Preview of New Features in Mplus Version 7.2 Including Mediation Modeling with Causally-Defined Effects

Bengt Muthén

Mplus
www.statmodel.com

bmuthen@statmodel.com

Presentation to PSMG November 5, 2013

1. Overview of New Features in Mplus Version 7.2

- New analysis features:
  1. Restructured routines for continuous-time survival analysis with latent variables
  2. Multiple-group CFA alignment with ML for binary outcomes and complex survey data
  3. Latent class and latent transition analysis with residual covariances
  4. Bootstrap SEs and CIs for ML with non-continuous outcomes leading to numerical integration
  5. Mediation analysis with effects based on potential outcomes (causal inference)

Overview of New Mplus Features Cont'd

- Convenience features:
  1. SEs for TECH4 and standardized coefficients with WLSMV and covariates
  2. Simplified language for factor analysis with nominal indicators
  3. New plots with estimated probabilities, Y-hat, and residuals
  4. More alignment output: invariance R-square, correlations between true and estimated factor scores, measurement invariance histogram plots
  5. Double DO loops for DEFINE, MODEL CONSTRAINT, MODEL TEST, and MODEL PRIORS
  6. Parameter names for parameter numbers listed as non-identified
  7. DEFINE for interaction terms using centering followed by multiplication

Plots can now be created in R using information from most of the Mplus PLOT command options. Mplus R functions read the Mplus GH5 file using the rhdf5 package from Bioconductor, thereby providing R with the necessary input data. See http://www.statmodel.com/mplus-R/

2. Latent Class and Latent Transition Analysis with Residual Covariances (Residual Associations) for Categorical Items

- Addition of a within-class two-way loglinear model, adding one association parameter per variable pair
  - Binary items: saturates the 2×2 table
  - Ordered polytomous items: Uniform association model (Goodman 1979)
- No need for numerical integration due to adding a factor behind the pair of items
- Association parameter can be equal of different across latent classes
- Covariates allowed, but not direct effects on items
- Asparouhov-Muthén web note

LCA Example: Deciding On The Number Of Classes For 17 Antisocial Behavior Items ($n = 7326$)

Five-Class Solution

The five-class solution is substantively meaningful:

Class 1	138.06985	0.01888	High Overall
Class 2	860.41897	0.11771	Property Offense
Class 3	1257.56652	0.17151	Drugs
Class 4	1909.32749	0.26219	Person Offense
Class 5	3160.61717	0.42971	Normative (Pot)

Six-Class Solution - adds a variation on Class 2 in the 5-class solution

Deciding On The Number Of Classes For 17 ASB Items

Number of classes	1	2	3	4	5	6
Loglikelihood	-48168.475	-42625.653	-41713.142	-41007.498	-40808.312	-40604.231
# par.	17	35	53	71	89	107
BIC	**96488**	**85563**	**83898**	**82647**	**82409**	**82161**

TECH10 bivariate tests in the 5-class run show need for adding residual covariances. Adding 4 residual covariances to the 5-class model:

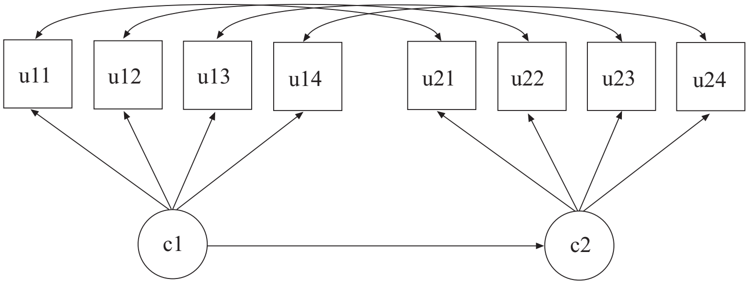
Loglikelihood = -40603, # parameters = 93, BIC = 82034

