Chapter 826

Tests of Mediation Effect in Poisson Regression

Introduction

This procedure computes power and sample size for a test of the mediation effect in a Poisson regression with a dependent (output) variable Y of counts and an independent (input) variable X. Interest focuses on the interrelationship between Y, X, and a third independent 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

Vittinghoff, Sen, and McCulloch (2009) derived sample size formulas for testing the mediation effect based on testing the significance of β_M in the Poisson regression model

$$\log(\frac{P}{1-P}) = \beta_0 + \beta_X X + \beta_M M.$$

They showed that testing $\beta_M = 0$ is equivalent to testing for a significant mediation effect. In addition to the notation above, they use ρ_{XM} as the correlation between the independent variables X and M and E(Y) to represent the marginal mean of Y.

Calculating the Power

Power calculations are based on the following assuming that if X is continuous, the variance of X is σ_X^2, or, if binary, the variance of X is $p_X(1-p_X)$ where $p_X = \text{Prob}(X=1)$. Similar notation is used for the mediator variable M.

1. Determine the critical value $z_{1-\alpha}$ from the standard normal distribution where α is the probability of a type-I error.
2. a. If X is continuous and M continuous, use $z_\beta = \sqrt{N\sigma_M^2\beta_M^2(1-\rho_{XM}^2)E(Y)} - z_{1-\alpha}$.

   b. If X is binary and M continuous, use $z_\beta = \sqrt{N\sigma_M^2\beta_M^2(1-\rho_{XM}^2)E(Y)} - z_{1-\alpha}$.

   c. If X is continuous and M binary, use $z_\beta = \sqrt{N\beta_M^2(1-\rho_{XM}^2)E(Y)F1} - z_{1-\alpha}$.

   d. If X is binary and M binary, use $z_\beta = \sqrt{N\beta_M^2(1-\rho_{XM}^2)E(Y)F2} - z_{1-\alpha}$.

where

$$F1 = [GH\sigma_{X.M}^2]/[(G+H)^2\sigma_{X.M}^2 + GH(\mu_1 - \mu_0)^2]$$

$$\mu_0 = -\rho_{XM}\sigma_X\sqrt{\frac{p_M}{1-p_M}} \text{ when } M = 0$$

$$\mu_1 = \rho_{XM}\sigma_X\sqrt{\frac{1-p_M}{p_M}} \text{ when } M = 1$$

$$\sigma_{X.M}^2 = \sigma_X^2(1 - \rho_{XM}^2)$$

$$G = p_M\exp(\beta_X\mu_1 + \beta_M)$$

$$H = (1 - p_M)\exp(\beta_X\mu_0)$$

$$F2 = [BCD + BCE + BDE + CDE]/[(B + C + D + E)(B + D)(C + E)]$$

$$B = p_{00}$$

$$C = p_{10}\exp(\beta_X)$$

$$D = p_{01}\exp(\beta_M)$$

$$E = p_{11}\exp(\beta_X + \beta_M)$$

$$p_{11} = p_Xp_M + \rho_{XM}\sqrt{p_X(1-p_X)p_M(1-p_M)} \text{ with } \rho_{XM}|0 < p_{11} < p_X^{p_M}$$

$$p_{10} = p_X - p_{11}$$

$$p_{01} = p_M - p_{11}$$

$$p_{00} = 1 - p_{01} - p_{10} - p_{11}$$

3. Calculate: Power = $\Phi(z_\beta)$.

Notes

1. Use $\frac{\alpha}{2}$ instead of α for two-sided test.

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

β_M (Reg Coef of M)

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

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

Y is the count (response), X is the primary predictor variable, and M is the mediator variable.

Range

β_M can be any value other than zero.

ρ_{XM} (Correlation of X and M)

Enter one or more values for the correlation of X and M.

Range

$0 < \rho_{XM} < 1$.

Data Types of X and M

Specify the data types (continuous or binary) of independent variables X and M.

X is the primary predictor variable.

M is the mediator variable.

Note

Here, 'Continuous' refers to data that are normally distributed.

σ_X (Standard Deviation of X)

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

Range

$0 < \sigma_M$.

σ_M (Standard Deviation of M)

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

Range

$0 < \sigma_M$.

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.

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.

E(Y) (Marginal Mean of Y)

Enter one or more values for the marginal mean of Y, where Y is the count variable. Y cannot be zero since the model uses log(Y).

Range

$0 < E(Y)$.

Example 1 – Finding Sample Size

Researchers are studying the relationship between a count-type dependent variable (Y) and a continuous independent variable (X). They want to understand the impact of a continuous 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 Poisson regression. Using prior analyses, they decide to use β_M = 0.3, 0.4, 0.5, ρ_{XM} = 0.4, σ_M = 1, and E(Y) = 0.5 0.8. 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 in Poisson Regression** 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**
Data Types of X and M	**X = Continuous, M = Continuous**
β_M (Reg Coef of M)	**0.3 0.4 0.5**
ρ_{XM} (Correlation of X and M)	**0.4**
σ_M (Standard Deviation of M)	**1**
E(Y) (Marginal Mean of Y)	**0.5 0.8**

Annotated Output

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

Numeric Results

Numeric Results
Data Types: X is continuous and M is continuous
Two-sided alternative hypothesis

