m_1) =$$
$$f_{1,M_1}(y_{11} | m_0, m_1) f_{1,M_0}(y_{10} | m_0, m_1) f_{0,M_0}(y_{00} | m_0, m_1).$$

Identification

Theorem

The joint posterior distribution of NIE and NDE is identified under Assumptions 1-5.

TOURS : weight management trial (Perri et al., 2008)

- Randomized trial to compare the effectiveness of extended care programs of weight management.
- After completing a standard six month lifestyle modification program, participants were randomly assigned to telephone counseling, face-to-face counseling or an education control group.
- Treatment : Face to face (FTF) vs Education control (EC).
- Mediator : The number of days with self-monitoring records for food intake (0-350) during 6 to 18 months.
- Binary outcome : Among those that lost $\geq 5\%$ of weight by 6 months, indicator of whether they maintained the $\geq 5\%$ weight loss from 6-18 months.

Table : Posterior means and credible intervals : ρ=0.7.

ϵ	χ	NDE	NIE	TE
50	1	0.077	0.007	0.085
		(-0.073,0.25)	(-0.092,0.13)	(-0.070,0.25)
50	1.3	0.088	-0.003	0.085
		(-0.085,0.26)	(-0.10,0.099)	(-0.070,0.25)
75	1	0.077	0.007	0.085
		(-0.066,0.25)	(-0.088,0.12)	(-0.070,0.25)
75	1.3	0.086	-0.001	0.085
		(-0.087,0.26)	(-0.097,0.10)	(-0.070,0.25)
100	1	0.078	0.007	0.085
		(-0.069,0.25)	(-0.091,0.12)	(-0.070,0.25)
100	1.3	0.084	0.0006	0.085
		(-0.084,0.26)	(-0.10,0.10)	(-0.070,0.25)

Summary

- Based on analysis, the number of self-monitoring food records completed was not a mediator.

Longitudinal Mediator

assume single treatment and longitudinal outcomes and mediators

Approaches

In the time-varying treatment setting

- Regression based approach: Maxwell et al. (2011). Structural models between Z_t, M_t and Y_t at each time and over all time periods.
- Marginal structural models: van der Laan and Petersen (2004). Estimating the marginal mean of potential outcome $E(Y^{\bar{z}, \bar{M}^{\bar{z}'}})$ where $\bar{z}$ denotes treatment histories.

Others

- Parallel-process model: Cheong et al. (2003). Mediators and outcomes are modeled as different parallel processes with latent factors. Specify structural equation models of latent factors.
- Principal strata approach: Lin et al. (2008). Use a mediator to form principal strata and draw principal strata direct and indirect effects.

Basic Framework

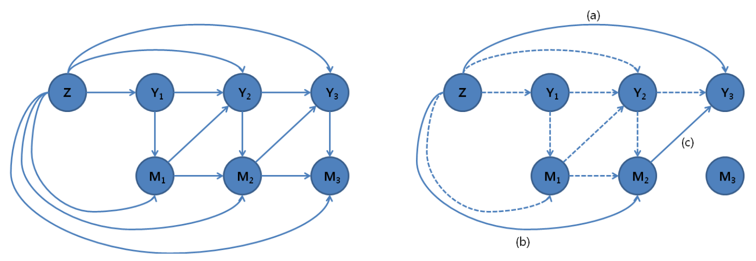

Figure : Mediation model with time $t = 1, 2, 3$.

- Outcome Y_t is affected by the preceding mediator M_{t-1} and both have their own causal pathways.
- Mediator M_t is affected by the outcome Y_t.
- Outcomes and mediators have direct paths from the treatment Z.

