Estimating and testing mediated effects with binary mediators and/or binary outcomes

Steve Gregorich

CAPS Methods Core

Main foci

. Simple and reasonable methods for estimating and testing mediated effects

. $\underline{a}\times\underline{b}$ method for estimation of indirect effects and standard errors

  . the $\underline{c}$ - $\underline{c}'$ method has some limitations (we will discuss them)

  . bootstrap standard error estimates & CIs are available; I do not cover them

. I do not consider SEM approaches

  . essentially requires reporting standardized coefficients

. 3-variable systems

  . extensions for multiple mediators and/or covariates are briefly addressed

. Limited Monte Carlo simulation to assess bias and coverage

If you do not already have a basic understanding of testing mediation and associated terminology, then this talk may move a little fast for you

8 combinations of X (explanatory var), M (mediator), and Y (outcome)

The current methodological 'lay of the land' (kind of)…

distribution of M	distribution of X	distribution of Y	
		continuous	binary
continuous	binary	Baron & Kenny (1986)	MacKinnon & Dwyer (1993)?
continuous	continuous		
binary	binary	Li et al (2007)	Huang et al (2004)
binary	continuous		Hmm…?

(note that I put M in the first table column)

8 combinations of X, M, and Y

…is not very satisfying

distribution of M	distribution of X	distribution of Y	
		continuous	binary
continuous	binary	Baron & Kenny (1986)	MacKinnon & Dwyer (1993)?
continuous	continuous		
binary	binary	Li et al (2007)	Huang et al (2004)
binary	continuous		Hmm…?

MacKinnon & Dwyer (1993) has shortcomings (c – c' method)

Nothing wrong with Huang et al (2004), per se, but Li et al (2007) covers it

I don't think continuous X with binary M and Y has been directly addressed

8 combinations of X, M, and Y

'Cut to the chase'

distribution of M	distribution of X	distribution of Y: continuous	distribution of Y: binary
continuous	continuous or binary	(1) Baron & Kenny (1986)	(2) Baron & Kenny (1986)
binary	binary	(3) Li et al (2007): Eq #14	
binary	continuous	(4) Li et al (2007): Eq #13	

The methods described by Baron & Kenny (1986) and Li et al (2007)

naturally generalize to models with binary Y,

but, to my knowledge, such applications have not been described although the last 2 sentences of Li et al (2007) state

"Our article provides the definition of the mediation effect for a binary mediator model with a continuous outcome. The definition can be generalized to a causal model where the set of link functions in equations (1)–(2) are specified by a generalized linear model."

8 combinations of X, M, and Y

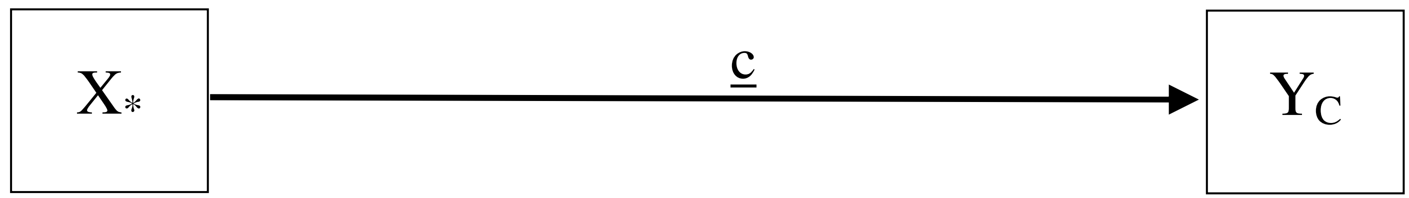
distribution of M	distribution of X	distribution of Y	
		continuous	binary
continuous	continuous or binary	Baron & Kenny (1986)	
binary	binary	Li et al (2007): Eq #14	
	continuous	Li et al (2007): Eq #13	

First—

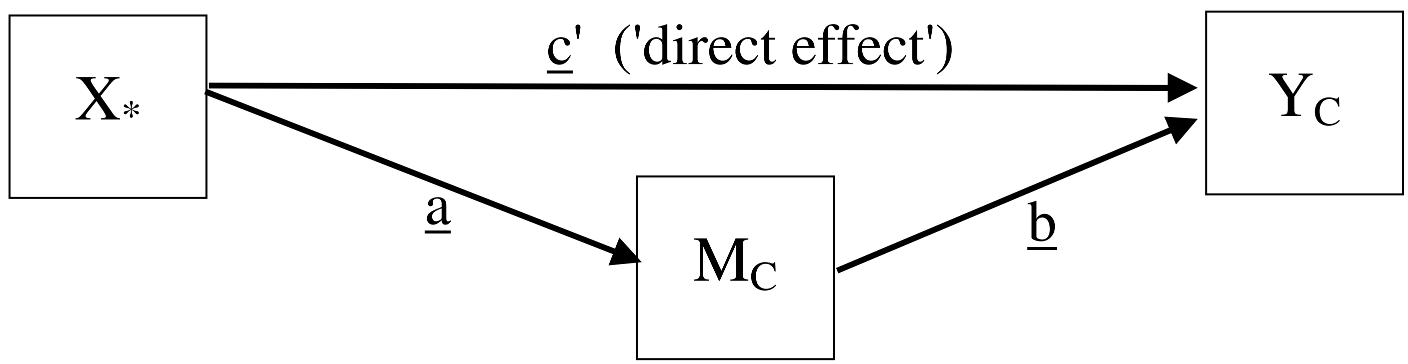
Mediation basics—Models with Continuous M and Y, and Continuous or Binary X

Mediation basics: Continuous M and Y:

Total effect model



3-variable path model



Mediated, or indirect effect

. $a \times b$, or

. $c - c'$

what is the logic behind $a \times b$?

$c = a \times b + c'$ (total effect = indirect effect + direct effect)

Mediation basics: Continuous M and Y modeling steps

<u>Steps</u>

1. Fit the total effect and 3-variable path model

2. Save the parameter and standard error estimates

3. Estimate the indirect (mediated) effect

$\underline{a} \times \underline{b}$, or

$\underline{c} - \underline{c}'$

4. Estimate the corresponding standard error and test the indirect effect

Mediation basics: Continuous M and Y modeling steps

1. Fit the total effect and 3-variable path model

Often this is done by fitting 3 models in a piecewise fashion

$$y_i = \alpha_0 + x_i\underline{c} + \varepsilon_{0i} \text{ (Model 1: total effect model)}$$

3-variable path model:

$$\begin{cases} y_i = \alpha_1 + x_i\underline{c}' + m_i\underline{b} + \varepsilon_{1i} & \text{(Model 2: direct effects model)} \\ m_i = \alpha_2 + x_i\underline{a} + \varepsilon_{2i} & \text{(Model 3: X} \rightarrow \text{M model)} \end{cases}$$

2. Save the parameter and standard error estimates

3. Estimate the indirect (mediated) effect as

$\underline{a}\times\underline{b}$, or

$\underline{c} - \underline{c}'$

Mediation basics: Continuous M and Y
Estimating the standard error of $\underline{c} - \underline{c}'$

Steps

4. Estimate the corresponding standard error and test the indirect effect

The $\underline{c} - \underline{c}'$ method

$$z_{\underline{c}-\underline{c}'} = \left(\underline{c} - \underline{c}'\right) \Big/ \sqrt{\sigma^2_{\underline{c}} + \sigma^2_{\underline{c}'} - 2\sigma_{MSE} \Big/ \left(N\sigma^2_X\right)},$$

where σ_{MSE} is the MSE from the direct effects model

I do not focus on the $\underline{c} - \underline{c}'$ method

it provides an '*indirect estimate of the indirect effect*'
that doesn't perform well in all situations

Mediation basics: Continuous M and Y
Estimating the standard error of $\underline{a}\times\underline{b}$

Steps

4. Estimate the corresponding standard error and test the indirect effect

$$z_{\underline{ab}} = \underline{a}\times\underline{b} / se_{\underline{ab}},$$

where, for se_{ab}, you have a choice between

. $se_{Sobel} = \sqrt{\underline{a}^2\sigma_{\underline{b}}^2 + \underline{b}^2\sigma_{\underline{a}}^2}$

. $se_{Aroian} = \sqrt{\underline{a}^2\sigma_{\underline{b}}^2 + \underline{b}^2\sigma_{\underline{a}}^2 + \sigma_{\underline{a}}^2\sigma_{\underline{b}}^2}$, 'unbiased' (generally preferred)

. $se_{Goodman} = \sqrt{\underline{a}^2\sigma_{\underline{b}}^2 + \underline{b}^2\sigma_{\underline{a}}^2 - \sigma_{\underline{a}}^2\sigma_{\underline{b}}^2}$, 'exact' (neg. var. estim. problem)

Mediation basics: Continuous M and Y
Estimating the standard error of $\underline{a} \times \underline{b}$

Assumptions of Sobel, Aroian, and Goodman variance estimators

. correctly specified model

. no measurement error in variables

. $\underline{a}$ and $\underline{b}$ are independent

Mediation basics: Continuous M and Y
Estimating the standard error of $\underline{a} \times \underline{b}$

. Sobel, Aroian, and Goodman var. estimates are asymptotic—bias < 5% with
   N>100 for a single mediator (MacKinnon et al, 1995)
   N>200 for a model with 7 indirect effects (Stone & Sobel, 1990)

. They assume $\underline{a} \times \underline{b}$ is normally distributed, but, $\underline{a} \times \underline{b}$ can be skewed
   So for confidence interval estimation you may want to bootstrap

. Under distributional violation, these tests are underpowered
   Therefore, conservative wrt power
   OK for 'quick' screening tests of indirect effects

. The Monte Carlo simulation results that I present today are based upon N=500

Next, Monte Carlo simulations for continuous and binary X

General method for data simulation

1. generate a continuous standard normal or binary (50/50) random X variate

2. generate random residuals, e_1, from a standard logit distribution $\left(\sigma_e^2 = \pi^2/3\right)$

3a. use X, e_1, and model parameters to generate continuous M values

$$m_i^* = \text{int}_1 + x_i \underline{a} + e_{1i}$$

3b. compute binary M values

if $m_i^*>0$ then $m_i=1$; else $m_i=0$

4a. use X, M, newly generated e_2, and model to generate continuous Y values

$y_i^* = \text{int}_2 + x_i \underline{c}' + m_i^* \underline{b} + e_{2i}$ (with continuous M) or

$y_i^* = \text{int}_2 + x_i \underline{c}' + m_i \underline{b} + e_{2i}$ (with binary M)

4b. compute binary Y values

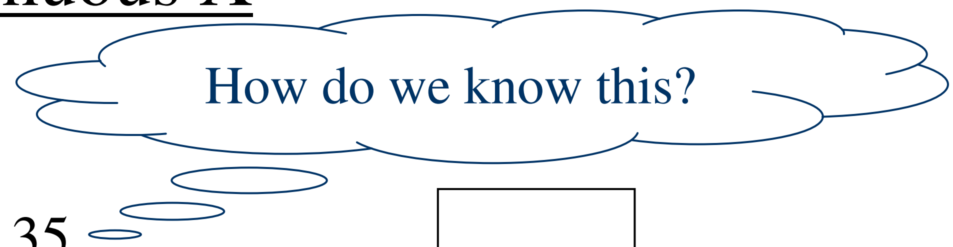
if $y_i^*>0$ then $y_i=1$; else $y_i=0$

Notes

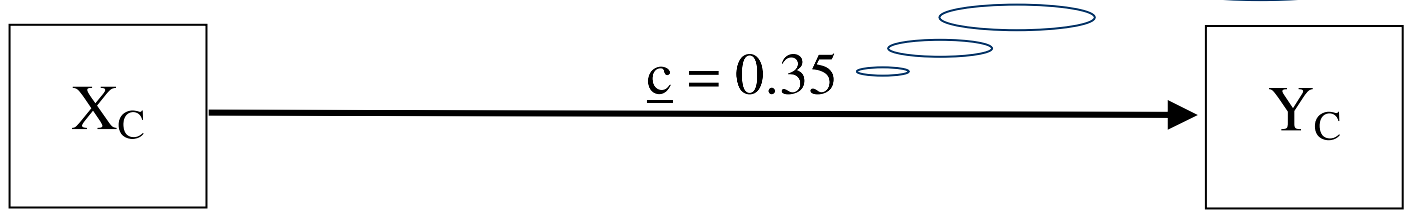
. total effect, $\underline{c}$, is never directly specified

