Chapter 824

Tests of Mediation Effect in Linear Regression

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 Vittinghoff, Sen, and McCulloch (2009). Note that their work has been extended in Vittinghoff and Neilands (2015). We are looking into adding those extensions in a later procedure.

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 $X \sim \mathrm{N}(\mu_x, \sigma_X^2)$.

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

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

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

Vittinghoff, Sen, and McCulloch (2009) derived sample size formulas based on testing the significance of β_M in model 3. They showed that, in the null case, testing $\beta_M = 0$ is equivalent to testing for mediation. In addition to the notation above, they use ρ_{XM} as the correlation between the independent variables X and M.

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: $z_\beta = \sqrt{\frac{N\sigma_M^2\beta_M^2(1-\rho_{XM}^2)}{\sigma_{eY}^2}} - z_{1-\alpha}$.
3. Calculate: $\text{Power} = \Phi(z_\beta)$.

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.

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 determine the sample size based on the significance test of the mediator term in a linear model. Using prior analyses, they decide to use β_M = 0.2, 0.3, 0.4, ρ_{XM} = 0.6, σ_M = 0.5, and σ_e = 0.3. They set the power at 0.9 and the two-sided significance level at 0.05.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 1** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab	
Solve For	**N (Sample Size)**
Alternative Hypothesis	**Two-Sided**
Power	**0.90**
Alpha	**0.05**
β_M (Reg Coef of M)	**0.2 0.3 0.4**
ρ_{XM} (Correlation of X and M)	**0.6**
Type of Mediator, M	**Continuous**
σ_M (Standard Deviation of M)	**0.5**
σ_e (Standard Deviation of e_Y)	**0.3**

Output

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

Numeric Reports

Numeric Results

Solve For:	N (Sample Size)
Alternative Hypothesis:	Two-Sided
Hypotheses:	H0: $\beta_M = 0$ versus H1: $\beta_M \neq 0$

Power	Sample Size N	Regression Coefficient of M β_M	Correlation of X and M ρ_{XM}	Standard Deviation M σ_M	Standard Deviation e_Y σ_e	Alpha
0.9005	148	0.2	0.6	0.5	0.3	0.05
0.9014	66	0.3	0.6	0.5	0.3	0.05
0.9005	37	0.4	0.6	0.5	0.3	0.05

Model	$Y = \beta 0 + \beta x(X) + \beta M(M) + eY$. The eY's are normally distributed.
X	The primary predictor. It is an independent variable.
M	The mediator. It is a continuous variable.
Power	The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
N	The number of observations on which the multiple regression is computed.
β_M	The regression coefficient of the mediator in the model.
ρ_{XM}	The correlation between X and M.
σ_M	The standard deviation of M.
σ_e	The standard deviation of eY in the model.
Alpha	The probability of rejecting a true null hypothesis.

Summary Statements

A mediation effect (single group, Y versus X with mediator M) design will be used to test whether the mediation effect (β_M) is different from 0 (H0: $\beta_M = 0$ versus H1: $\beta_M \neq 0$). The comparison will be made using a two-sided linear regression test of the mediation effect coefficient (β_M), with a Type I error rate (α) of 0.05. The continuous mediator, M, is assumed to have a standard deviation of 0.5. The correlation between X (primary predictor) and M (mediator) is assumed to be 0.6. The standard deviation of the residuals from the two-variable regression model is assumed to be 0.3. To detect a mediation effect (mediator regression coefficient, β_M) of 0.2 with 90% power, the number of needed subjects will be 148.

Dropout-Inflated Sample Size

Dropout Rate	Sample Size N	Dropout-Inflated Enrollment Sample Size N'	Expected Number of Dropouts D
20%	148	185	37
20%	66	83	17
20%	37	47	10

Dropout Rate	The percentage of subjects (or items) that are expected to be lost at random during the course of the study and for whom no response data will be collected (i.e., will be treated as "missing"). Abbreviated as DR.
N	The evaluable sample size at which power is computed. If N subjects are evaluated out of the N' subjects that are enrolled in the study, the design will achieve the stated power.
N'	The total number of subjects that should be enrolled in the study in order to obtain N evaluable subjects, based on the assumed dropout rate. After solving for N, N' is calculated by inflating N using the formula N' = N / (1 - DR), with N' always rounded up. (See Julious, S.A. (2010) pages 52-53, or Chow, S.C., Shao, J., Wang, H., and Lokhnygina, Y. (2018) pages 32-33.)
D	The expected number of dropouts. D = N' - N.