Mplus Version 7.2 LCA Input for 17 ASB items

```
VARIABLE:    NAMES = property fight shoplift lt50 gt50 force threat injure
             pot drug soldpot solddrug con auto bldg goods gambling
             dsm1-dsm22 sex black hisp single divorce dropout college
             onset f1 f2 f3 age94 cohort dep abuse;
             USEVARIABLES = property-gambling;
             CATEGORICAL = property-gambling;
             CLASSES = c(5);
ANALYSIS:    TYPE = MIXTURE;
             STARTS = 1200 300;
             PARAMETERIZATION = RESCOV;
MODEL:       %OVERALL%
             threat WITH injure;
             drug WITH soldpot;
             drug WITH solddrug;
             soldpot WITH solddrug;
OUTPUT:      TECH1 TECH8 TECH10;
```

Residual covariances can also be specified as class specific.

LTA with Correlated Residuals



Allowing across-time correlation for each item changes the estimated latent transition probabilities.

3. Mediation Analysis with Effects Based on Counterfactuals/Potential Outcomes (Causal Inference; Causally-Defined Effects)

Overview:

- Software
- The issues, intuitively
  - Continuous Y, continuous M with ”exposure-mediator interaction” influencing Y
  - Binary Y, Continuous M
- The causal effect definitions using the mediation formula
  - Specific case: Binary Y, continuous M
- Applications
  - Hopkins GBG data
  - MacKinnon smoking data
- Sensitivity analysis (M-Y confounding)

3.1 Causal Effects in Software

- Focus on:
  - Binary and count Y and M
  - Single Y, single M
  - Binary (treatment/control) X or continuous (exposure) X
  - Covariates

- Valeri-VanderWeele SAS/SPSS macros (Psych Methods, 2013)
- Tingley et al. R package mediation (forthcoming in JSS)
- Mplus
  - Muthén (2011). Applications of Causally Defined Direct and Indirect Effects in Mediation Analysis using SEM in Mplus (the paper, an appendix with formulas, and Mplus scripts are available at www.statmodel.com under Papers, Mediational Modeling.)
  - Mplus Version 7.2 simplifies the input for the single M, single Y case
  - Mplus is unique in allowing latent Y, M, and X (latent exposure), logit link without rare Y assumption, and nominal M or Y

Mplus Implementation of Causally-Defined Effects

- The effects can be estimated in Mplus using maximum-likelihood or Bayes
- ML:
  - Standard errors of the direct and indirect causal effects are obtained by the delta method using the Mplus MODEL CONSTRAINT command
  - Bootstrapped standard errors and confidence intervals are also available, taking into account possible non-normality of the effect distributions
- Bayes
  - Bayesian analysis is available in order to describe the possible non-normal posterior distributions
- Mplus Version 7.2 greatly simplifies how to get the causally-defined effects using MODEL INDIRECT (available for ML, including bootstrapping) instead of user-specified MODEL CONSTRAINT formulas

Mplus Version 7.2 MODEL INDIRECT Commands for Causal Effects

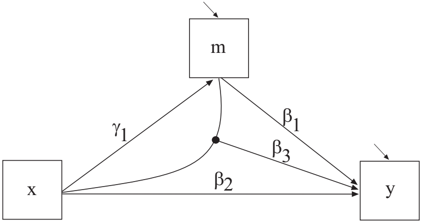
1. No moderation:
    - Y IND M X;
    - - all 3 can be latent
2. Moderation with X*M:
    - y MOD M XM X;
    - Y can be latent
3. Moderation with Z involving X and M:
    - Y MOD M Z(low, high, increment) MZ XZ X;
    - - only Y can be latent
4. Moderation with Z involving M and not X:
    - Y MOD M Z(low, high, increment) MZ X;
    - - X and Y can be latent
5. Moderation with Z involving X and not M:
    - Y MOD M Z(low, high, increment) XZ X;
    - - M and Y can be latent

For controlled direct effects an M value is placed in parenthesis: M(m).