. equivalent parameters for linear and logistic models

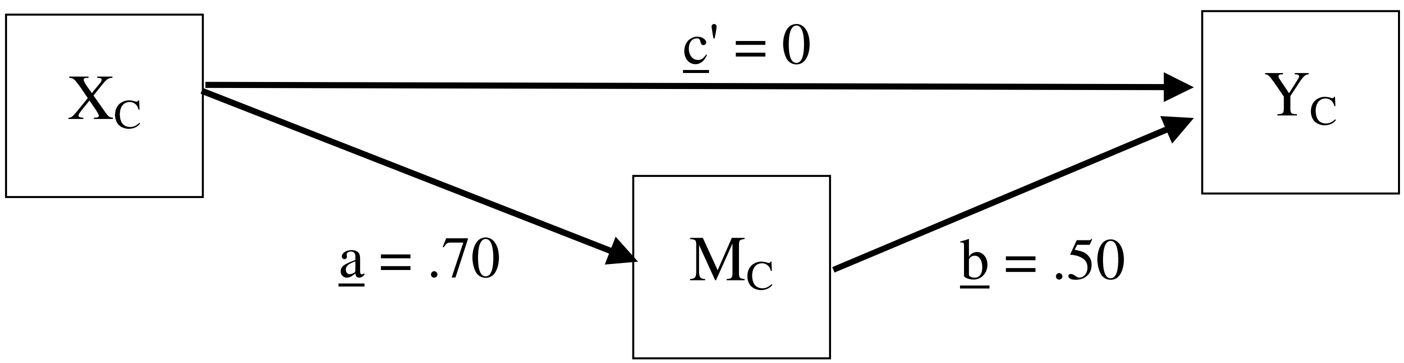
Mediation basics: Continuous M and Y: Simulation: Set-up with Continuous X



Total effect linear model



3-variable linear path model



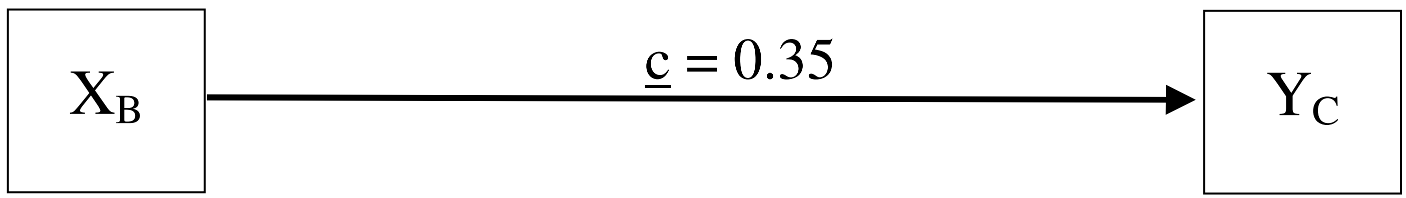
$$\sigma_X^2 = 1.0, \ \sigma_M^2 = 3.78, \ \sigma_Y^2 = 4.24$$

10,000 replicate samples of N=500

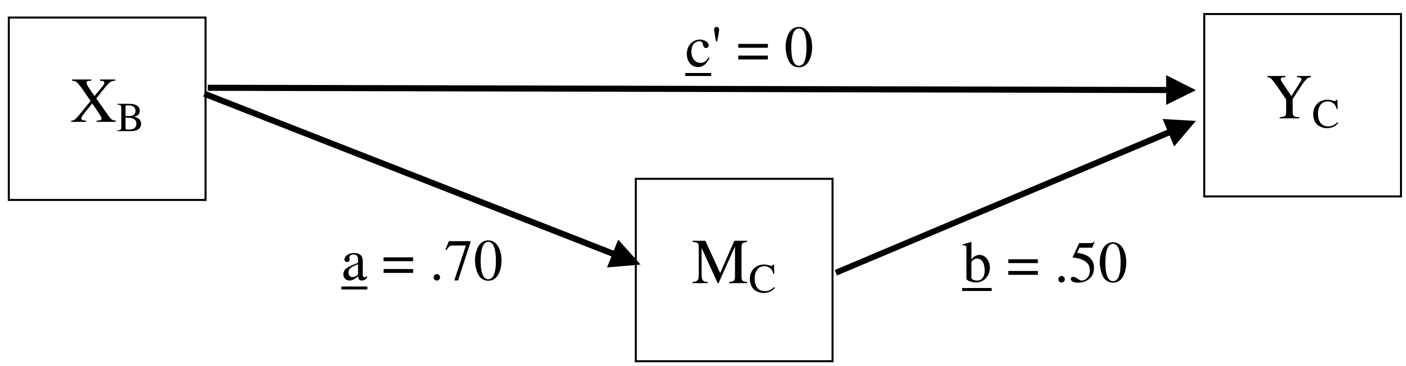
Direct effect=0 in the population, so Total effect=Indirect effect

Mediation basics: Continuous M and Y
Simulation: Set-up with Binary X

Total effect linear model



3-variable linear path model



$$\sigma_X^2 = 0.25,\ \sigma_M^2 = 3.41,\ \sigma_Y^2 = 4.14$$

10,000 replicate samples of *N*=500

Direct effect=0 in the population, so Total effect=Indirect effect

Mediation basics: Continuous M and Y: Simulation: Parameter estimates

X		c total effect	a X→M	b M→Y	c' direct effect	a×b indirect effect	a×b+c' total effect
contin-uous	pop. value	.3500	.7000	.5000	0	.3500	.3500
	estimate	.3499	.7001	.4992	.0004	.3495	.3499
binary	pop. value	.3500	.7000	.5000	0	.3500	.3500
	estimate	.3479	.7000	.4992	-.0015	.3493	.3479

Mediation basics: Continuous M and Y: Simulation: Standard error estimates

X		c total effect	a X→M	b M→Y	c' direct effect	a×b indirect effect Sobel	 Aroian	 Goodman
contin-uous	pop. val.	.0911	.0816	.0444	.0879	.0514	.0514	.0514
	estimate	.0909	.0813	.0449	.0872	.0515	.0516	.0514
binary	pop. val.	.1821	.1624	.0444	.1669	.0875	.0875	.0875
	estimate	.1815	.1623	.0449	.1656	.0872	.0875	.0869

Notes

. Each standard error population value represents the standard deviation of the corresponding parameter (point) estimate across the 10,000 replicate samples

. Estimated standard errors represent the mean standard error estimate across the 10,000 replicate samples

Mediation basics: Continuous M and Y: Simulation: Coverage

X	c total effect	a X→M	b M→Y	c' direct effect	a×b indirect effect Sobel	 Aroian	 Goodman
continuous	.9503	.9470	.9515	.9486	.9505	.9511	.9492
binary	.9507	.9494	.9512	.9476	.9478	.9490	.9471

coverage estimates of known population values

Continuous M and Y
Summary

With continuous M, and Y

the $\underline{a} \times \underline{b}$ and $\underline{c} - \underline{c}'$ methods are equivalent, unbiased

In this simulation (*N*=500)…

Sobel, Aroian, and Goodman standard error estimates unbiased

w/ good coverage

Next—Continuous M and Binary Y

8 combinations of X, M, and Y

distribution of M	distribution of X	distribution of Y	
		continuous	binary
continuous	continuous or binary	Baron & Kenny (1986)	
binary	binary	Li et al (2007): Eq #14	
	continuous	Li et al (2007): Eq #13	

OK, so we've covered Continuous M and Y

Next—Continuous M and Binary Y

We will be using the a×b method to estimate indirect effects,
but first we will review the shortfall of the $\underline{c} - \underline{c}'$ *method with Binary Y*

Background for Binary Y

The $\underline{c} - \underline{c}'$ method with Binary Y

When Y is binary, the $\underline{c} - \underline{c}'$ estimate of the indirect effect is problematic

Note that the $\underline{c} - \underline{c}'$ estimate is based upon comparing parameter estimates across two nested logistic regression models

$$\text{logit}(y_i=1 \mid x_i) = \alpha_0 + x_i\underline{c} \qquad \text{[total effect model]}$$

$$\text{logit}(y_i=1 \mid x_i, m_i) = \alpha_1 + x_i\underline{c}' + m_i\underline{b} \qquad \text{[direct effects model]}$$

This comparison of parameters across nested logistic models is the source of the problem

Background for Binary Y

The c – c' method with Binary Y

Comparing coefficients across nested models

Linear models (included here for context)

. variance of the outcome variable is observed
. variance of the residual is estimated
. as significant explanatory variables are added to the model
  the residual variance decreases

Logistic models

. variance of the continuous variate underlying the outcome is estimated
. variance of the residual is fixed
. as significant explanatory variables are added to the model
  the implied outcome variance increases

Result for logistic models

. as explanatory variables are added to logistic models, the scaling of the
  outcome variable and, thus, the scaling of parameter estimates is altered.
. cannot directly compare parameter estimates across nested logistic models

Background for Binary Y

The c – c' method with Binary Y

Comparing coefficients across nested logistic models

Simulated data—a somewhat different simulation approach just for this topic

. A single sample with N=500,000

. x_1 and x_2 are bivariate normal with the following sample statistics (exactly)

. $\bar{x}_1 = \bar{x}_2 = 0$

. $\sigma^2_{x1} = 1, \quad \sigma^2_{x2} = 4$

. $r_{x_1 x_2} = 0$

Next, I used x_1 and x_2 values to generate a continuous y^* variate as

$y_i^* = x_{1i} + x_{2i} + e_i,$ (i.e., both regression parameters equaled unity)

where the e_i followed a standard logistic distribution

Finally, I created a binary version of y^* as

$y = 1$ if $y^* > 0$;

y = 0 otherwise

Background for Binary Y

The <u>c</u> – <u>c</u>' method with Binary Y

Results of <u>Linear</u> models regressing y^* onto $x1$ and $x2$

M1: *Multivariate model*

$$\hat{B}_0 = 0.00$$

$$\hat{B}_1 = 1.00$$

$$\hat{B}_2 = 1.00$$

M2: *Bivariate effect of x1*

$$\hat{B}_0 = 0.00$$

$$\hat{B}_1 = 1.00$$

M3: *Bivariate effect of x2*

$$\hat{B}_0 = 0.00$$

$$\hat{B}_2 = 1.00$$

$$r_{x1x2} = 0$$

Background for Binary Y

The c – c' method with binary Y

Results of Logistic models regressing *y* onto *x*1 and *x*2

M0: *Multivariate model*

$$\hat{B}_{0(M0)} = 0.00$$
$$\hat{B}_{1(M0)} = 0.99$$
$$\hat{B}_{2(M0)} = 0.99$$

M1: *Bivariate effect of x1*

$$\hat{B}_{0(M1)} = 0.00$$
$$\hat{B}_{1(M1)} = 0.61$$

M2: *Bivariate effect of x2*

$$\hat{B}_{0(M2)} = 0.00$$
$$\hat{B}_{2(M2)} = 0.85$$

but… r_{x1x2} is exactly zero…!?