Longitudinal Mediation Analysis

- Potential mediator : M_t^z as the value of the mediator at time $t = 1, \cdots, T$ under treatment $Z = z$.
- Full histories of mediators : $M^1 = (M_1^1, \cdots, M_T^1)$ and $M^0 = (M_1^0, \cdots, M_T^0)$.
- Potential outcome : Y_t^{z, M_{t-1}^z} denotes the outcome that would be observed at time t if $Z = z$.
- Then, causal effects at time t,

$$
\begin{aligned}
NDE_t &= E(Y_t^{1,M_{t-1}^0} - Y_t^{0,M_{t-1}^0}) \\
NIE_t &= E(Y_t^{1,M_{t-1}^1} - Y_t^{1,M_{t-1}^0}) \\
TE_t &= NDE_t + NIE_t = E(Y_t^{1,M_{t-1}^1} - Y_t^{0,M_{t-1}^0}).
\end{aligned}
$$

- We update these by a Bayesian dynamic model and the observed data.

Bayesian Dynamic Models (West & Harrison, 1989)

Define $V_t = (Y_t, M_{t-1})$ to be the response at time t and mediator at time $t-1$ assume a model $p_t(V_t|\boldsymbol{\theta}_t)$ parameterized by $\boldsymbol{\theta}_t$.

Modeling Assumption

Conditional on a vector of state parameters, $\boldsymbol{\theta}_t$, $V_t = (Y_t, M_{t-1})$ is independent of V_{s-1} and $\boldsymbol{\theta}_s$ for all values of $s < t$.

- Observation Model : $V_t \sim p_t(V_t|\boldsymbol{\theta}_t, V_{t-1})$.
- Evolution Model : $(\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-\mathbf{1}}) \sim p_e(\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-\mathbf{1}})$ for $t = 1, \cdots, T$.

Assumptions

Let $D_{t-1} = M_{t-1}^1 - M_{t-1}^0$. And let $y_t^{zz'}$ denote $y_t^{z,M_{t-1}^{z'}}$ for notational simplicity.

Assumption 1a

For a fixed ϵ at each time t,

$$f_{1,M_{t-1}^0}(y_t^{10} | M_{t-1}^0 = m, |D_{t-1}| \leq \epsilon, \boldsymbol{\theta}_t) = f_{1,M_{t-1}^1}(y_t^{11} | M_{t-1}^1 = m, |D_{t-1}| \leq \epsilon, \boldsymbol{\theta}_t).$$

- Among the subject for whom the treatment would have minimal impact on the mediator (quantified by ϵ), the distributions of the outcomes are the same whether the mediator value induced by $z = 1$ or $z = 0$ conditional on $\boldsymbol{\theta}_t$.

Assumption 1b

For a fixed ϵ and χ,

$$f_{1,M_{t-1}^{0}}(y_t^{10}|M_{t-1}^{0}=m,|D_{t-1}|\geq\epsilon,\boldsymbol{\theta}_t)\propto \exp(\log(\chi^{sgn(D_{t-1})})y_t^{11})f_{1,M_{t-1}^{1}}(y_t^{11}|M_{t-1}^{1}=m,|D_{t-1}|\geq\epsilon,\boldsymbol{\theta}_t).$$

- The second assumption is for the subgroup of subjects for whom the intervention has a greater than ϵ effect on M_{t-1}.

Thus, Assumption 1a and b stratify the population into those where the intervention has a large versus small effect on the mediator.

Assumption 2

For a fixed z at each time t,

$$f_{z,M_{t-1}^{z}}(y_t^{zz} | m_{t-1}^{z}, m_{t-1}^{(1-z)}, \boldsymbol{\theta}_t) = f_{z,M_{t-1}^{z}}(y_t^{zz} | m_{t-1}^{z}, \boldsymbol{\theta}_t).$$

- The potential outcomes $Y_t^{z,M_{t-1}^{z}}$ are independent of the mediator under the other treatment $M_{t-1}^{(1-z)}$ conditional on the mediator associated with the potential outcomes M_{t-1}^{z} and the vector of state parameters $\boldsymbol{\theta}_t$.