3.2 The Issues, Intuitively

- Causally-defined effects based on counterfactuals and potential outcomes using expectations have been developed by Robins, Greenland, Pearl, VanderWeele, Vansteelandt, Imai etc
  - Total, direct, and indirect causal effects
- Different results than SEM with for instance "exposure-treatment interaction" (Y=X*M) or categorical DVs
- The effects are causal only under strong assumptions (if assumptions don't hold, are the causal methods better/useful anyway?)

Continuous Y and M with Exposure-Mediator Interaction



$$\textit{Total indirect effect}: \; TIE = \beta_1 \, \gamma_1 + \beta_3 \, \gamma_1. \quad (1)$$

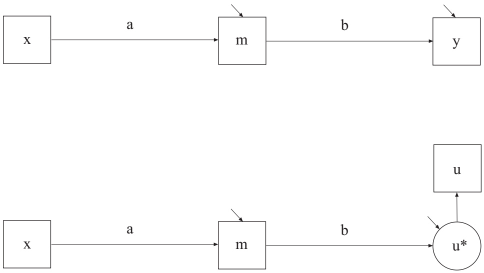
$$\textit{Pure indirect effect}: \; PIE = \beta_1 \, \gamma_1. \quad (2)$$

$$\textit{Direct effect}: \; DE = \beta_2 + \beta_3 \, \gamma_0. \quad (3)$$

$$\textit{Total direct effect}: \; TDE = \beta_2 + \beta_3 \, \gamma_0 + \beta_3 \, \gamma_1. \quad (4)$$

$$\textit{Total effect} = (1) + (3) = (2) + (4)$$

Continuous versus Binary Distal Outcome



Conventional versus Causal Mediation Effects with a Categorical Distal Outcome

With a categorical distal outcome, the conventional product formula for an indirect effect is only valid for an underlying continuous latent response variable behind the categorical observed outcome (2 linear regressions), not for the observed categorical outcome itself (linear plus non-linear regression).

Similarly, with a categorical mediator, conventional product formulas for indirect effects are only relevant/valid for a continuous latent response variable behind the mediator.

Comparing Causal and Naive Effects

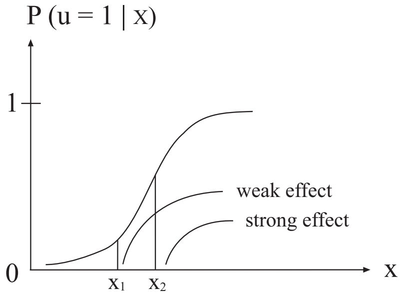
The difference between the causal effects and the effects obtained by what is called the naive approach has been studied in Imai et al. (2010a) and Pearl (2011c). Imai et al. (2010a, Appendix E, p. 23) conducted a Monte Carlo simulation study to show the biases, while Pearl (2011c) presented graphs showing the differences.

We will look at the differences in some examples.

Previous use of the Product Approach with Binary Distal Outcome

- MacKinnon & Dwyer (1993). Estimating mediated effects in prevention studies. Evaluation Review, 17, 144-158
- MacKinnon, D.P., Lockwood, C.M., Brown, C.H., Wang, W., & Hoffman, J.M. (2007). The intermediate endpoint effect in logistic and probit regression. Clinical Trials, 4, 499-513
- Mplus MODEL INDIRECT

The Problem with $a \times b$ for a Binary Outcome: 2 Parameters when 5 are Needed



- The $a \times b$ indirect effect faces the problem of non-constant effect due to ignoring the level parameters (the intercept for the mediator and threshold for the distal outcome)
- The causally-defined indirect effect uses these level parameters, focusing on the expected values of the observed binary outcome - the probabilities