Background for Binary Y

The c – c' method with binary Y

Explanation for results on previous slide

In this simplified example, $x1$ and $x2$ are orthogonal, so the implied variance of y^* equals

Multivariate model

$$\sigma^2_{y^*(\mathrm{M0})} = \sigma^2_{x1}\mathrm{B}^2_{1(\mathrm{M0})} + \sigma^2_{x2}\mathrm{B}^2_{2(\mathrm{M0})} + \pi^2/3 = 8.29$$

Bivariate model with $x1$

$$\sigma^2_{y^*(\mathrm{M1})} = \sigma^2_{x1}\mathrm{B}^2_{1(\mathrm{M1})} + \pi^2/3 = 4.29$$

Bivariate model with $x2$

$$\sigma^2_{y^*(\mathrm{M2})} = \sigma^2_{x2}\mathrm{B}^2_{2(\mathrm{M2})} + \pi^2/3 = 7.29$$

Scaling of the outcome and parameter estimates is not equivalent across models

Background for Binary Y

The $\underline{c} - \underline{c}'$ method with binary Y

Comparing coefficients across nested regression models

MacKinnon & Dwyer (1993) suggested a rescaling of model parameters to allow use of $\underline{c} - \underline{c}'$ with binary Y

However, the MacKinnon & Dwyer method

. requires saving intermediate results and applying some matrix algebra

. essentially standardizes the parameters of the total and direct effects models

. does not work very well; $\underline{c} - \underline{c}' \approx \underline{a} \times \underline{b}$ even after parameter rescaling

Standardization of model parameters complicates subsequent interpretation

However, *that* issue is easily addressed by scaling the parameters of the total effects model to approximate the scale of the direct effects model

Regardless, I want something that works better and is easier to apply

8 combinations of X, M, and Y

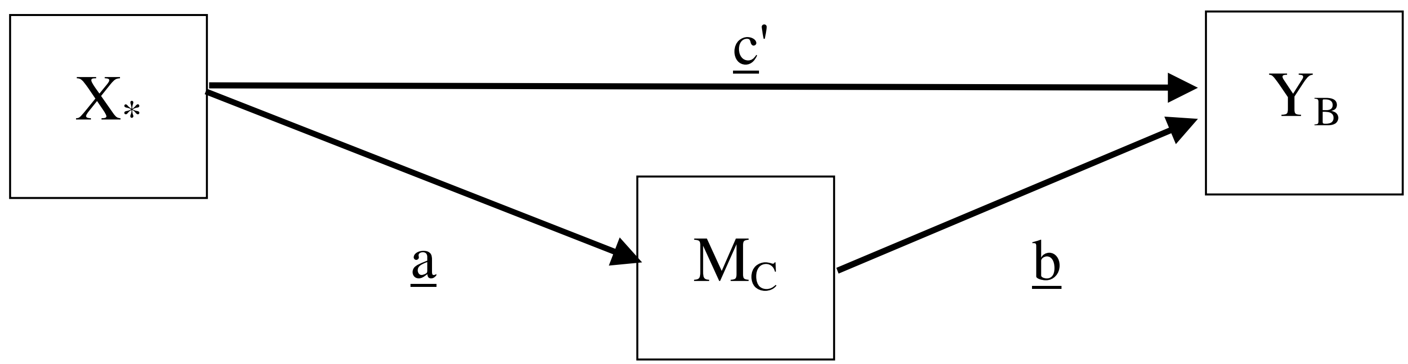
distribution of M	distribution of X	distribution of Y	
		continuous	binary
continuous	continuous or binary	Baron & Kenny (1986)	
binary	binary	Li et al (2007): Eq #14	
	continuous	Li et al (2007): Eq #13	

Now (finally) we move on to consider models with Continuous M and Binary Y

But, we will not base tests upon MacKinnon & Dwyer (1993)

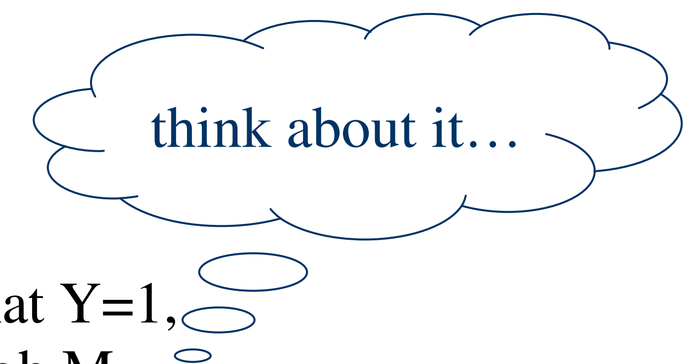
Instead, for models with Continuous M and Binary Y we will use the classic method of Baron & Kenny

Continuous M and Binary Y



. b equals the expected change in the log odds that Y=1 (versus 0), *caused* by a one-unit change in M

. a represents the expected change in M, *caused* by a one-unit change in X



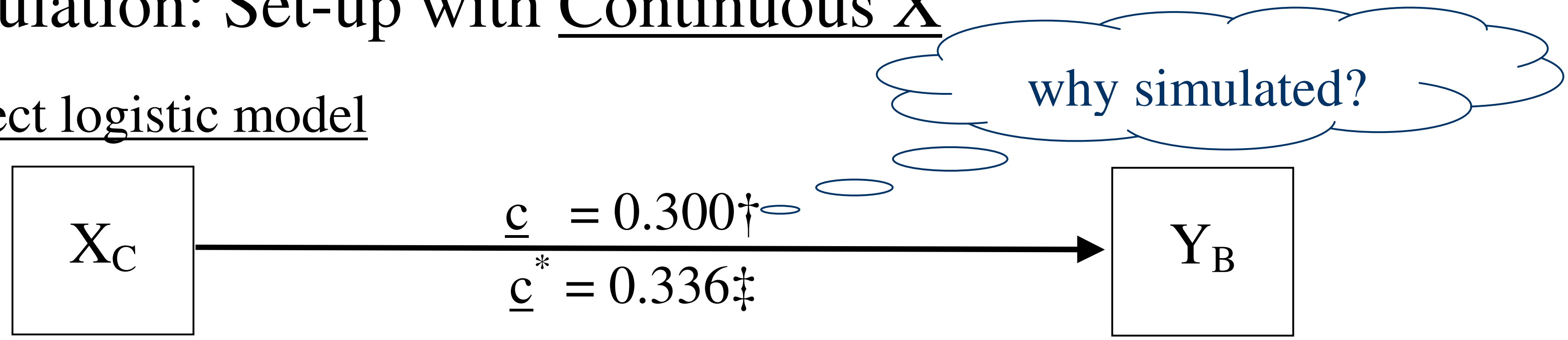
. a×b represents the expected change in the log odds that Y=1, caused by a one-unit change in X channeled through M

. b represents the effect on Y given a one-unit change in M.
However, a one-unit change in X results in an a-unit change in M
So, a one-unit change in X has an a×b effect on Y (via M)

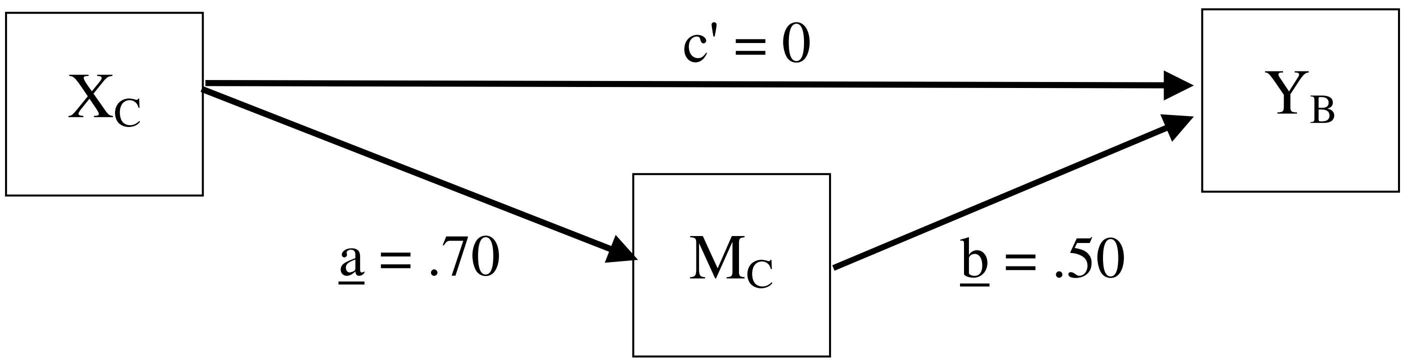
. Estimate the indirect effect as a×b and use Sobel or Aroian test

Continuous M and Binary Y:
Simulation: Set-up with <u>Continuous X</u>

<u>Total effect logistic model</u>



<u>3-variable linear and logistic path model</u>



$$\sigma_X^2 = 1.0,\ \sigma_M^2 = 3.78,\ \sigma_Y^2 = 0.25$$

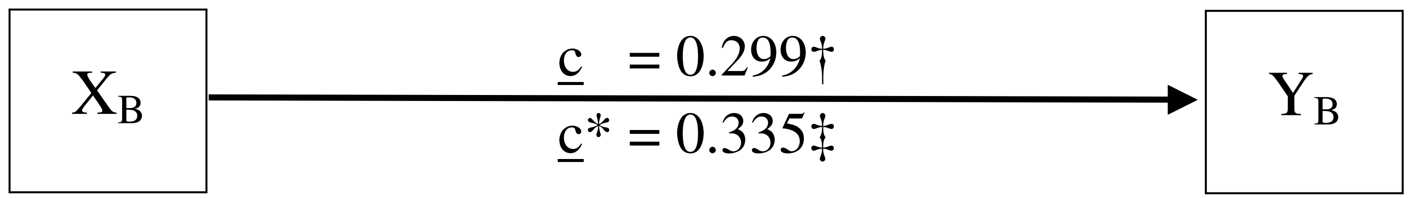
10,000 replicate samples of N=500

The total and direct effects models have different scales, so $\underline{c}^* - \underline{c}' \approx \underline{a} \times \underline{b} = 0.35$

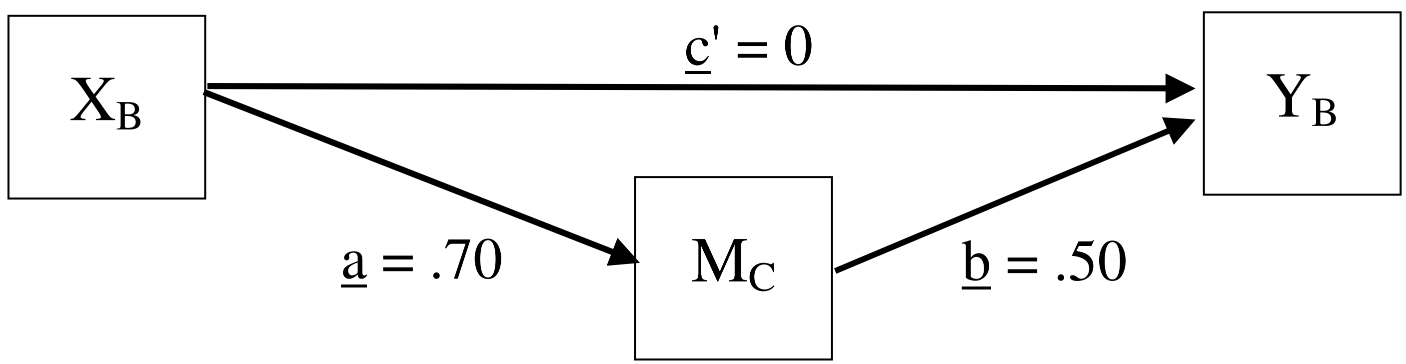
† estimated by simulation; ‡ rescaled by factor 1.119

Continuous M and Binary Y: Simulation: Set-up with Binary X

Total effect logistic model



3-variable linear and logistic path model



$\sigma_X^2 = 0.25, \ \sigma_M^2 = 3.41, \ \sigma_Y^2 = 0.25$

10,000 replicate samples of N=500

The total and direct effects models have different scales, so c - c' ≈ a×b = 0.35

† estimated by simulation; ‡ rescaled by factor 1.118

Continuous M and Binary Y: Simulation: Parameter estimates

X		c total effect	a X→M	b M→Y	c' direct effect	a×b indirect effect	a×b+c' total effect
contin-uous	pop. value	.3004†	.7000	.5000	0	.3500	.3500
	estimate	.3005‡	.7001	.5051	-.0008	.3537	.3529
binary	pop. value	.2992†	.7000	.5000	0	.3500	.3500
	estimate	.3028‡	.6997	.5049	-.0028	.3534	.3506

† population value of total effect estimated by simulation with 5M records. Not rescaled.

