Chapter 822

Tests of Mediation Effect using the Sobel Test

Introduction

This procedure computes power and sample size for a mediation analysis of a continuous dependent (output) variable Y and an independent (input) variable X. Interest focuses on the interrelationship between Y, X, and a third variable called the mediator M. The sample size calculations are based on the work of Sobel (1982).

Mediation Model

An in-depth discussion of mediation can be found in Hayes (2018). A popular method for testing for mediation is that of Baron and Kenny (1986). In this method, three regression models are fit where $M \sim \mathrm{N}(\mu_M, \sigma_M^2)$ and $X \sim \mathrm{N}(\mu_x, \sigma_X^2)$.

(1) $M = \theta_0 + \theta_X X + e_M,\ e_M \sim \mathrm{N}\left(0, \sigma_{e_M}^2\right)$

(2) $Y = \beta_0^* + \beta_X^* X + e_{Y^*}$

(3) $Y = \beta_0 + \beta_X X + \beta_M M + e_Y,\ e_Y \sim \mathrm{N}\left(0, \sigma_{e_Y}^2\right)$

These coefficients have the following interpretations: β_X is called the *direct effect* of X on Y, $\theta_X \beta_M$ is called the *indirect effect* of X on Y, and β_X^* is called the *total effect* of X on Y.

The indirect effect comes from substituting the equation for M in (1) into equation (3) and then rearranging terms.

$$\begin{aligned} Y &= \beta_0 + \beta_X X + \beta_M (M) + e_Y \\ &= \beta_0 + \beta_X X + \beta_M (\theta_0 + \theta_X X + e_M) + e_Y \\ &= \beta_0 + \beta_X X + \beta_M \theta_0 + \beta_M \theta_X X + \beta_M e_M + e_Y \\ &= (\beta_0 + \beta_M \theta_0) + (\beta_X + \boldsymbol{\beta_M \theta_X}) X + (\beta_M e_M + e_Y) \end{aligned}$$

Comparing coefficients, we can see that the *total effect* is exactly equal to the *direct effect* plus the *indirect effect*.

Mediation is likely if all four of the following tests are significant:

1. Test of θ_X.
2. Test of β_X^*.
3. Test of β_M.
4. Sober's (1982) Test of whether β_X is significantly smaller than β_X^* using

   $z = \theta_X \beta_M / \sqrt{\theta_X^2 \mathrm{V}(\beta_M) + \beta_M^2 \mathrm{V}(\theta_X)}$.

   where $\mathrm{V}(\theta_X) = \frac{\sigma_{e_M}^2}{N \sigma_X^2}, \mathrm{V}(\beta_M) = \frac{\sigma_{e_Y}^2}{N \sigma_{e_M}^2},\ \rho_{XM} = \frac{\theta_X \sigma_X}{\sigma_M}$, and $\sigma_{e_M}^2 = \sigma_M^2 (1 - \rho_{XM}^2)$

This procedure provides a power analysis and sample size calculation of the z test in step 4.

Calculating the Power

Power calculations are based on standard normal distribution. They proceed as follows:

1. Determine the critical value $z_{1-\alpha}$ from the standard normal distribution where α is the probability of a type-I error.
2. Calculate: $z_\beta = \frac{\theta_X \beta_M}{\sqrt{\theta_X^2 \mathrm{V}(\beta_M) + \beta_M^2 \mathrm{V}(\theta_X)}} - z_{1-\alpha}$.
3. Calculate: Power $= \Phi(z_\beta)$.

Notes

1. Use $\frac{\alpha}{2}$ instead of α for two-sided test.
2. $\sigma_M^2 = \Pr(M = 1)\Pr(M = 0)$ if M is binary.
3. $\sigma_X^2 = \Pr(X = 1)\Pr(X = 0)$ if X is binary.

Procedure Options

This section describes the options that are specific to this procedure. These are located on the Design tab. For more information about the options of other tabs, go to the Procedure Window chapter.

Design Tab

The Design tab contains most of the parameters and options that you will be concerned with.

Solve For

Solve For

This option specifies the parameter to be solved for from the other parameters. Under most situations, you will select either *Power* or *N* (*Sample Size)*.

Select *Sample Size* when you want to calculate the sample size needed to achieve a given power and alpha level.

Select *Power* when you want to calculate the power of an experiment.

Test Direction

Alternative Hypothesis

Specify whether the hypothesis test is one-sided or two-sided. When a two-sided test is selected, the value of alpha is automatically divided by two. A two-sided test requires alpha to be less than 0.50.