The Probability of $u = 1|x$

- Conditional on x, $P(u = 1|x)$ is obtained by integrating over the residual of the mediator and apart from the regression coefficients the probability involves the residual variance, the mediator intercept, and the distal outcome threshold: 5 parameters
- One can compute $P(u = 1|x = 1)$ - $P(u = 1|x = 0)$ to compare treatment and control
- This is TE (Total Effect)
- But what are the indirect and direct effects?
- This is where the counterfactual definitions of causal effects come in

3.3 Causal Effect Definitions

- $Y_i(x)$: Potential outcome that would have been observed for that subject had the treatment variable X been set at the value x, where x is 0 or 1 in the example considered here
- The $Y_i(x)$ outcome may not be the outcome that is observed for the subject and is therefore possibly counterfactual
- The causal effect of treatment for a subject can be seen as $Y_i(1) - Y_i(0)$, but is clearly not identified given that a subject only experiences one of the two treatments
- The average effect $E[Y(1) - Y(0)]$ is, however, identifiable
- Similarly, let Y(x, m) denote the potential outcome that would have been observed if the treatment for the subject was x and the value of the mediator M was m

The Controlled Direct Effect (CDE)

The controlled direct effect is defined as

$$CDE(m) = E[Y(1,m) - Y(0,m) \mid C = c]. \qquad (5)$$

where $M = m$ for a fixed value m. The first index of the first term is 1 corresponding to the treatment group and the first index of the second term is 0 corresponding to the control group.
VanderWeele-Vansteelandt (2009):

> *While controlled direct effects are often of greater interest in policy evaluation (Pearl, 2001; Robins, 2003), natural direct and indirect effects may be of greater interest in evaluating the action of various mechanisms (Robins, 2003; Joffe et al., 2007).*

The Direct Causal Effect (DE)

The direct effect (often called the pure or **natural** direct effect) does not hold the mediator constant, but instead allows the mediator to vary over subjects in the way it would vary if the subjects were given the control condition. The direct effect is expressed as

$$DE = E[Y(1,M(0)) - Y(0,M(0)) \mid C = c] = \tag{6}$$

$$= \int_{-\infty}^{\infty} \{E[Y \mid C = c, X = 1, M = m] - E[Y \mid C = c, X = 0, M = m]\}$$

$$\times f(M \mid C = c, X = 0)\, \partial M, \tag{7}$$

where f is the density of M. A simple way to view this is to note that in Y's first argument, that is x, changes values, but the second does not, implying that Y is influenced by X only directly. The right-hand side of (7) is part of what is referred to as the Mediation Formula in Pearl (2009, 2011c).

The Total Indirect Effect (TIE)

The total indirect effect is defined as (Robins, 2003)

$$TIE = E[Y(1,M(1)) - Y(1,M(0)) \mid C = c] = \tag{8}$$

$$= \int_{-\infty}^{\infty} E[Y \mid C = c, X = 1, M = m] \times f(M \mid C = c, X = 1)\, \partial M$$

$$- \int_{-\infty}^{\infty} E[Y \mid C = c, X = 1, M = m] \times f(M \mid C = c, X = 0)\, \partial M. \tag{9}$$

A simple way to view this is to note that the first argument of Y does not change, but the second does, implying that Y is influenced by X due to its influence on M.

The Total Effect (TE)

The total effect is (Robins, 2003)

$$TE = E[Y(1) - Y(0) \mid C = c] \tag{10}$$
$$= E[Y(1, M(1)) - Y(0, M(0)) \mid C = c]. \tag{11}$$

A simple way to view this is to note that both indices are 1 in the first term and 0 in the second term. In other words, the treatment effect on Y comes both directly and indirectly due to M. The total effect is the sum of the direct effect and the total indirect effect (Robins, 2003),

$$TE = DE + TIE. \tag{12}$$

The Pure Indirect Effect (PIE)

The pure indirect effect (Robins, 2003) is defined as

$$PIE = E[Y(0, M(1)) - Y(0, M(0)) \mid C = c] \tag{13}$$

Here, the effect of X on Y is only indirect via M. This is called the natural indirect effect in Pearl (2001) and VanderWeele and Vansteelandt (2009).