‡ not rescaled

Continuous M and Binary Y: Simulation: Standard error estimates

X		c total effect	a X→M	b M→Y	c' direct effect	a×b indirect effect Sobel	Aroian	Goodman
contin-uous	pop. val.	.0946‡	.0816	.0664	.1089	.0626	.0626	.0626
	estimate	.0930‡	.0813	.0661	.1064	.0621	.0623	.0618
binary	pop. val.	.1802‡	.1624	.0658	.2021	.0955	.0955	.0955
	estimate	.1805‡	.1623	.0656	.1983	.0946	.0952	.0940

. Each standard error population value represents the standard deviation of the corresponding parameter estimate across the 10,000 replicate samples

. Estimated standard errors represent the mean standard error estimate across the 10,000 replicate samples

‡ not rescaled

Continuous M and Binary Y: Simulation: Coverage

X	c total effect	a X→M	b M→Y	c' direct effect	a×b indirect effect Sobel	a×b indirect effect Aroian	a×b indirect effect Goodman
continuous	.9490†	.9473	.9511	.9435	.9471	.9482	.9462
binary	.9498†	.9494	.9505	.9471	.9467	.9482	.9452

† coverage of simulated population value

Continuous M
Summary

distribution of M	distribution of X	distribution of Y	
		continuous	binary
continuous	continuous or binary	Baron & Kenny (1986)	
binary	binary	Li et al (2007): Eq #14	
	continuous	Li et al (2007): Eq #13	

Main points

. With continuous M, regardless of the whether X and Y are binary or continuous
use the classical Baron & Kenny approach
estimate of the indirect as $\underline{a}\times\underline{b}$ and test via Sobel or Aroian

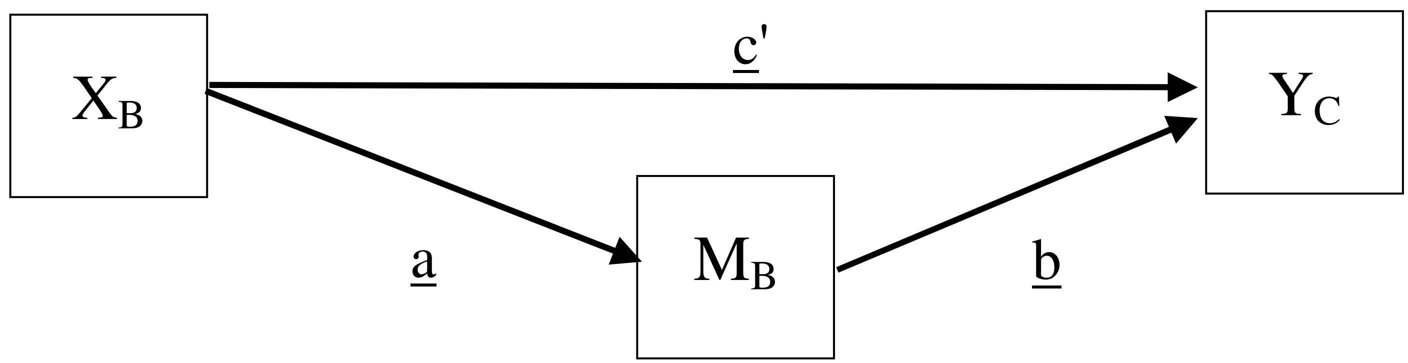
. Of course, you can bootstrap, as well

8 combinations of X, M, and Y

distribution of M	distribution of X	distribution of Y	
		continuous	binary
continuous	continuous or binary	Baron & Kenny (1986)	
binary	binary	Li et al (2007): Eq #14	
	continuous	Li et al (2007): Eq #13	

Next—a slight wrinkle: Binary X and M

Binary X and M with Continuous Y



. b equals the expected change in Y
caused by a one-unit change in binary M (from 0 to 1)

. a represents the expected change in the log odds that M=1 (versus 0)
caused by a one-unit change in binary X

. a×b represents…, hmm…

Binary X and M with Continuous Y

Simply taking the product a×b won't help—not interpretable

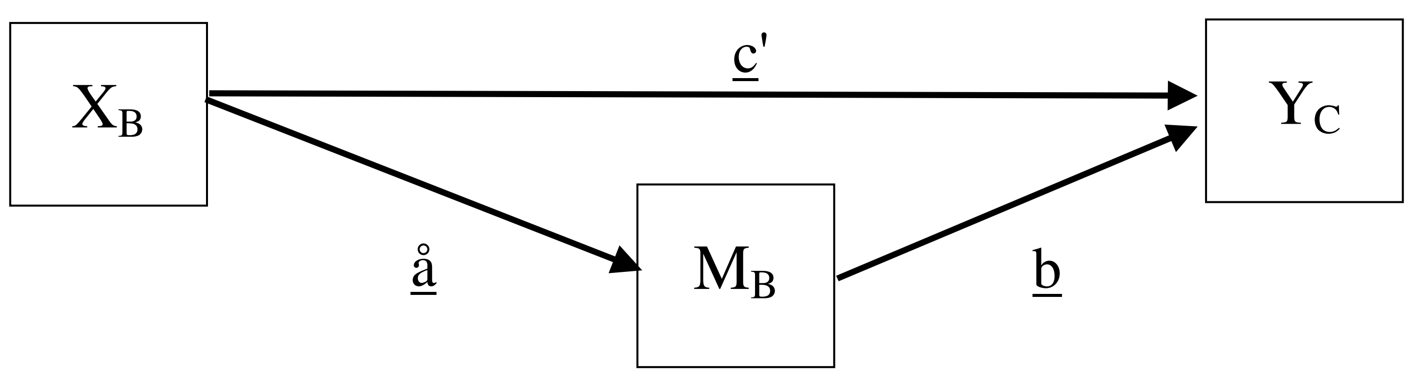
Huang et al (2004) and Li et al (2007) suggested that the effect of X on M be represented as the expected change in the probability that M=1 (versus 0), for a one unit increase in X.

Assuming no covariates…

$$\mathring{a} = \Pr(M=1|X=1) - \Pr(M=1|X=0)$$

$\mathring{a} \times b$ *is interpretable*

For continuous Y, it represents the expected change in Y, caused by a one-unit change in X channeled through M

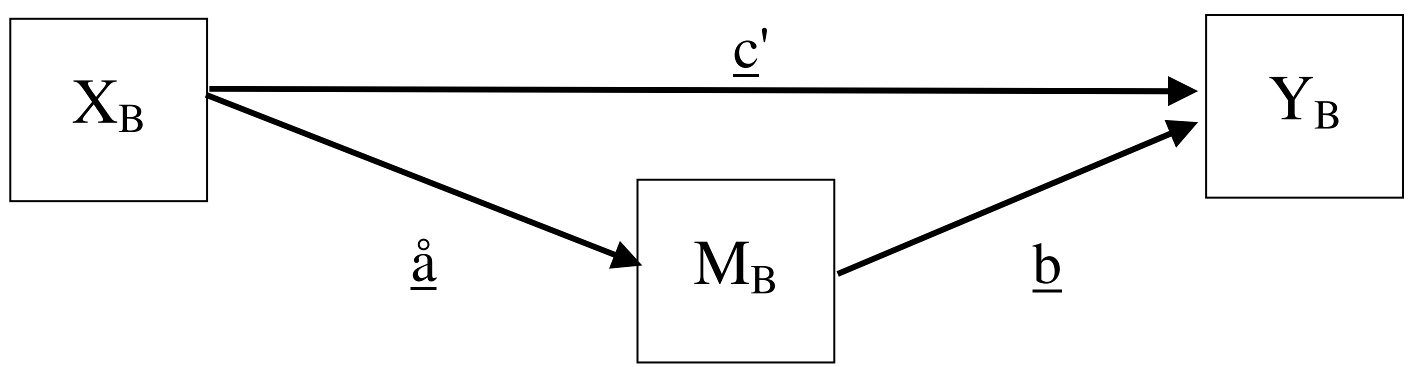


Binary X and M with Binary Y

. b equals the expected change in the log odds that Y=1 (versus 0), *caused* by a one-unit change in M

. å represents expected change in the probability that M=1 (versus 0), *caused* by a one unit increase in binary X

. å×b represents the expected change in the log odds that Y=1, caused by a one-unit change in X channeled through M



. Not (really) covered by Li et al (2007)

Binary X and M

Obtaining a standard error estimate for $\underline{å}\times\underline{b}$

. $\underline{å}$ can be regarded as a rescaling of $\underline{a}$

. The scaling factor, s, equals

$$s = \underline{å} \div \underline{a}$$

. It follows that

$$\underline{a} \times s = \underline{å}, \text{ and}$$

$$\sigma_{\underline{a}} \times s = \sigma_{\underline{å}}$$

similar to rescaling regression parameter and standard error estimates so they correspond to changes in X larger (or smaller) than one unit

Given estimates of $\underline{å}$, $\sigma_{\underline{å}}$, $\underline{b}$, and $\sigma_{\underline{b}}$,

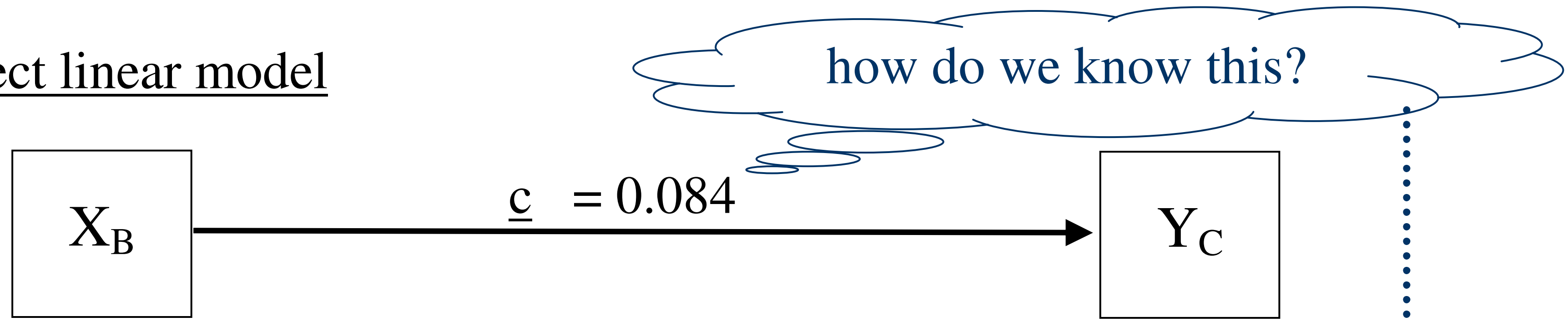
a standard error estimate of $\underline{å}\times\underline{b}$ can be obtained by Sobel, Aroian, etc.