Assumption 3 (Conditional Independence)

$$
\begin{aligned}
&f_{(1,M^1_{t-1}),(1,M^0_{t-1}),(0,M^0_{t-1})}(y^{11}_t, y^{10}_t, y^{00}_t | m^0_{t-1}, m^1_{t-1}, \boldsymbol{\theta}_t) = \\
&\quad f_{1,M^1_{t-1}}(y^{11}_t | m^0_{t-1}, m^1_{t-1}, \boldsymbol{\theta}_t) \times f_{1,M^0_{t-1}}(y^{10}_t | m^0_{t-1}, m^1_{t-1}, \boldsymbol{\theta}_t) \\
&\quad \times f_{0,M^0_{t-1}}(y^{00}_t | m^0_{t-1}, m^1_{t-1}, \boldsymbol{\theta}_t).
\end{aligned}
$$

- Not necessary to estimate $E[NIE_t|data]$ and $E[NDE_t|data]$.
- Necessary to estimate other features of the posterior distribution of NIE_t and NDE_t such as an upper bound on the variance.

Assumption 4

For the joint distribution of mediators, we assume a Gaussian copula model

$$F_{M_t^0, M_t^1}(m_t^0, m_t^1 | \boldsymbol{\theta}_{t+1}) = \Phi_2[\Phi_1^{-1}\{F_{M_t^0}(m_t^0|\boldsymbol{\theta}_{t+1})\}, \Phi_1^{-1}\{F_{M_t^1}(m_t^1|\boldsymbol{\theta}_{t+1})\}]$$

- ϕ_1 is the univariate standard normal CDF and ϕ_2 is the bivariate normal CDF with mean $(0,0)^T$, variance $(1,1)^T$ and correlation ρ.
- ρ is a sensitivity parameter since the two mediators are not observed at the same time.

Identification

Theorem

The joint posterior distribution of NIE_t and NDE_t for each $t = 1, \cdots, T$ is identified under Assumptions 1-4 and the Bayesian dynamic model.

Specification of Models

Observation model

- $(Y_t^{zz}, Y_{t-1}^{zz}) \sim \text{Mult}(\pi_{1t}^z, \pi_{2t}^z, \pi_{3t}^z)$ for $t = 2, \cdots, T$.
- For mediators, we need two conditional distributions,

$$f_{M_{t-2}^z}(m_{t-2}^z | y_t^{zz}, y_{t-1}^{zz}) \text{ and } f_{M_{t-1}^z}(m_{t-1}^z | m_{t-2}^z, y_t^{zz}, y_{t-1}^{zz}).$$

- We specify conjugate normal-normal MDP models (MacEachern and Müller, 1998) for these.

Evolution model

- $\boldsymbol{\pi_t}^z | \boldsymbol{\pi_{t-1}}^z \sim \text{Dir}(\delta_t^z \boldsymbol{\pi_{t-1}}^z)$, where the hyperparameter δ_t is given a uniform shrinkage prior (Daniels, 1999).
- For the mediator, we update parameters of the base measures of the MDP models (Observation models).

Application: CTQ II (Marcus et al., 2005)

- A randomized clinical trial to study the effect of moderate-intensity exercise on smoking cessation over 8 weeks.
- Moderate intensity exercise intervention ($n_1 = 109$) vs. equivalent staff contact time control ($n_0 = 108$).
- Potential mediator at week t is the difference between baseline weight and weight measured at week t.
- Binary outcome at week t is quit status (quit=1).
- Since the target quit week was week 3, we assess the effects from week 4 to week 8.
- Assume missingness is ignorable.

Sensitivity Parameters

- The difference in the average weight changes between two groups at week 8 was 0.7. with S.D. of 1.65.
- Consider differences more than 0.7 as the case that Assumption 1a does not hold.
- ϵ varies from 0.7 to 2.3 (average$+$ 1 S.D.).
- We assume the impact of treatment on the mediator being more than $1 \sim 2$ pounds could reflect a negative impact up to the odds ratio of about 0.5.
- χ varies from 0.5 to 1.
- We assume the positive correlation between mediators. $\rho \in [0, 1)$.