Power and Alpha

Power

This option specifies one or more values for power. Power is the probability of rejecting a false null hypothesis and is equal to one minus Beta. Beta is the probability of a type-II error, which occurs when a false null hypothesis is not rejected.

Values must be between zero and one. Historically, the value of 0.80 (Beta = 0.20) was used for power. Now, 0.90 (Beta = 0.10) is also commonly used. A single value may be entered here or a range of values such as *0.8 to 0.95 by 0.05* may be entered.

Alpha

This option specifies one or more values for the probability of a type-I error (alpha). A type-I error occurs when you reject the null hypothesis when in fact it is true.

Values of alpha must be between zero and one. Historically, the value of 0.05 has been used for alpha. This means that about one test in twenty will falsely reject the null hypothesis. You should pick a value for alpha that represents the risk of a type-I error you are willing to take in your experimental situation.

You may enter a range of values such as *0.01 0.05 0.10* or *0.01 to 0.10 by 0.01*.

Sample Size

N (Sample Size)

This option specifies the value(s) for *N*, the sample size. Note that $3 < N$.

Effect Size

θ_X (Reg Coef of X)

Enter one or more values for θ_X, which is the coefficient of X in the regression model.

$M = \theta_0 + \theta_X(X) + e_M$

where X is the primary predictor variable and M is the mediator variable. The product of θ_X and β_M is the quantity being tested. This product gives the *indirect effect* of X on Y. Hence, the effect size is $\theta_X(\beta_M)$ divided by its standard error.

Range

θ_X can be any value other than zero. It is in the same scale as σ_X.

β_M (Reg Coef of M)

Enter one or more values for β_M, which is the coefficient of M in the regression model.

$Y = \beta_0 + \beta_X(X) + \beta_M(M) + e_Y$

where X is the primary predictor variable and M is the mediator variable.

The product of θ_X and β_M is the quantity being tested. This product gives the *indirect effect* of X on Y. Hence, the effect size is $\theta_X(\beta_M)$ divided by its standard error.

Range

β_M can be any value other than zero. It is in the same scale as σ_M.

Type of Primary Predictor, X

Indicate whether the primary predictor variable, X, is continuous or binary.

σX (Standard Deviation of X)

Enter one or more values for the standard deviation of X, the primary predictor variable.

Range

$0 < \sigma_X$.

Probability X = 1

Enter one or more values of the probability in the population that X = 1 when X is a binary variable that takes on the values 0 or 1. This value is used to compute the variance of X using the formula: $\sigma_X^2 = P(X=1) \times (1 - P(X=1))$.

Range

Since this is a probability, it can be any value between 0 and 1.

Type of Mediator, M

Indicate whether the mediator variable, M, is continuous or binary.

σM (Standard Deviation of X)

Enter one or more values for the standard deviation of X, the primary predictor variable.

Range

$0 < \sigma_M$.

Probability M = 1

Enter one or more values of the probability in the population that M = 1 when M is a binary variable that takes on the values 0 or 1. This value is used to compute the variance of M using the formula: $\sigma_M^2 = P(M=1) \times (1 - P(M=1))$.

Range

Since this is a probability, it can be any value between 0 and 1.

σe (Standard Deviation of eY)

Enter one or more values for the standard deviation of e_Y from the model

$Y = \beta 0 + \beta_X(X) + \beta_M(M) + e_Y$

Range

$0 < \sigma e$.

Example 1 – Finding Sample Size

Researchers are studying the relationship between a dependent variable (Y) and an independent variable (X). They want to understand the impact of a third variable (M) on the relationship between X and Y, so they decide to carry out a mediation analysis. They decide to use Sobel's test for their sample calculation. Using prior analyses, they decide to use θ_X = 0.2, 0.3, 0.5; β_M = 0.2, 0.4, 0.6, σ_X = 0.6, σ_M = 0.5, and σ_e = 0.2. They set the power at 0.9 and the two-sided significance level at 0.05.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Tests of Mediation Effect using the Sobel Test** procedure. You may then make the appropriate entries as listed below, or open **Example 1** by going to the **File** menu and choosing **Open Example Template**.