. Not (really) covered by Li et al (2007)

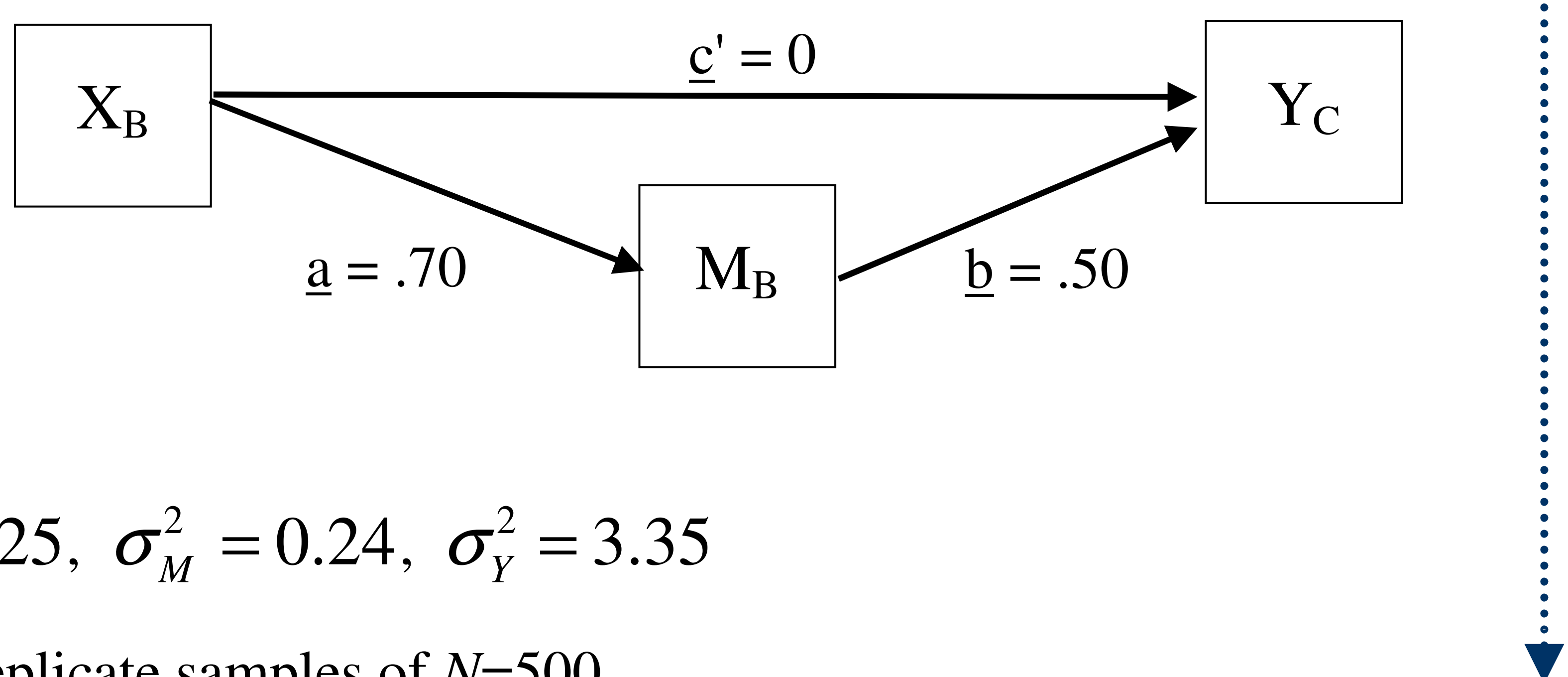
"One can also use the delta method to construct asymptotically valid confidence intervals."

Binary X and M
Simulation: Set-up with Continuous Y

Total effect linear model



3-variable linear and logistic path model

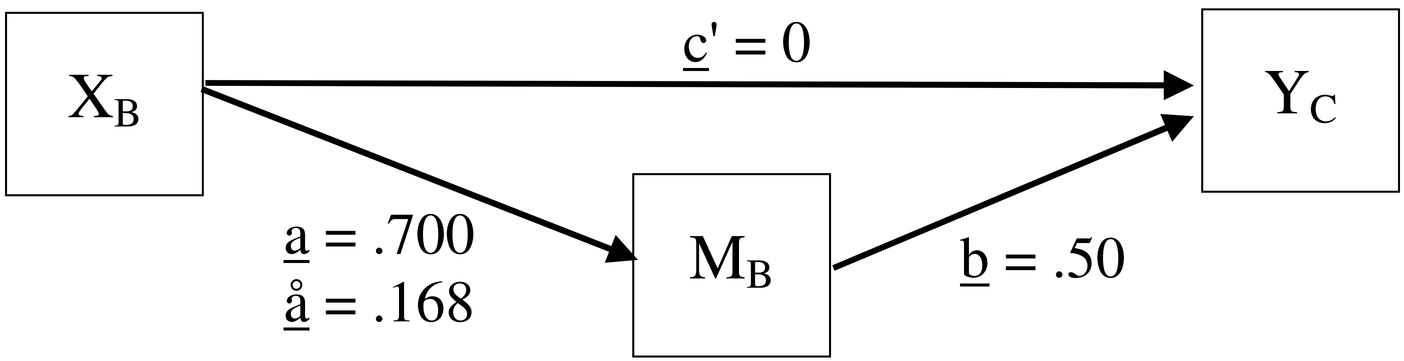


$$\sigma_X^2 = 0.25,\ \sigma_M^2 = 0.24,\ \sigma_Y^2 = 3.35$$

10,000 replicate samples of *N*=500

Binary X and M
Simulation: Set-up with Continuous Y

3-variable linear and logistic path model



All intercepts equal zero in the population, so the expected value of å equals

$$\mathring{a} = \frac{1}{\left(1+e^{-.70}\right)} - \frac{1}{\left(1+e^{0}\right)} = 0.168.$$

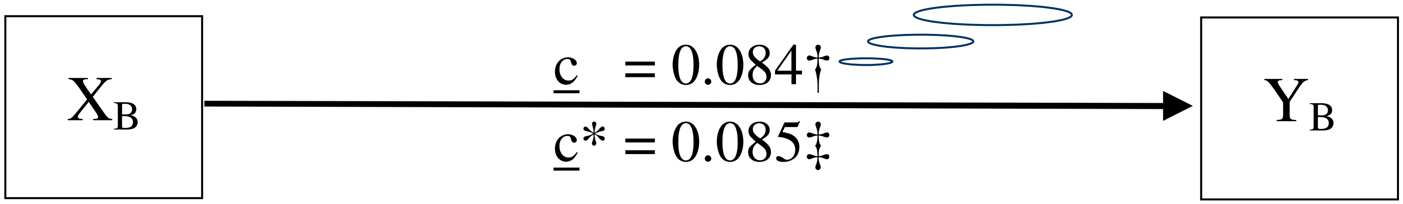
And,

å×b = 0.168×0.50 = 0.0841 = the indirect (and total) effect estimates because c' = 0

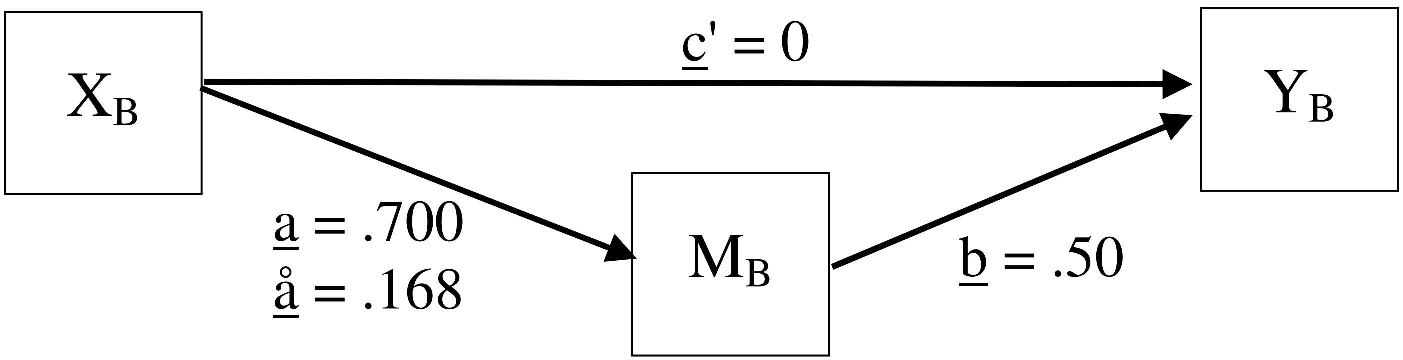
Binary X and M
Simulation: Set-up with Binary Y

why *don't* we know this?

Total effect logistic model



3-variable linear and logistic path model



$\sigma_X^2 = 0.25,\ \sigma_M^2 = 0.24,\ \sigma_Y^2 = 0.25$

10,000 replicate samples of N=500

Indirect effect = 0.168×0.5 = 0.0841

† estimated by simulation; ‡ rescaled by factor 1.009

Binary X and M
Simulation: Parameter estimates

Y		c total effect	a X→M	b M→Y	c' direct effect	å×b indirect effect	å×b+c' total effect
contin-uous	pop. value	.0841	.7000	.5000	0	.0841	.0841
	estimate	.0819	.7021	.4985	-.0019	.0839	.0820
binary	pop. value	.0842*	.7000	.5000	0	.0841	.0841
	estimate	.0845	.6990	.5021	.0003	.0840	.0843

Notes

* Population value of total effect estimated by simulation with 5M records

Binary X and M
Simulation: Standard errors

Y		c	a	b	c'	â×b		
		total effect	X→M	M→Y	direct effect	indirect effect		
						Sobel	Aroian	Goodman
contin-uous	pop. val.	.1650	.1843	.1675	.1663	.0366	.0366	.0366
	estimate	.1638	.1852	.1674	.1649	.0366	.0373	.0358
binary	pop. val.	.1828	.1856	.1870	.1878	.0392	.0392	.0392
	estimate	.1814	.1851	.1877	.1858	.0394	.0403	.0385

Notes

. Each standard error population value represents the standard deviation of the corresponding parameter (point) estimate across the 10,000 replicate samples

. Estimated standard errors represent the mean standard error estimate across the 10,000 replicate samples

Binary X and M
Simulation: Coverage

Y	c	a	b	c'	â×b		
	total effect	X→M	M→Y	direct effect	indirect effect		
					Sobel	Aroian	Goodman
continuous	.9485	.9496	.9513	.9480	.9329	.9402	.9264
binary	.9486	.9487	.9512	.9508	.9335	.9410	.9251

Binary X and M
Summary

distribution of M	distribution of X	distribution of Y	
		continuous	binary
continuous	continuous or binary	Baron & Kenny (1986)	
binary	binary	Li et al (2007): Eq #14	
	continuous	Li et al (2007): Eq #13	

Main points

. to use the a×b method with binary X and M replace a with å,
(å = change in predicted probability that M=1 given a one-unit increase in X)

. å can be considered a rescaling of a, so
a standard error estimate of å is readily available and
can be used to test the indirect effect

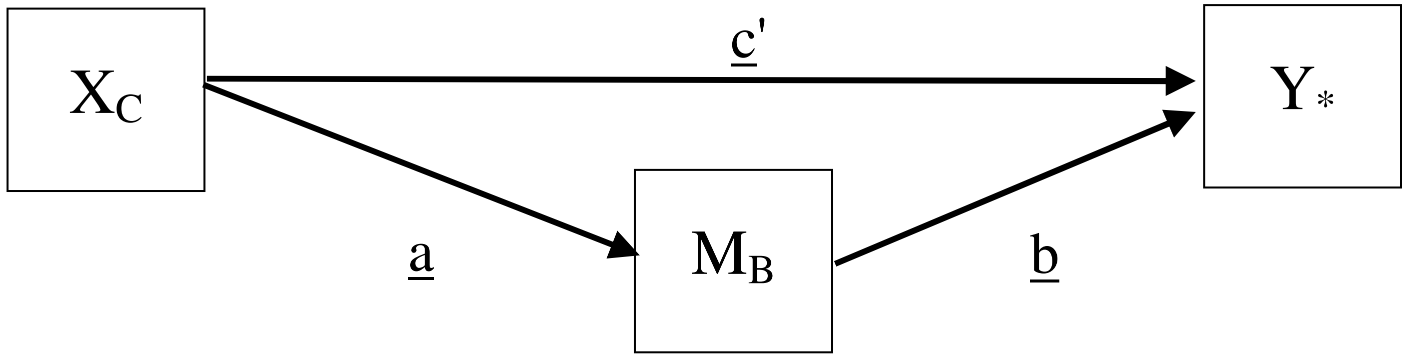
8 combinations of X, M, & Y

distribution of M	distribution of X	distribution of Y	
		continuous	binary
continuous	continuous or binary	Baron & Kenny (1986)	
binary	binary	Li et al (2007): Eq #14	
	continuous	Li et al (2007): Eq #13	