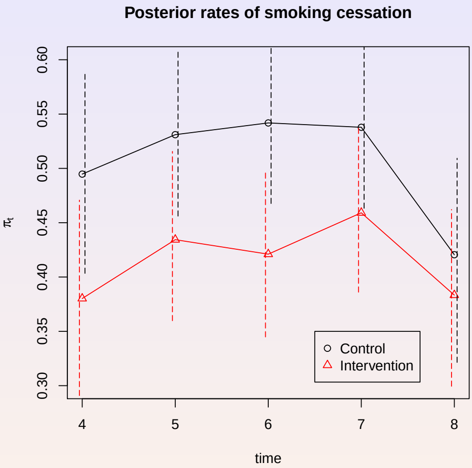

Figure : Cessation rates of the exercise intervention and the contact control under ignorability

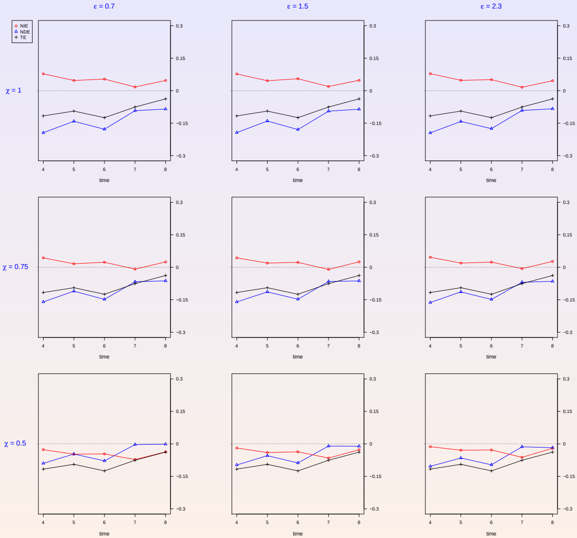

Figure : NIE, NDE and TE for week 4,5,6,7,8 when $\rho = 0.3$

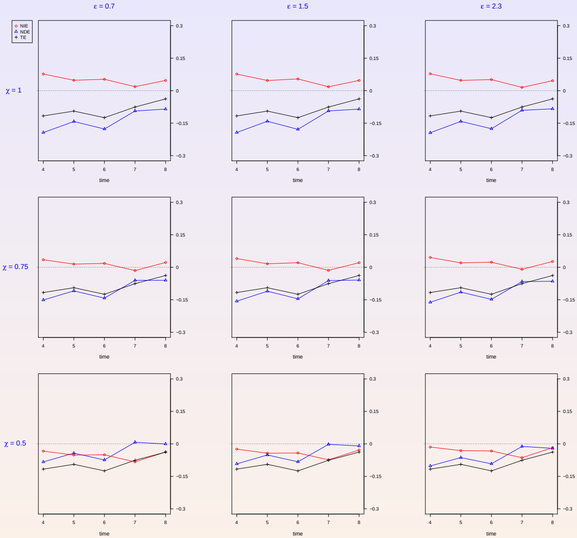

Figure : NIE, NDE and TE for week 4,5,6,7,8 when $\rho = 0.5$

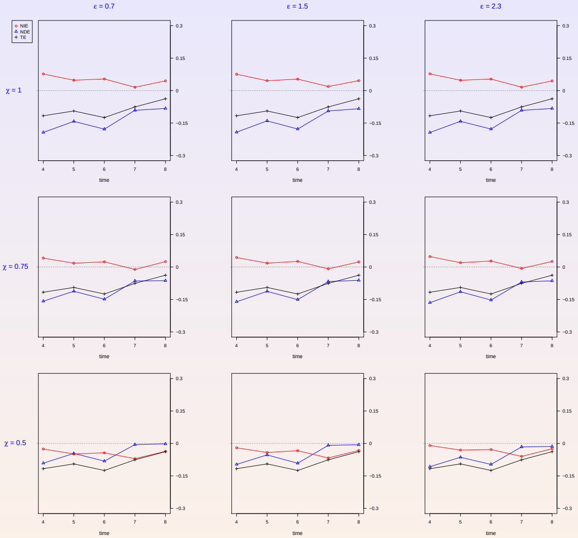

Figure : NIE, NDE and TE for week 4,5,6,7,8 when $\rho = 0.7$

Discussion

- Based on analysis, the effect of moderate exercise intervention vs. the staff contact control was marginally significant (in a negative way).
- The longitudinal mediator, weight change, had a significant positive impact
- For future work,
  - Explore time-varying treatments instead of a single time treatment.
  - Also, incorporate time-varying covariates to weaken some of assumptions.