Translation of Different Terms by Different Authors

Expectation Setting	Names of Effects: Imai	Pearl/VanderWeele	Robins
1,1 - 1,0	**ACME(treated)**	**TNIE**	**TIE**
0,1 - 0,0	ACME(control)	PNIE	PIE
1,0 - 0,0	**ADE(control)**	**PNDE**	**DE**
1,1 - 0,1	ADE(treated)	TNDE	-
1,1 - 0,0	Total	Total	Total

ACME - Average causal mediated effect
ADE - Average direct effect
TNIE - Total natural indirect effect
PNIE - Pure natural indirect effect
PNDE - Pure natural direct effect
TNDE - Total natural direct effect
TIE - Total indirect effect
PIE - Pure indirect effect
DE - direct effect

Causal Effect Decomposition using VanderWeele Terms

Natural (N) direct (D) and indirect (I) effect decompositions of the total effect (TE) can be expressed in two ways:

1. TE=Pure NDE +Total NIE = PNDE+TNIE (= DE + TIE)
   - E[Y(1,M(1))-Y(0,M(0))] = E[Y(1,M(0))-Y(0,M(0))]+E[Y(1,M(1))-Y(1,M(0))]

2. TE= Total NDE + Pure NIE = TNDE+PNIE
   - E[Y(1,M(1))-Y(0,M(0))] = E[Y(1,M(1))-Y(0,M(1))]+E[Y(0,M(1))-Y(0,M(0))]

1 is the focus of Valeri-VanderWeele (2013).

The Causal Effect Approach is General

The causal effects are expressed in a general way using expectations and can be applied to many different settings:

- Continuous mediator, continuous distal outcome (gives the usual SEM formulas)
- Categorical mediator, continuous distal outcome
- Continuous mediator, categorical distal outcome
- Categorical mediator, categorical distal outcome
- Count distal outcome
- Nominal mediator, nominal outcome
- Survival distal outcome

Binary Y, Continuous M

Using the general definition, the causal natural indirect effect (total indirect effect) is expressed as the probability difference

$$Total\ NIE = TIE = \Phi[probit(1,1)] - \Phi[probit(1,0)], \qquad (14)$$

where Φ is the standard normal distribution function, the argument (a, b) = (x, M(x)), and probit is defined on the next slide.
The pure natural indirect effect is expressed as the probability difference

$$Pure\ NIE = PIE = \Phi[probit(0,1)] - \Phi[probit(0,0)]. \qquad (15)$$

and the pure natural direct effect expressed as the probability difference

$$Pure\ NDE = DE = \Phi[probit(1,0)] - \Phi[probit(0,0)], \qquad (16)$$

$$TE = Pure\ NDE + Total\ NIE = DE + TIE = \Phi[probit(1,1)] - \Phi[probit(0,0)]. \qquad (17)$$

Consider a mediation model for a binary outcome u and a continuous mediator m. Assume a probit link for the binary outcome *u*,

$$probit(u_i) = \beta_0 + \beta_1\, m_i + \beta_2\, x_i + \beta_3\, x_i\, m_i + \beta_4\, c_i, \tag{18}$$

$$m_i = \gamma_0 + \gamma_1\, x_i + \gamma_2\, c_i + \varepsilon_{2i}, \tag{19}$$

where the residual ε_2 is assumed normally distributed. For x, x' = 0, 1 corresponding to the control and treatment group,

$$probit(x,x') = [\beta_0 + \beta_2\, x + \beta_4\, c + (\beta_1 + \beta_3\, x)(\gamma_0 + \gamma_1\, x' + \gamma_2\, c)]/\sqrt{v(x)}, \tag{20}$$

where the variance $v(x)$ for x = 0, 1 is

$$v(x) = (\beta_1 + \beta_3\, x)^2\, \sigma_2^2 + 1. \tag{21}$$

where σ_2^2 is the residual variance for the continuous mediator m. Although not expressed in simple functions of model parameters, the quantity of (14) can be computed and corresponds to the change in the y=1 probability due to the indirect effect of the treatment (conditionally on c when that covariate is present).