OK, so we've covered Binary X and M

Next—the wrinkle expands: Continuous X and Binary M

Continuous X and Binary M



. Binary Y: b equals the expected change in the logs odds that Y=1 (vs. 0) *caused* by a one-unit change in binary M (0 v 1)

or

. Continuous Y: b equals the expected change in Y *caused* by a one-unit change in binary M (0 v 1)

. a represents the expected change in the log odds that M=1 (vs. 0) *caused* by a one-unit change in continuous X

. a×b represents…, hmm…

We need a rescaling of a that reflects the change in predicted probability that M=1, given a one-unit increase in *continuous* X

Continuous X and Binary M

Rescaling $\underline{a}$

Li et al (2007) suggested a rescaling of $\underline{a}$, where

$\underline{\dot{a}} = \underline{a} \times \overline{s}$, and with no covariates the scaling factor, $\overline{s}$, is estimated as

$$\overline{s} = \frac{1}{N} \sum_{i=1}^{N} \frac{e^{(\alpha+\beta x_i)}}{\left(1 + e^{(\alpha+\beta x_i)}\right)^2}$$

With continuous X and binary M, the mediation effect is not constant across X—
The effect of X on the Pr(M=1) decreases as X approaches its extreme values

∴ estimate the scaling factor for each record in the data set and take the average

$\underline{\dot{a}}$ is analogous to $\underline{\mathring{a}}$
it represents the *average* change in predicted probability that M=1,
given a one-unit increase in X

Continuous X and Binary M

Obtaining a standard error estimate for $\underline{\dot{a}} \times \underline{b}$

. analogous to the case with $\underline{\mathring{a}}$, we can use the rescaling factor, $\overline{s}$, to obtain a standard error estimate for $\underline{\dot{a}}$

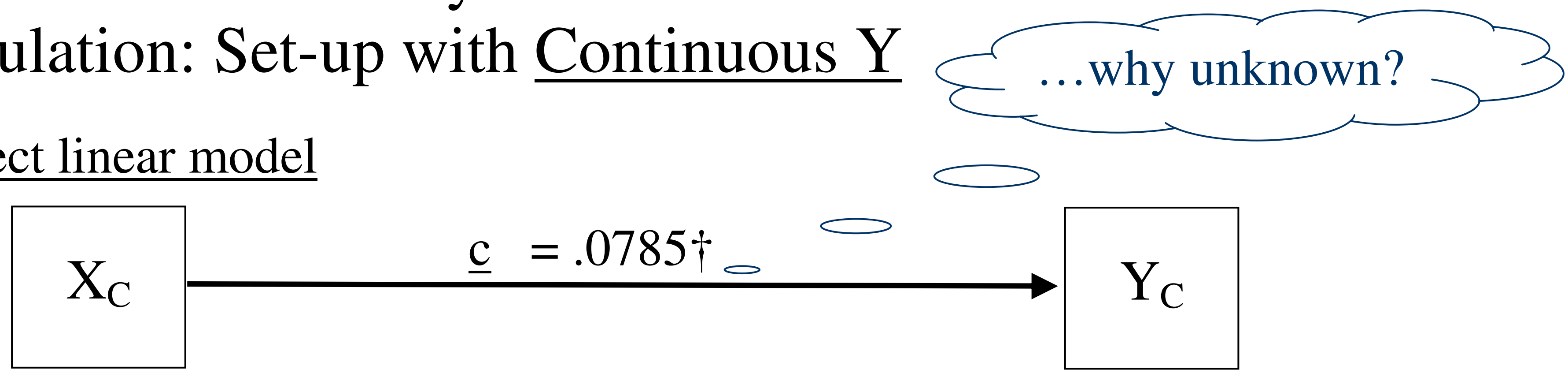
$$\sigma_{\underline{a}} \times \overline{s} = \sigma_{\underline{\dot{a}}}$$

Given estimates of $\underline{\dot{a}}$, $\sigma_{\underline{\dot{a}}}$, $\underline{b}$, and $\sigma_{\underline{b}}$,
a standard error estimate of $\underline{\dot{a}} \times \underline{b}$ can be obtained by Sobel, Aroian, etc.

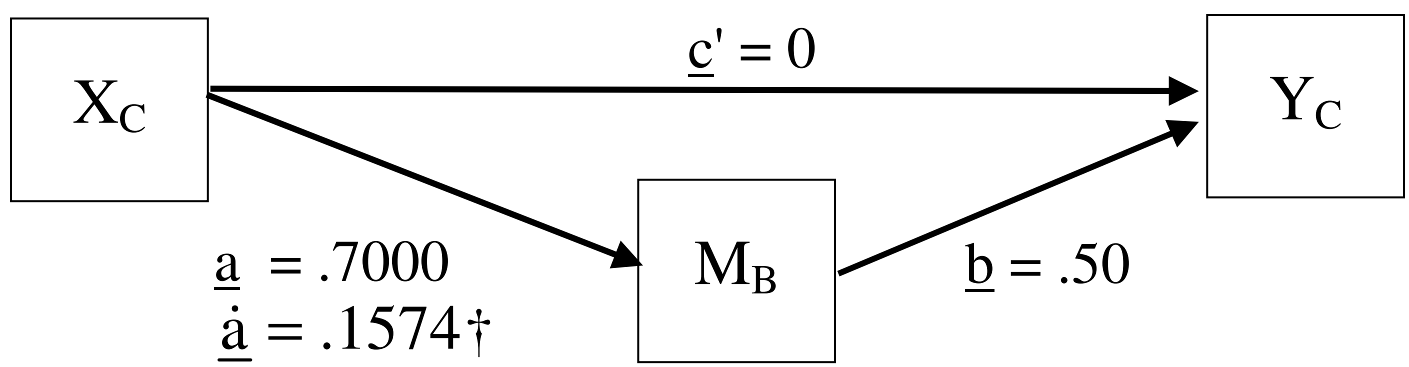
Not covered by Li et al (2007)

Continuous X and Binary M
Simulation: Set-up with Continuous Y

Total effect linear model



3-variable linear and logistic path model



$$\sigma_X^2 = 1.00,\ \sigma_M^2 = 0.25,\ \sigma_Y^2 = 3.36$$

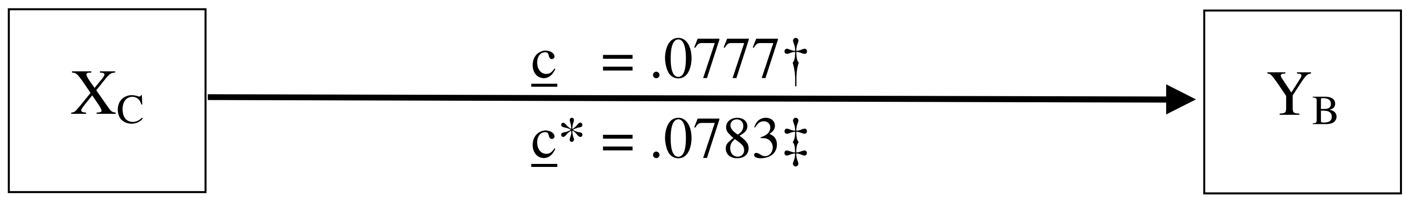
10,000 replicate samples of N=500

† estimated by simulation

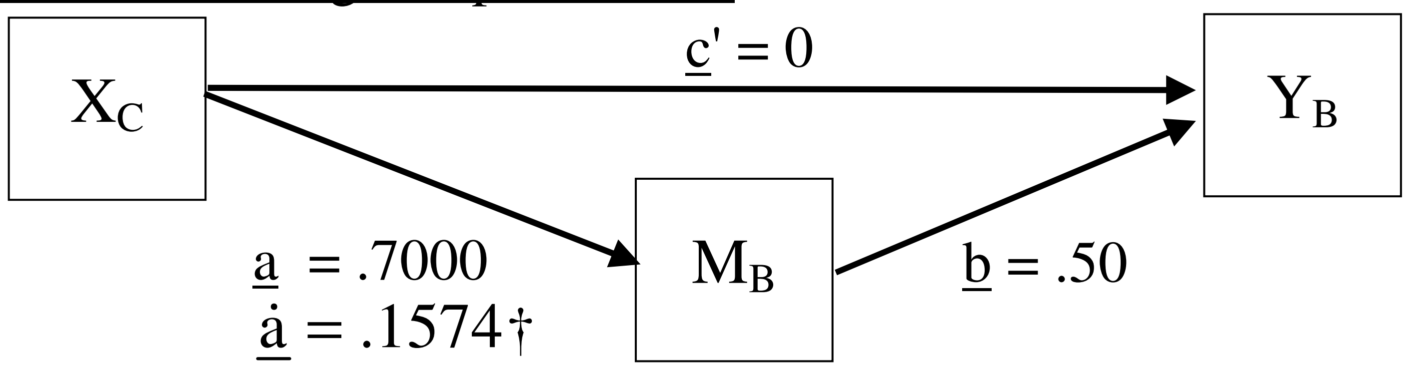
Given the simulated value for ȧ, the expected indirect effect $.1574 \times .50 = .0787$

Continuous X and Binary M
Simulation: Set-up with Binary Y

Total effect linear model



3-variable linear and logistic path model



$$\sigma_X^2 = 1.00,\ \sigma_M^2 = 0.25,\ \sigma_Y^2 = 0.25$$

10,000 replicate samples of *N*=500

† estimated by simulation; ‡ rescaled by factor 1.008

Given the simulated value for $\dot{\underline{a}}$, the expected indirect effect .1574 × .50 = .0787

Continuous X and Binary M
Simulation: Parameter estimates

Y		$\underline{c}$ total effect	$\underline{a}$ X→M	$\underline{b}$ M→Y	$\underline{c}'$ direct effect	$\underline{\dot{a}}\times\underline{b}$ indirect effect	$\underline{\dot{a}}\times\underline{b}+\underline{c}'$ total effect
contin-uous	pop. value	.0785*	.7000	.5000	0	.0787†	.0787†
	estimate	.0786	.7052	.4975	.0002	.0784	.0785
binary	pop. value	.0777*	.7000	.5000	0	.0787†	.0787†
	estimate	.0784	.7052	.5017	-.0005	.0790	.0785

* population value of total effect estimated by simulation with 5M records

† population value of $\underline{\dot{a}}$ estimated by simulation with 5M records

Continuous X and Binary M
Simulation: Standard errors

Y		c total effect	a X→M	b M→Y	c' direct effect	$\dot{a}\times b$ indirect effect Sobel	 Aroian	 Goodman
contin-uous	pop. val.	.0824	.1056	.1700	.0864	.0287	.0287	.0287
	estimate	.0820	.1050	.1713	.0857	.0297	.0300	.0294
binary	pop. val.	.0921	.1056	.1914	.0981	.0320	.0320	.0320
	estimate	.0910	.1050	.1925	.0965	.0329	.0332	.0325

Notes

. Population values of standard errors represent the standard deviation of the corresponding parameter (point) estimate across the 10,000 replicate samples

. Estimated standard errors represent the mean standard error estimate across the 10,000 replicate samples

Continuous X and Binary M
Simulation: Coverage

Y	$\underline{c}$ total effect	$\underline{a}$ X→M	$\underline{b}$ M→Y	$\underline{c}'$ direct effect	$\underline{\dot{a}}\times\underline{b}$* indirect effect Sobel	Aroian	Goodman
continuous	.9499	.9504	.9501	.9488	.9538	.9555	.9517
binary	.9465	.9504	.9536	.9467	.9578	.9602	.9544

* coverage of $\underline{\dot{a}}\times\underline{b}$ was around the simulated population value.
w/ continuous X and binary M, the population value was not known a priori

Continuous X, Binary M
Summary

distribution of M	distribution of X	distribution of Y: continuous	distribution of Y: binary
continuous	continuous or binary	Baron & Kenny (1986)	
binary	binary	Li et al (2007): Eq #14	
	continuous	Li et al (2007): Eq #13	

<u>Main points</u>

. With continuous X and Binary M,
the X→M effect varies as a function of X values.