Option	Value
Design Tab	
Solve For	**N (Sample Size)**
Alternative Hypothesis	**Two-Sided**
Power	**0.90**
Alpha	**0.05**
θ_X (Reg Coef of X)	**0.2 0.3 0.5**
β_M (Reg Coef of M)	**0.2 0.4 0.6**
Type of Primary Predictor, X	**Continuous**
σ_X (Standard Deviation of X)	**0.6**
Type of Mediator, M	**Continuous**
σ_M (Standard Deviation of M)	**0.5**
σ_e (Standard Deviation of e_Y)	**0.2**

Annotated Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Results

Numeric Results
Two-sided alternative hypothesis

Power	Sample Size N	Reg Coef of X θ_X	Reg Coef of M β_M	Coef Product $\theta_X(\beta_M)$	Std Dev of X σ_X	Std Dev of M σ_M	Std Dev of e_Y σ_e	Std Dev of Coef Product $\sigma(\theta_X(\beta_M))$	Alpha
0.9006	217	0.200	0.200	0.040	0.600	0.500	0.200	0.012	0.050
0.9014	184	0.200	0.400	0.080	0.600	0.500	0.200	0.025	0.050
0.9002	177	0.200	0.600	0.120	0.600	0.500	0.200	0.037	0.050
0.9003	119	0.300	0.200	0.060	0.600	0.500	0.200	0.018	0.050
0.9012	83	0.300	0.400	0.120	0.600	0.500	0.200	0.037	0.050
0.9002	76	0.300	0.600	0.180	0.600	0.500	0.200	0.056	0.050
0.9022	85	0.500	0.200	0.100	0.600	0.500	0.200	0.031	0.050
0.9071	36	0.500	0.400	0.200	0.600	0.500	0.200	0.061	0.050
0.9003	26	0.500	0.600	0.300	0.600	0.500	0.200	0.093	0.050

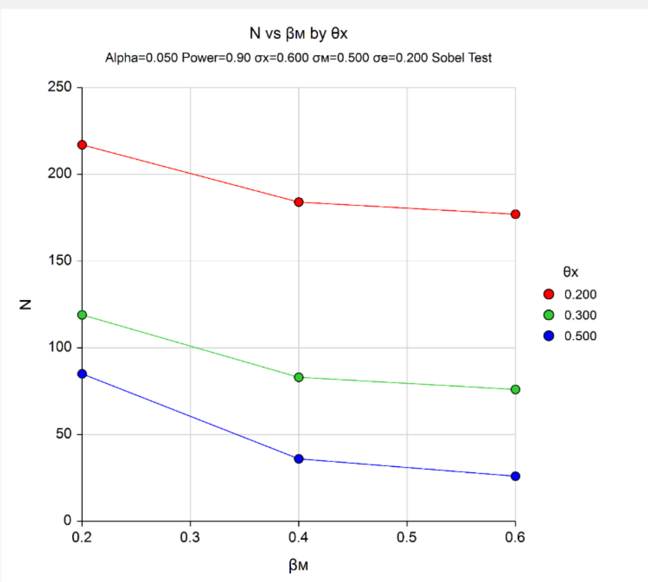
References
Fritz, M.S. and MacKinnon, D.P. 2007. 'Required Sample Size to Detect the Mediated Effect.' Psychological Science, Vol. 18, No. 3, Pages 233-239.
Hayes, A.F. 2018. Introduction to Mediation, Moderation, and Conditional Process Analysis, Second Edition. CRC The Guilford Press. New York.
Sobel, M. E. 1982. 'Asymptotic confidence intervals for indirect effects in structural equation models.' Sociological Methodology. Vol. 13, Pages 290-312.
Vittinghoff, E., Sen, S., and McCulloch, C.E. 2009. 'Sample size calculations for evaluating mediation.' Statistics in Medicine, Vol. 28, Pages 541-557.

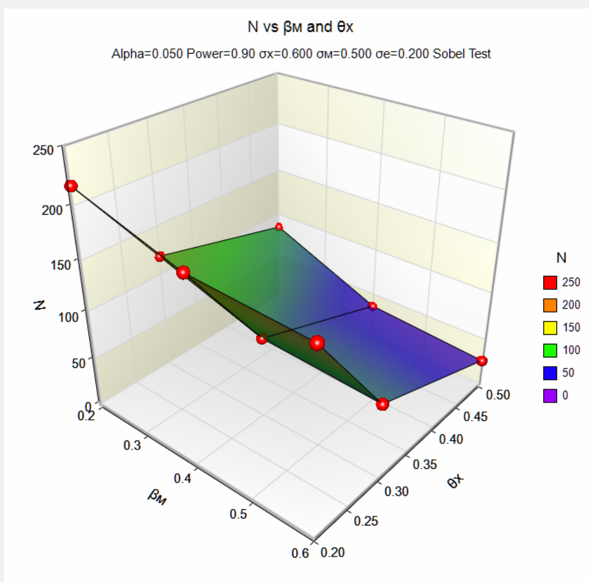
Report Definitions
Hypotheses: H0: $\theta_X(\beta_M) = 0$ versus H1: $\theta_X(\beta_M) \neq 0$. (Two-Sided).
X is the primary predictor. It is a continuous, independent variable.
M is the mediator. It is a continuous variable.
Model 1: $M = \theta_0 + \theta_X(X) + e_M$
Model 2: is $Y = \beta_0 + \beta_X(X) + \beta_M(M) + e_Y$. The e_Y's are normally distributed.
Power is the probability of rejecting a false null hypothesis.
N is the number of observations on which the multiple regression is computed.
θ_X is the regression coefficient of the primary predictor in Model 1. It is sometimes refered to as a path coefficient.
β_M is the regression coefficient of the mediator in Model 2. It is sometimes refered to as a path coefficient.
σ_X is the standard deviation of X.
σ_M is the standard deviation of M.
σ_e is the standard deviation of the ei in model 2.
$\sigma(\theta_X(\beta_M))$ is the standard deviation of the coefficient product $\theta_X(\beta_M)$.
Alpha is the probability of rejecting a true null hypothesis. It should be small.