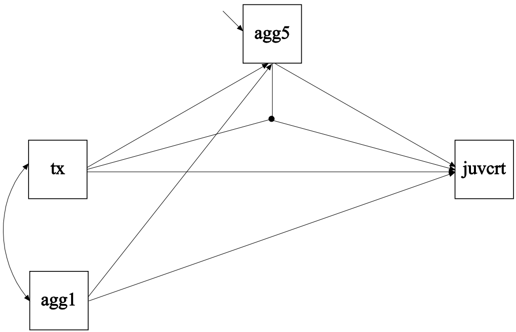
3.4 Example: Aggressive Behavior and Juvenile Court Record

- Randomized field experiment in Baltimore public schools
- Classroom-based intervention aimed at reducing aggressive-disruptive behavior among elementary school students
- Covariate is the Grade 1 aggression score before the intervention started
- Mediator is the aggression score in Grade 5 after the intervention ended
- Distal outcome is a binary variable indicating whether or not the student obtained a juvenile court record by age 18 or an adult criminal record
- $n = 250$ boys in treatment and control classrooms

A Mediation Model for Aggressive Behavior and a Binary Juvenile Court Outcome

Two reasons for causal effects:

- ”Exposure-mediator interaction” (tx*agg5)
- Binary outcome (juvcrt)



A Mediation Model for Aggressive Behavior and a Binary Juvenile Court Outcome

$$juvcrt_i^* = \beta_0 + \beta_1\, agg5_i + \beta_2\, tx_i + \beta_3\, tx_i\, agg5_i + \beta_4\, agg1_i + \varepsilon_{1i}, \quad (22)$$
$$agg5_i = \gamma_0 + \gamma_1\, tx_i + \gamma_2\, agg1_i + \varepsilon_{2i}. \quad (23)$$

The juvcrt outcome is not rare, but is observed for 50% of the sample. The mediator agg5 is not normally distributed, but is quite skewed with a heavy concentration at low values. The normality assumption, however, pertains to the mediator residual ε_2. Because the covariate agg1 has a distribution similar to the mediator agg5, the agg5 distribution is to some extent produced by the agg1 distribution so that the normality assumption for the residual may be a reasonable approximation.

Aggressive Behavior and Juvenile Court Record: Mplus Input for Causal Effects

```
Analysis:
        estimator = mlr;
        link = probit;
        integration = montecarlo;

model:
        [juvcrt$1] (mbeta0);
        juvcrt on tx (beta2)
        agg5 (beta1)
        xm (beta3)
        agg1 (beta4);
        [agg5] (gamma0);
        agg5 on tx (gamma1)
        agg1 (gamma2);
        agg5 (sig2);
```

Aggressive Behavior and Juvenile Court Record: Mplus Input for Causal Effects, Continued - Done Automatically in Mplus Version 7.2

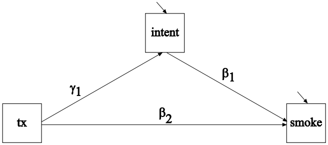
```
model constraint:

new(ind dir arg11 arg10 arg00 v1 v0
probit11 probit10 probit00 indirect direct
total iete dete compdete orind ordir);
dir=beta3*gamma0+beta2;
ind=beta1*gamma1+beta3*gamma1;
arg11=-mbeta0+beta2+beta4*0+(beta1+beta3)*(gamma0+gamma1+gamma2*0);
arg10=-mbeta0+beta2+(beta1+beta3)*gamma0;
arg00=-mbeta0+beta1*gamma0;
v1=(beta1+beta3)^2*sig2+1;
v0=beta1^2*sig2+1;
probit11=arg11/sqrt(v1);
probit10=arg10/sqrt(v1);
probit00=arg00/sqrt(v0);
! Version 6.12 Phi function needed below:
indirect=phi(probit11)-phi(probit10);
direct=phi(probit10)-phi(probit00);
total=phi(probit11)-phi(probit00);
orind=(phi(probit11)/(1-phi(probit11)))/(phi(probit10)/(1-phi(probit10)));
ordir=(phi(probit10)/(1-phi(probit10)))/(phi(probit00)/(1-phi(probit00)));
```