. Therefore, we average across the indirect effects for each data record

. $\underline{\dot{a}}$ is analogous to $\underline{\mathring{a}}$
after rescaling the standard error for <u>a</u>, use the Sobel or Aroian method

Final wrinkles….

Conditioning on covariates: Binary X and M

Binary X without covariates

$$\underline{\mathring{a}} = \frac{1}{\left(1+e^{-(\beta_0+\underline{a})}\right)} - \frac{1}{\left(1+e^{-(\beta_0)}\right)},$$

change in predicted probability that M=1, given a one-unit change in binary X

Binary X with covariates

$$\underline{\mathring{a}} = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{1}{\left(1+e^{-(\beta_0+\underline{a}+\mathbf{\Gamma Z}_i)}\right)} - \frac{1}{\left(1+e^{-(\beta_0+\mathbf{\Gamma Z}_i)}\right)}\right],$$

where $\mathbf{Z}_i$ hold covariates and Γ is a corresponding vector of model parameters.

Final wrinkles….
Conditioning on covariates: Continuous X and Binary M

Continuous X without covariates

$$\overline{s} = \frac{1}{N}\sum_{i=1}^{N}\frac{e^{(\beta_0+\underline{a}x_i)}}{\left(1+e^{(\beta_0+\underline{a}x_i)}\right)^2}$$

Continuous X with covariates

$$\overline{s} = \frac{1}{N}\sum_{i=1}^{N}\frac{e^{(\beta_0+\underline{a}x_i+\mathbf{\Gamma Z}_i)}}{\left(1+e^{(\beta_0+\underline{a}x_i+\mathbf{\Gamma Z}_i)}\right)^2},$$

Either way, $\dot{\underline{a}} = \underline{a}\times\overline{s}$

Final wrinkles….
Multiple mediators

Obtain appropriate estimates of each indirect effect

The sum of the indirect effect estimates is the omnibus indirect effect estimate

Estimate the omnibus indirect effect standard error/confidence interval via bootstrap

Overall Summary

distribution of X	distribution of M	distribution of Y	
		continuous	binary
binary	continuous	Baron & Kenny (1986)	
continuous	continuous		
binary	binary	Li et al (2007)	
continuous	binary		

With continuous M

Calculate $\underline{a} \times \underline{b}$ directly, and

Apply the Sobel or Aroian test

With binary M

Calculate $\underline{\mathring{a}}$ and $\underline{\mathring{a}} \times \underline{b}$ (with binary X), or $\underline{\dot{a}}$ and $\underline{\dot{a}} \times \underline{b}$ (continuous X);

Rescale the standard error estimate for $\underline{a}$ (to correspond to $\underline{\mathring{a}}$ or $\underline{\dot{a}}$); and

Apply the Sobel or Aroian test

. You can save time by using the 2 *t*-test method as a screening tool (next….)

Alternative ways to test indirect effects
Performing the Sobel test of $\underline{a}\times\underline{b}$ via 2 *t*-tests

Given,

$$z_{Sobel} = \underline{a}\times\underline{b} \Big/ \sqrt{\underline{a}^2\sigma_{\underline{b}}^2 + \underline{b}^2\sigma_{\underline{a}}^2}\,,$$

and the standard t-test formula

$$t_{\underline{a}} = \underline{a} \Big/ \sqrt{\sigma_{\underline{a}}^2}\,, \text{ and}$$

$$t_{\text{b}} = \underline{b} \Big/ \sqrt{\sigma_{\underline{b}}^2}\,,$$

it follows that

$$z_{Sobel} = 1 \Big/ \sqrt{1/t_{\underline{a}}^2 + 1/t_{\underline{b}}^2}$$

Alternative ways to test indirect effects

Performing the Aroian & Goodman tests of $\underline{a}\times\underline{b}$ via 2 t-tests

Similarly,

$$z_{Aroian} = \underline{a}\times\underline{b} \Big/ \sqrt{\underline{a}^2\sigma_{\underline{b}}^2 + \underline{b}^2\sigma_{\underline{a}}^2 + \sigma_{\underline{a}}^2\sigma_{\underline{b}}^2}$$

$$= 1 \Big/ \sqrt{1/t_{\underline{a}}^2 + 1/t_{\underline{b}}^2 + 1/\left(t_{\underline{a}}^2 t_{\underline{b}}^2\right)}$$

$$z_{Goodman} = \underline{a}\times\underline{b} \Big/ \sqrt{\underline{a}^2\sigma_{\underline{b}}^2 + \underline{b}^2\sigma_{\underline{a}}^2 - \sigma_{\underline{a}}^2\sigma_{\underline{b}}^2}$$

$$= 1 \Big/ \sqrt{1/t_{\underline{a}}^2 + 1/t_{\underline{b}}^2 - 1/\left(t_{\underline{a}}^2 t_{\underline{b}}^2\right)}$$

Thank you, Kris Preacher!

Alternative ways to test indirect effects: Extensions

Performing tests of $\underline{a}\times\underline{b}$ via 2 χ^2 tests

. With 1 df, $t^2 = \chi^2$

$$z_{Sobel} = 1 \Big/ \sqrt{1/\chi^2_{\underline{a}} + 1/\chi^2_{\underline{b}}}$$

$$z_{Aroian} = 1 \Big/ \sqrt{1/\chi^2_{\underline{a}} + 1/\chi^2_{\underline{b}} + 1/\left(\chi^2_{\underline{a}}\chi^2_{\underline{b}}\right)}$$

$$z_{Goodman} = 1 \Big/ \sqrt{1/\chi^2_{\underline{a}} + 1/\chi^2_{\underline{b}} - 1/\left(\chi^2_{\underline{a}}\chi^2_{\underline{b}}\right)}$$

no big deal, just for completeness

Alternative ways to test indirect effects: <u>Extensions</u>
Performing the Sobel test of $\underline{a}\times\underline{b}$ via 2 *t*-tests

Given a value for $t_{\bullet}$ solve for $t_{\bullet\bullet}$ that will yield a critical *z*-value for $\underline{a}\times\underline{b}$, $cv_{\underline{a}\times\underline{b}}$

First, define the magnitude of $t_{\bullet\bullet}$ as a factor of $t_{\bullet}$, such that $t_{\bullet\bullet} = \gamma t_{\bullet}$

Then, use the Sobel 2 *t*-test method and solve for γ.

$$cv_{\underline{a}\times\underline{b}} = \frac{1}{\sqrt{\frac{1}{t_{\bullet}^2} + \frac{1}{\gamma^2 t_{\bullet}^2}}}$$

rearranging, we get

$$\gamma = \sqrt{\frac{cv_{\underline{a}\times\underline{b}}^2}{t_{\bullet}^2 - cv_{\underline{a}\times\underline{b}}^2}}$$

Alternative way to test indirect effects: Extensions

Testing $\underline{a}\times\underline{b}$ via 2 t-tests

$$\gamma = \sqrt{\frac{cv_{\underline{a}\times\underline{b}}^{2}}{t_{\bullet}^{2} - cv_{\underline{a}\times\underline{b}}^{2}}}, \quad \text{and} \quad t_{\bullet\bullet} = \gamma t_{\bullet}$$

Implications

. If $t_{\bullet} \le cv_{\underline{a}\times\underline{b}}$, then γ is undefined—

So, if $t_{\bullet}$ is not significant or *just* significant, then $\underline{a}\times\underline{b}$ is not significant

. As $t_{\bullet} \to cv_{\underline{a}\times\underline{b}}$ from above, $t_{\bullet\bullet} \to \infty$

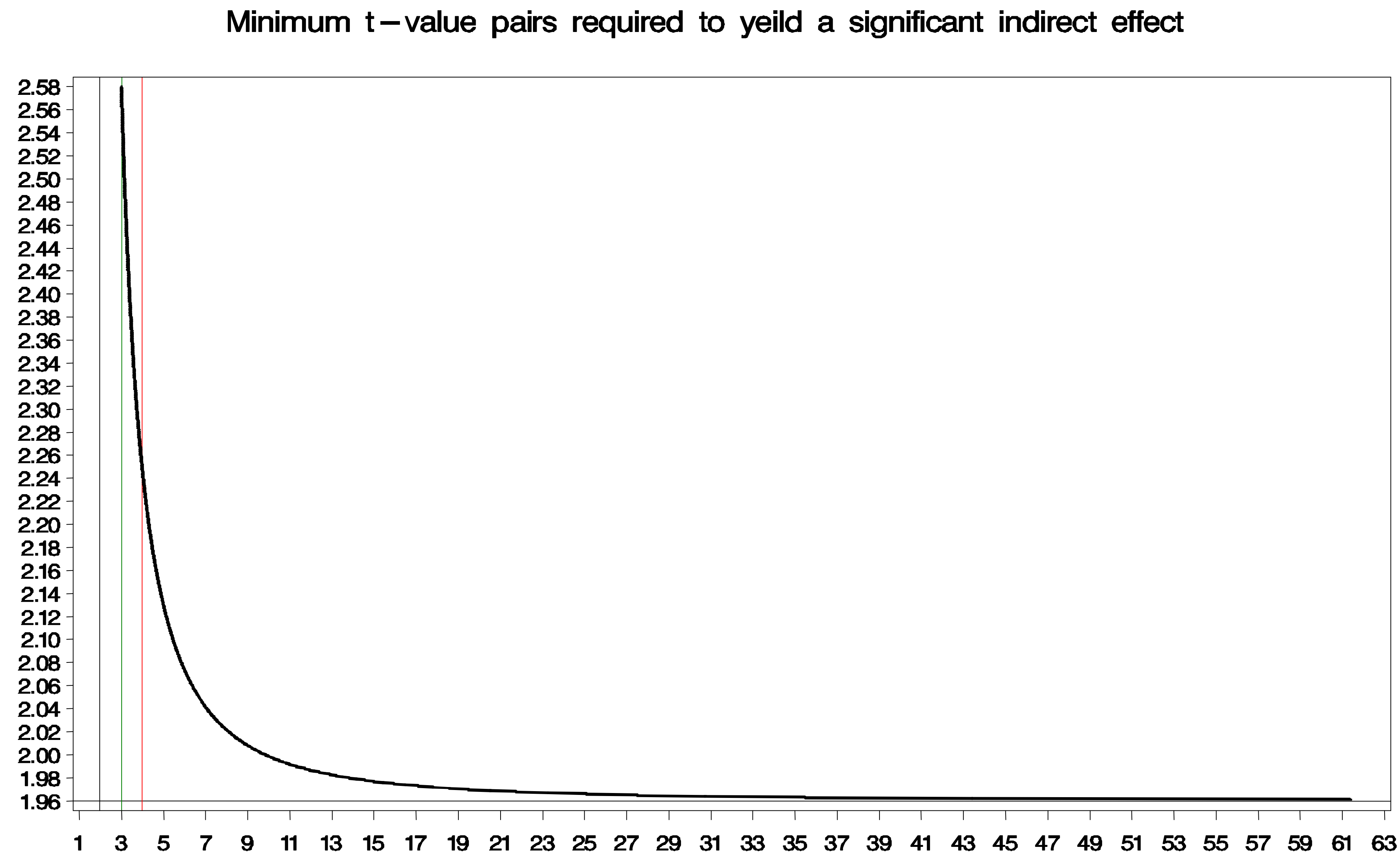
. When $t_{\bullet}^{2} = 2cv_{\underline{a}\times\underline{b}}^{2}$, then γ=1 and $t_{\bullet} = t_{\bullet\bullet}$: the 'balanced' case

If $cv_{\underline{a}\times\underline{b}} = 1.96$, then $t_{\bullet} = t_{\bullet\bullet} = 2.772$ (for each t, $p \approx .0055$)

. Crazy rules: e.g., if $t_{\underline{a}} + t_{\underline{b}} < 5.54$, then the indirect effect must be n.s.