Summary Statements
A sample size of 217 achieves 90% power to detect a mediation effect of at least 0.040 as measured by the product of two regression coefficients 0.200 (primary predictor) and 0.200 (mediator) when the significance level (alpha) is 0.050. The primary predictor, X, has a standard deviation of 0.600. The medicator, M, has a standard deviation of 0.500. The standard deviation of the residuals from the two-variable regression model is 0.200.

This report shows the necessary sample sizes. The definitions of each of the columns is given in the Report Definitions section.

Plots Section





This plot shows the relationship between sample size and effect size.

Example 2 – Validation using Hand Calculations

We were unable to find a validation example in the literature, so we will present an example calculated by hand. We will use N = 100, $\theta_X = 0.3$; $\beta_M = 0.6$, $\sigma_X = 0.7$, $\sigma_M = 0.8$, and $\sigma_e = 0.2$. The two-sided significance level is set at 0.05. The calculation could proceed as follows:

$$\rho_{XM} = \frac{\theta_X \sigma_X}{\sigma_M} = \frac{(0.3)(0.7)}{0.8} = 0.2625$$

$$\sigma_{e_M}^2 = \sigma_M^2(1 - \rho_{XM}^2) = 0.8^2(1 - 0.2625^2) = 0.5959$$

$$\mathrm{V}(\theta_X) = \frac{\sigma_{e_M}^2}{N\sigma_X^2} = \frac{0.5959}{100(0.7^2)} = 0.01216122$$

$$\mathrm{V}(\beta_M) = \frac{\sigma_{e_Y}^2}{N\sigma_{e_M}^2} = \frac{0.2^2}{100(0.5959)} = 0.00067125357$$

$$\begin{aligned} z_\beta &= \frac{\theta_X \beta_M}{\sqrt{\theta_X^2 \mathrm{V}(\beta_M) + \beta_M^2 \mathrm{V}(\theta_X)}} - z_{1-\frac{\alpha}{2}} \\ &= \frac{(0.3)(0.6)}{\sqrt{0.3^2(0.00067125357) + 0.6^2(0.01216122)}} - 1.959964 \\ &= \frac{0.18}{\sqrt{0.00006041282 + 0.0043780392}} - 1.959964 \\ &= 0.741858 \end{aligned}$$

$$\text{Power} = \Phi(0.741858) = 0.7709$$

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Tests of Mediation Effect using the Sobel Test** procedure. You may then make the appropriate entries as listed below, or open **Example 2** by going to the **File** menu and choosing **Open Example Template**.

Option	Value
Design Tab	
Solve For	**Power**
Alternative Hypothesis	**Two-Sided**
Alpha	**0.05**
N (Sample Size)	**100**
θ_X (Reg Coef of X)	**0.3**
β_M (Reg Coef of M)	**0.6**
Type of Primary Predictor, X	**Continuous**
σ_X (Standard Deviation of X)	**0.7**
Type of Mediator, M	**Continuous**
σ_M (Standard Deviation of M)	**0.8**
σ_e (Standard Deviation of e_Y)	**0.2**

Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Results

Numeric Results
Two-sided alternative hypothesis

Power	Sample Size N	Reg Coef of X θ_X	Reg Coef of M β_M	Coef Product $\theta_X(\beta_M)$	Std Dev of X σ_X	Std Dev of M σ_M	Std Dev of e_Y σ_e	Std Dev of Coef Product $\sigma(\theta_X(\beta_M))$	Alpha
0.7709	100	0.300	0.600	0.180	0.700	0.800	0.200	0.067	0.050

PASS matches the calculation by hand of power = 0.7709.