Multiple Mediators

assume multiple mediators on the causal pathway

Setting of Multiple Mediators

- We can extend the previous method to the case of multiple mediators.

$$
\begin{array}{ccccc}
Z & & \longrightarrow & & Y \\
& \searrow & \vdots & \nearrow & \\
& & M^1, \cdots, M^n & &
\end{array}
$$

- We focus on the situation that mediators are measured at the same time (not sequentially).

Framework for Multiple Mediators

- M_z^k: k-th potential mediator under treatment $Z = z$.
- $Y_{z, M_{z_1}^1 M_{z_2}^2 \cdots, M_{z_K}^K}$: Potential outcome under randomization to intervention level z and realized mediator values $M_{z_1}^1 M_{z_2}^2 \cdots, M_{z_K}^K$ where $\{z, z_1, \cdots, z_K\} \in \{0,1\}^{\otimes(K+1)}$.
- Observed mediator: $M^k = ZM_1^k + (1-Z)M_0^k$.
- Observed outcome:
  $Y = ZY_{1, M_1^1 M_1^2 \cdots M_1^K} + (1-Z)Y_{0, M_0^1 M_0^2 \cdots M_0^K}$.
- Let $\mathbf{M}_z = \{M_z^1, M_z^2, \cdots, M_z^K\}$ for notational simplicity.
- The potential outcome $Y_{z, M_z^1 M_z^2 \cdots M_z^K}$ can be re-written as $Y_{z, \mathbf{M}_z}$.

Decomposition of the Total Effect

- Total Effect (TE) :

$$TE = E(Y_{1,\mathbf{M}_1} - Y_{0,\mathbf{M}_0}).$$

- Natural Direct Effect (NDE) :

$$NDE = E(Y_{1,\mathbf{M}_0} - Y_{0,\mathbf{M}_0}).$$

- Joint Natural Indirect Effect (JNIE) :

$$JNIE = TE - NDE = E(Y_{1,\mathbf{M}_1} - Y_{1,\mathbf{M}_0})$$

  which is the aggregate effect of all mediators.

Also, decompose $JNIE$ further to get mediator-specific indirect effects (e.g., $NIE_1, \cdots, NIE_K$).

Decomposition of JNIE

WLOG, assume 3 mediators on the causal pathway,

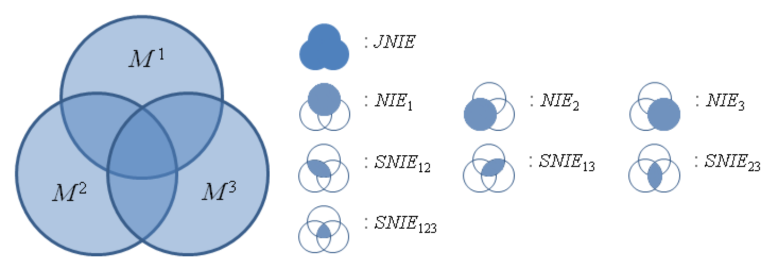

Figure : Partitioning of the *JNIE*

Mediator-specific indirect effect (NIE_k) is

- $NIE_1 = E(Y_{1,\mathbf{M}_1} - Y_{1,M_0^1 M_1^2 M_1^3})$.

Simultaneous indirect effect of two mediators (SE_{jk}) is

- $SNIE_{12} = E(Y_{1,\mathbf{M}_1} - Y_{1,M_0^1 M_1^2 M_1^3} - Y_{1,M_1^1 M_0^2 M_1^3} + Y_{1,M_0^1 M_0^2 M_1^3})$.