Dropout Summary Statements

Anticipating a 20% dropout rate, 185 subjects should be enrolled to obtain a final sample size of 148 subjects.

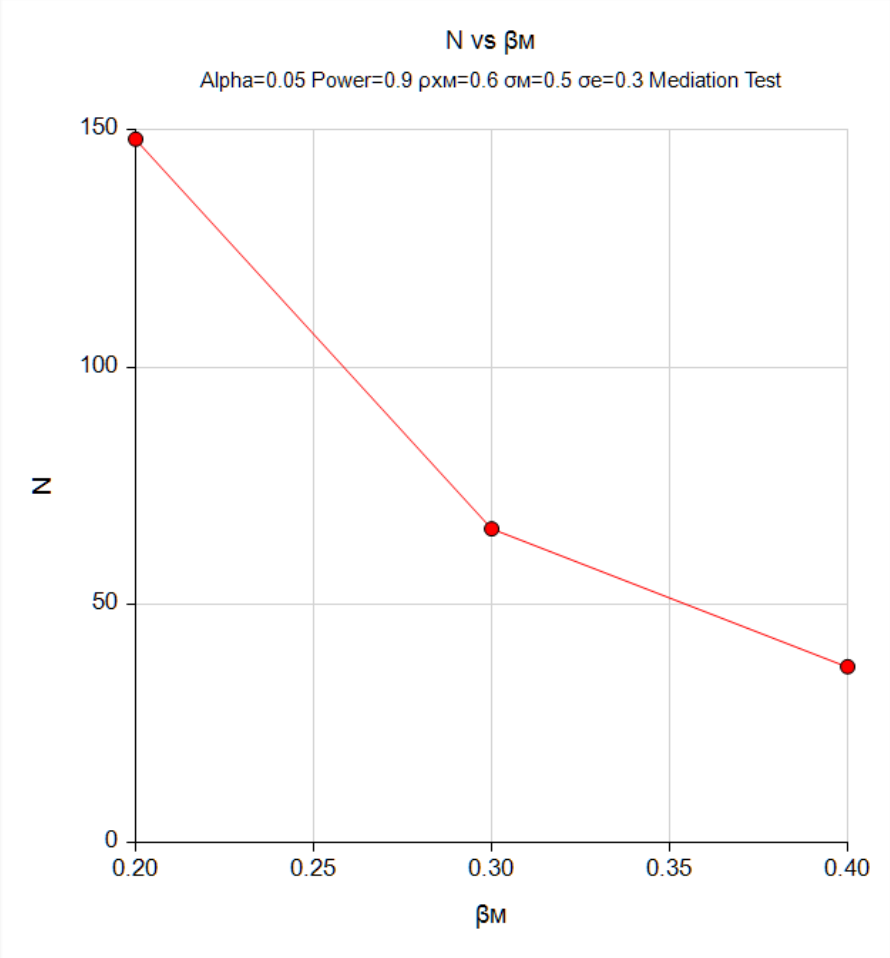
References

Vittinghoff, E., Sen, S., and McCulloch, C.E. 2009. 'Sample size calculations for evaluating mediation.' Statistics in Medicine, Vol. 28, Pages 541-557.

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

Plots Section

Plots



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

Example 2 – Validation using Vittinghoff (2009)

Vittinghoff et al. (2009) present an example on page 544 in which βM = 0.1, ρXM = 0.3, σM = 1, and σe = 1. They set the power at 0.8 and the one-sided significance level at 0.025. The computed sample size is 863.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 2** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab	
Solve For	**N (Sample Size)**
Alternative Hypothesis	**One-Sided**
Power	**0.80**
Alpha	**0.025**
βM (Reg Coef of M)	**0.1**
ρXM (Correlation of X and M)	**0.3**
Type of Mediator, M	**Continuous**
σM (Standard Deviation of M)	**1**
σe (Standard Deviation of eY)	**1**

Output

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

Numeric Results

Solve For:	N (Sample Size)
Alternative Hypothesis:	One-Sided
Hypotheses:	H0: H0: βM ≤ 0 versus H1: βM > 0 or H0: βM ≥ 0 versus H1: βM < 0

Power	Sample Size N	Regression Coefficient of M βM	Correlation of X and M ρXM	Standard Deviation M σM	Standard Deviation eY σe	Alpha
0.8002	863	0.1	0.3	1	1	0.025

PASS matches the calculation of N = 863.