Alternative way to test indirect effects: Extensions

Testing a×b via 2 *t*-tests



Alternative way to test indirect effects: Extensions

Testing $\underline{a} \times \underline{b}$ via 2 p-values

. Compute the absolute t-values from the corresponding p-values (assuming a two-tailed test)

$$\mathrm{ABS}\left(t_{\underline{a}}\right) = -\Phi\left(p_{\underline{a}}/2\right), \text{ and}$$

$$\mathrm{ABS}\left(t_{\underline{b}}\right) = -\Phi\left(p_{\underline{b}}/2\right),$$

where $\Phi(.)$ represents the probit function

. Then, apply the Sobel, Aroian, or Goodman 2 t-test method

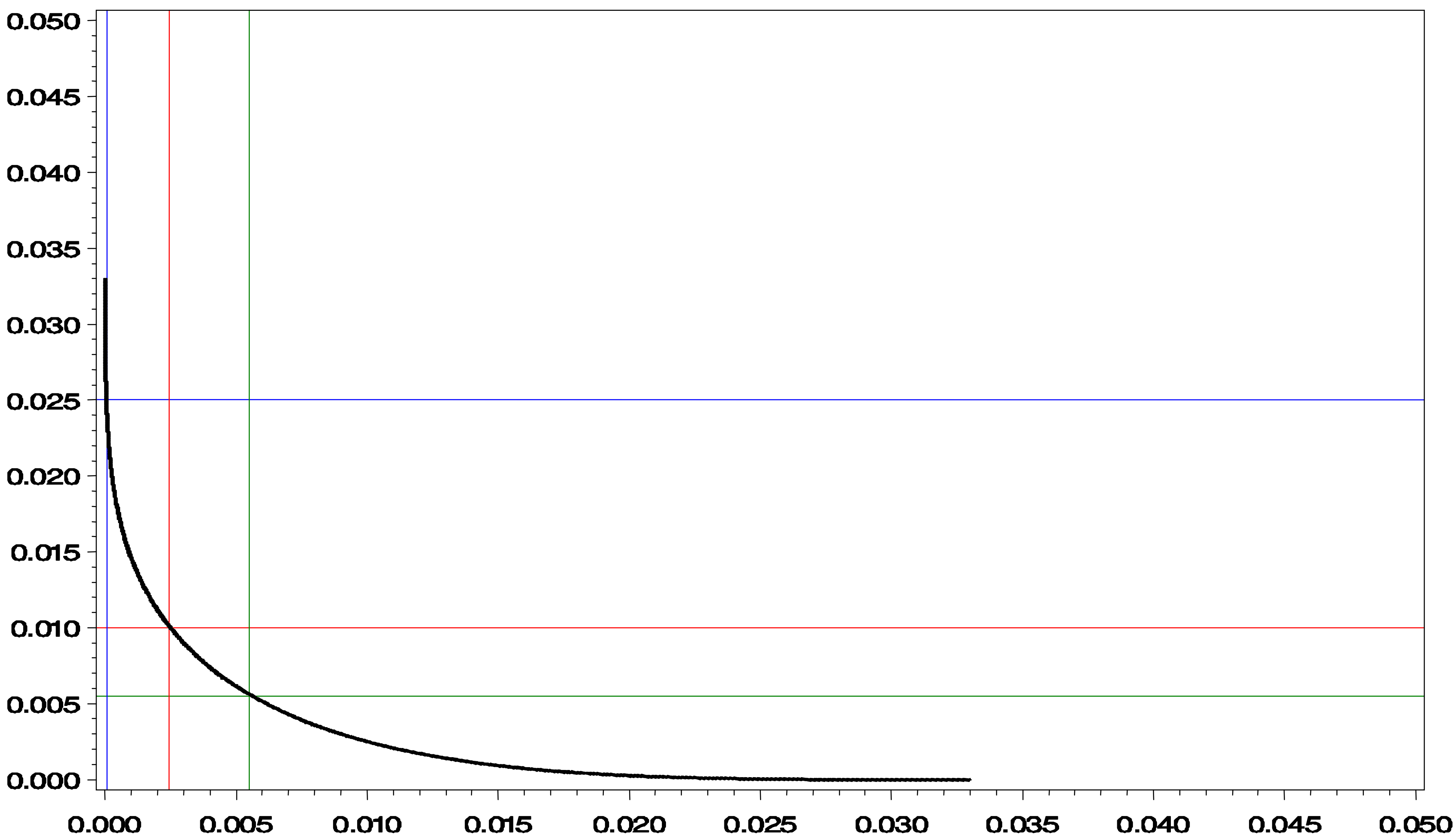
Minimum conditions for significant indirect effect ($p < .05$)

. balanced case: two p-values ≈ .0055 (t-values = 2.772)
. if one p=.010 (t=2.58), then the other must be <.0025 (t=3.01)
. if one p=.025 (t=2.25), then the other must be <.0001 (t=3.99)
. if one p=.030 (t=2.17), then the other must be <.000005 (t=4.57)
. if one p=.040 (t=2.05), then the other must be <.00000000005 (t=6.56)
. if one p=.045 (t=2.01), then the other must → 0 (t=9.34)

Alternative way to test indirect effects: Extensions

Performing the Sobel test of a×b via 2 *p*-values

Minimum p−value pairs required to yeild a significant indirect effect



Alternative way to test indirect effects
Summary

The 2 *t*-test (2 *p*-value) methods and corresponding a×b Sobel, Aroian, and Goodman tests will give equivalent results (within rounding error)

The 2 *t*-test (2 *p*-value) method allows a quick test of the indirect effect, even in the absence of an indirect effect point estimate

You can use the 2 *t*-value or 2 *p*-value tests with any combination of fixed effect models

linear, logistic, Poisson, neg-binomial, gamma, you name it

Whenever you have continuous Y, but binary M,
you can estimate the indirect effect as c – c'
and test the indirect effect with the 2-*t*-test method;
no need to estimate å or ȧ

References

Baron and Kenny (1986). The moderator-mediator distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173-82.

Bin, Sivaganesan, Succop, and Goodman (2004). Statistical assessment of mediational effects for logistic mediational models. *Statist. Med.*, 23, 2713–2728

Li, Schneider and Bennett (2007). Estimation of the mediation effect with a binary mediator. *Statist. Med.*, 26, 3398–3414.

Mackinnon and Dwyer (1993). Estimating Mediated Effects in Prevention Studies. *Evaluation Review*, 17, 144-158.

SAS Code

```
*------------------------------------------------------------------------------------*;
* CONTINUOUS X AND BINARY M;
*------------------------------------------------------------------------------------*;
title 'Continuous X and Binary M';

data x;                           * GENERATE DATA FROM LOGISTIC MODEL;
do j = 1 to 5000;                 * N=5000;
  x   = rannor(1234567);          * GENERATE CONTINUOUS X;
  cov = rannor(1234567);          * GENERATE CONTINUOUS COVARIATE;
  e = ranuni(1234566);            * GENERATE STANDARD LOGISTIC RESIDUAL;
  e = log(e / (1-e));             * GENERATE STANDARD LOGISTIC RESIDUAL;
  y_star = 1 + .5*x + .2*cov + e; * GENERATE Y-STAR;
  if y_star >0 then y=1;          * GENERATE BINARY Y;
  else              y=0;          * GENERATE BINARY Y;
  _x_=1;                          * USED FOR MERGE LATER;
  output;
end;
run;
*;

proc logistic descending data=x;  * FIT X-->M MODEL;
  model y = x cov;
  output out=o xbeta=lin_pred;
  ods output parameterestimates=a_ContX;
run;
*;

data o;                           * CALCULATE S_BAR FOR EACH DATA RECORD;
 set o;
s_bar = exp(lin_pred) / (1+exp(lin_pred))**2;

proc means data=o noprint;        * AVERAGE S_BAR ACROSS DATA RECORDS;
 var s_bar;
 output out=s_bar mean=s_bar;
run;

data a_ContX;                     * SAVE THE POINT AND STANDARD ERROR ESTIMATES;
 set a_ContX;
if variable='x';
keep variable estimate stderr;
rename estimate=B_a stderr=B_a_se;

data a_dot;                       * ESTIMATE A_DOT AND ITS STANDARD ERROR;
 merge a_ContX s_bar;
B_a_dot    = B_a*s_bar;
B_a_dot_se = B_a_se*s_bar;
drop _type_ _freq_ variable;

proc print;
run;
*-----------------------------------------------------------------------------*;
* END: CONTINUOUS X AND BINARY M;
*-----------------------------------------------------------------------------*;
```

```
*---------------------------------------------------------------------------*;
* BINARY X AND BINARY M;
*---------------------------------------------------------------------------*;
title 'Binary X and M';

data x;                              * GENERATE DATA FROM LOGISTIC MODEL;
do j = 1 to 5000;                    * N=5000;
  x   = rannor(1234567);             * GENERATE CONTINUOUS X;
  if x>0 then x_bin=1;               * GENERATE BINARY X;
  else        x_bin=0;               * GENERATE BINARY X;
  cov = rannor(1234567);             * GENERATE CONTINUOUS COVARIATE;
  e = ranuni(1234566);               * GENERATE STANDARD LOGISTIC RESIDUAL;
  e = log(e / (1-e));                * GENERATE STANDARD LOGISTIC RESIDUAL;
  y_star = 1 + .5*x_bin + .2*cov + e; * GENERATE Y-STAR;
  if y_star >0 then y=1;             * GENERATE BINARY Y;
  else              y=0;             * GENERATE BINARY Y;
  _x_=1;                             * USED FOR MERGE LATER;
  output;
end;
run;
*;

proc logistic descending data=x;      * FIT X-->M MODEL;
  model y = x_bin cov;
  output out=o xbeta=lin_pred;
  ods output parameterestimates=a_BinX; * OUTPUT THE LINEAR PREDICTOR;
run;
*;

data a_2;                             * SAVE THE POINT AND STANDARD ERROR ESTIMATES;
 set a_BinX;
_x_=1;
if variable='x_bin';
keep variable estimate stderr _x_;
rename estimate=B_a stderr=B_a_se;

data o;                               * CALCULATE RECORD-SPECIFIC A_CIRCLE VALUES;
 merge o a_2;
 by _x_;
if x_bin=0 then lin_pred_int_cov = lin_pred;       ** CALC THE LINEAR PREDICTOR FOR INTERCEPT AND COVARIATES ONLY;
if x_bin=1 then lin_pred_int_cov = lin_pred - B_a; ** CALC THE LINEAR PREDICTOR FOR INTERCEPT AND COVARIATES ONLY;
B_a_circle = (1 / (1+exp(-(lin_pred_int_cov+B_a)))) - (1 / (1+exp(-lin_pred_int_cov)));

proc means data=o noprint;           * AVERAGE A_CIRCLE VALUES;
 var                     B_a B_a_se B_a_circle;
 output out=B_a_circle mean=B_a B_a_se B_a_circle;
run;

data B_a_circle;                     * CALCULATE S AND STANDARD ERROR ESTIMATE FOR A_CIRCLE;
 set B_a_circle;
s_nobar       = B_a_circle/B_a;
B_a_circle_se = B_a_se*s_nobar;
drop _type_ _freq_;

proc print;
 var B_a B_a_se s_nobar B_a_circle B_a_circle_se;
run;
```