The causal direct effect is not significant. The causal indirect effect is estimated as -0.064 and is significant ($z = -2.120$). This is the drop in the probability of a juvenile court record due to the indirect effect of treatment.

The odds ratio for the indirect effect is estimated as 0.773 which is significantly different from one ($z = (0.773 - 1)/0.092 = -2.467$).

The conventional direct effect is not significant and the conventional product indirect effect is -0.191 (z= -1.98).

Muthén (2011): MacKinnon et al (2007) smoking data (binary Y)



	Intention	No Use	Use	Total
		Cigarette use		
Ctrl	4 (Yes)	9	20	29
	3 (Probably)	14	20	34
	2 (Don't think so)	36	13	49
	1 (No)	229	30	259
Tx	4 (Yes)	9	19	28
	3 (Probably)	15	11	26
	2 (Don't think so)	43	11	54
	1 (No)	353	32	385

Smoking Data Continued

Different approaches with an ordinal mediator:

- Ordered polytomous variable treated as continuous (non-normal residual issue)
- Latent response variable behind ordered polytomous variable
- Dichotomized variable
- Latent response variable behind dichotomized variable

See Muthén (2011)

3.6 Binary Mediator and Binary Distal Outcome

Recalling that the general formulas for the direct, total indirect, and pure indirect effects are defined as

$$DE = E[Y(1, M(0)) - Y(0, M(0)) \mid C], \quad (24)$$

$$TIE = E[Y(1, M(1)) - Y(1, M(0)) \mid C], \quad (25)$$

$$PIE = E[Y(0, M(1)) - Y(0, M(0)) \mid C], \quad (26)$$

it can be shown that with a binary mediator and a binary outcome these formulas lead to the expressions

$$DE = [F_Y(1,0) - F_Y(0,0)]\,[1 - F_M(0)] + [F_Y(1,1) - F_Y(0,1)]\,F_M(0), \quad (27)$$

$$TIE = [F_Y(1,1) - F_y(1,0)]\,[F_M(1) - F_m(0)], \quad (28)$$

$$PIE = [F_Y(0,1) - F_y(0,0)]\,[F_M(1) - F_m(0)]. \quad (29)$$

where $F_Y(x,m)$ denotes $P(Y = 1 \mid X = x, M = m)$ and $F_M(x)$ denotes $P(M = 1 \mid X = x)$, where F denotes either the standard normal or the logistic distribution function corresponding to using probit or logistic regression. These formulas agree with those of Pearl (2010, 2011a).

Pearl's Hypothetical Binary-Binary Case

Pearl (2010, 2011a) provided a hypothetical example with a binary treatment X, a binary mediator M corresponding to the enzyme level in the subject's blood stream, and a binary outcome Y corresponding to being cured or not. This example was also hotly debated on SEMNET in September 2011.

Pearl's Hypothetical Binary-Binary Case, Continued

Treatment X	Enzyme M	Percentage Cured Y = 1
1	1	$F_Y(1,1) = 80\%$
1	0	$F_Y(1,0) = 40\%$
0	1	$F_Y(0,1) = 30\%$
0	0	$F_Y(0,0) = 20\%$

Treatment	Percentage M=1
0	$F_M(0) = 40\%$
1	$F_M(1) = 75\%$

The top part of the table suggests that the percentage cured is higher in the treatment group for both enzyme levels and that the effect of treatment is higher at enzyme level 1 than enzyme level 0: Treatment-mediator interaction.