Simultaneous indirect effect of three mediators (SE_{123}) is

- $SNIE_{123} = E(Y_{1,\mathbf{M}_1} - Y_{1,M_0^1 M_1^2 M_1^3} - Y_{1,M_1^1 M_0^2 M_1^3} - Y_{1,M_1^1 M_1^2 M_0^3} + Y_{1,M_0^1 M_0^2 M_1^3} + Y_{1,M_0^1 M_1^2 M_0^3} + Y_{1,M_1^1 M_0^2 M_0^3} - Y_{1,\mathbf{M}_0})$.

Then,

$$\begin{aligned} JNIE &= \sum_k NIE_k - \sum_{j,k:j\neq k} SE_{jk} + SE_{123} = E(Y_{1,\mathbf{M}_1} - Y_{1,\mathbf{M}_0}) \\ &= TE - NDE. \end{aligned}$$

Assumptions

All assumptions are presented in the presence of $K = 3$ mediators.

Assumption1

For a given mediators under intervention $z = 1$, the conditional distributions of the outcome are the same whether those mediator values were induced by $z = 1$ or $z = 0$.

- $f_{1,M_0^1 M_0^2 M_0^3}(y_{1,\mathbf{M}_0} | \mathbf{M}_0 = \mathbf{m}_0, \mathbf{M}_1) = f_{1,M_1^1 M_1^2 M_1^3}(y_{1,\mathbf{M}_1} | \mathbf{M}_0, \mathbf{M}_1 = \mathbf{m}_0)$. Other cases are in page 49.

Assumption 2

$$f_{z,M_z^1 M_z^2 M_z^3}(y_{z,\mathbf{M}_z} | \mathbf{m}_z, \mathbf{m}_{1-z}) = f_{z,M_z^1 M_z^2 M_z^3}(y_{z,\mathbf{M}_z} | \mathbf{m}_z).$$

Assumption 3 (Joint Distribution of Mediators)

$$
\begin{aligned}
&F_{\mathbf{M}_0,\mathbf{M}_1}(\mathbf{m}_0,\mathbf{m}_0) = \\
&\qquad \Phi_6[\Phi_1^{-1}\{F_{M_0^1}(m_0^1)\}, \Phi_1^{-1}\{F_{M_0^2}(m_0^2)\}, \Phi_1^{-1}\{F_{M_0^3}(m_0^3)\}, \\
&\qquad\qquad \Phi_1^{-1}\{F_{M_1^1}(m_1^1)\}, \Phi_1^{-1}\{F_{M_1^2}(m_1^2)\}, \Phi_1^{-1}\{F_{M_1^3}(m_1^3)].
\end{aligned}
$$

Assumption 4 (Conditional Independence)

All potential outcomes that define JNIE, NIE_k, SNIE_{jk}, etc. are conditionally independent given all potential mediators under $z = 0, 1$.

Identification

Theorem

The joint posterior distribution of NDE, JNIE, NIE_k, SNIE_{jk} and SNIE_{123} is identified under Assumption 1-4.

Thank You.

Appendix

Assumption1

For a given mediators under intervention $z = 1$, the conditional distributions of the outcome are the same whether those mediator values were induced by $z = 1$ or $z = 0$.

- $f_{1,M_0^1,M_1^2,M_0^3}(y_{1,M_0^1,M_1^2,M_0^3} | M_0^1 = m^1, M_0^2, M_0^3, M_1^1, M_1^2 = m^2, M_1^3 = m^3)$
$= f_{1,M_1^1,M_1^2,M_1^3}(y_{1,\mathbf{M}_1} | \mathbf{M}_0, M_1^1 = m^1, M_1^2 = m^2, M_1^3 = m^3)$.
- $f_{1,M_1^1,M_1^2,M_0^3}(y_{1,M_1^1,M_1^2,M_0^3} | M_0^1, M_0^2, M_0^3 = m^3, M_1^1 = m^1, M_1^2 = m^2, M_1^3)$
$= f_{1,M_1^1,M_1^2,M_1^3}(y_{1,\mathbf{M}_1} | \mathbf{M}_0, M_1^1 = m^1, M_1^2 = m^2, M_1^3 = m^3)$.