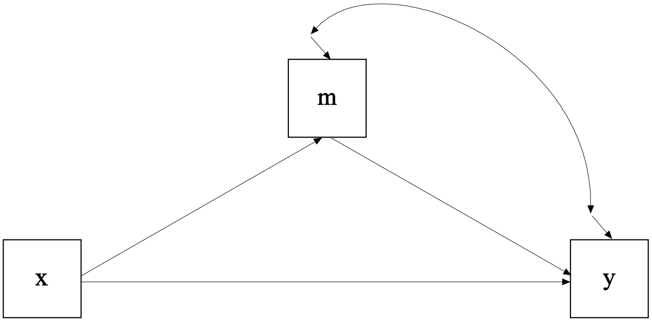
3.7 Sensitivity Analysis of Mediator-Outcome Confounding

To claim that effects are causal, it is not sufficient to simply use the causally-derived effects

The underlying assumptions need to be fulfilled, such as no mediator-outcome confounding

Violation of the no mediator-outcome confounding can be seen as an unmeasured (latent) variable Z influencing both the mediator M and the outcome Y. When Z is not included in the model, a covariance is created between the residuals in the two equations of the regular mediation model. Including the residual covariance, however, makes the model not identified.

Mediator-Outcome Confounding (Residual Correlation $\rho \neq 0$)

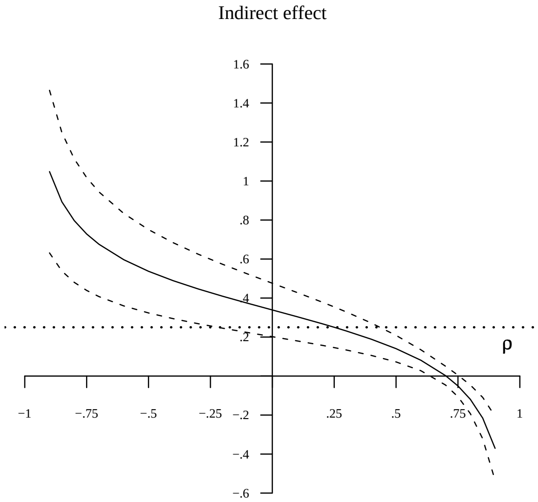


Imai Sensitivity Analysis

Imai et al. (2010a, b) proposed a sensitivity analysis where causal effects are computed given different fixed values of the residual covariance. This is useful both in real-data analyses as well as in planning studies. As for the latter, the approach can answer questions such as how large does your sample and effects have to be for the lower confidence band on the indirect effect to not include zero when allowing for a certain degree of mediator-outcome confounding?

Sensitivity plots can be made in Mplus using LOOP in the PLOT command.

Indirect Effect Based on Imai Sensitivity Analysis with ρ Varying from -0.9 to +0.9 and True Residual Correlation 0.25



Explaining the Sensitivity Figure

- The correct value for the indirect effect is 0.25 (marked with a horizontal broken line)
- The biased estimate assuming $\rho = 0$ is 0.3287, an overestimation due to ignoring the positive residual correlation
- The sensitivity analysis varies the ρ values from -0.9 to $+0.9$:
  - Using $\rho = 0$, the biased estimate of 0.3287 is obtained
  - Using the correct value of $\rho = 0.25$, the correct indirect effect value of 0.25 is obtained
  - For lower ρ values the effect is overestimated and for larger ρ values the effect is underestimated
- The graph provides useful information for planning new studies:
  - At this sample size ($n = 400$) and effect size, the lower confidence limit does not include zero until about $\rho = 0.6$
  - This means that a rather high degree of confounding is needed for the effect to not be detected
  - In the range of ρ from about -0.1 to +0.4 the confidence interval covers the correct value of 0.25 for the indirect effect