Power	Sample Size N	Reg Coef of M β_M	Corr of X and M ρ_{XM}	Std Dev of M σ_M	Marg Mean of Y E(Y)	Alpha
0.9000	278	0.300	0.400	1.00	0.50	0.050
0.9004	174	0.300	0.400	1.00	0.80	0.050
0.9012	157	0.400	0.400	1.00	0.50	0.050
0.9008	98	0.400	0.400	1.00	0.80	0.050
0.9026	101	0.500	0.400	1.00	0.50	0.050
0.9021	63	0.500	0.400	1.00	0.80	0.050

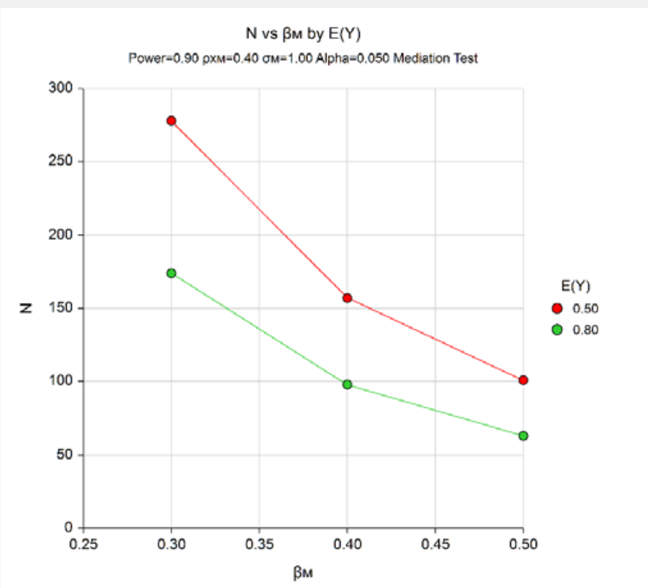
References
Vittinghoff, E., Sen, S., and McCulloch, C.E. 2009. 'Sample size calculations for evaluating mediation.' Statistics in Medicine, Vol. 28, Pages 541-557.

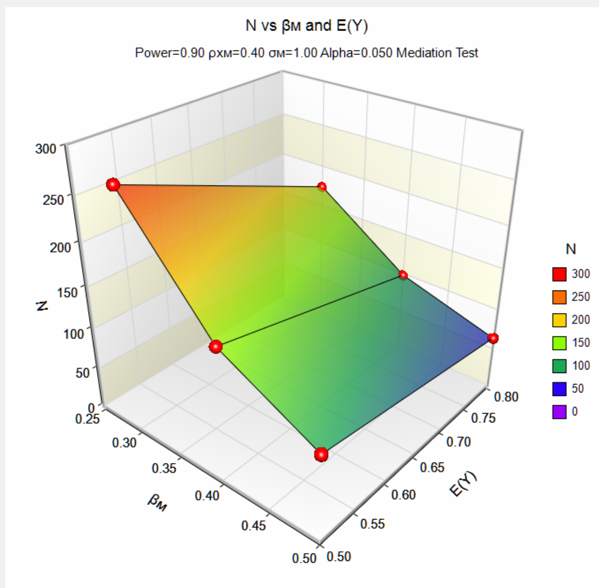
Report Definitions
Hypotheses: H0: β_M = 0 versus H1: $\beta_M \neq 0$. (Two-Sided).
X is the primary predictor. It is a continuous variable.
M is the mediator. It is a continuous variable.
Model: log(Y) = $\beta 0 + \beta x(X) + \beta_M(M)$.
Power is the probability of rejecting a false null hypothesis.
N is the number of observations on which the multiple regression is computed.
β_M is the regression coefficient of the mediator in the model.
ρ_{XM} is the correlation between X and M.
σ_M is the standard deviation of M.
E(Y) is the marginal mean of the outcome, Y.
Alpha is the probability of rejecting a true null hypothesis. It should be small.

Summary Statements
A sample size of 278 achieves 90% power to detect a mediation effect (as measured by the regression coefficient of M, β_M) of at least 0.300 when the two-sided significance level (alpha) is 0.050. The correlation between X (the primary predictor) and M (mediator) is 0.400. The medicator, M, has a standard deviation of 1.00. The marginal mean of the outcome, E(Y), is 0.50.

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

Plots Section





These plots show the relationship between sample size and the regression coefficient.

Example 2 – Validation using Vittinghoff (2009)

Vittinghoff et al. (2009) present an example on page 546 in which β_X = log 1.4 (0.3365), β_M = log 1.35 (0.3001), ρ_{XM} = 0.5, σ_X = 1, P(M=1) = 0.25, and E(Y) = 0.5. They set the power at 0.8 and the one-sided significance level at 0.025. The computed sample size is 1037.

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 in Poisson Regression** 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	**N (Sample Size)**
Alternative Hypothesis	**One-Sided**
Power	**0.80**
Alpha	**0.025**
Data Types of X and M	**X = Continuous, M = Binary**
β_X (Reg Coef of X)	**0.3365**
β_M (Reg Coef of M)	**0.3001**
ρ_{XM} (Correlation of X and M)	**0.5**
σ_X (Standard Deviation of X)	**1**
Probability M = 1	**0.25**
E(Y) (Marginal Mean of Y)	**0.5**

Output

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

Numeric Results

Numeric Results
Data Types: X is continuous and M is binary
One-sided alternative hypothesis

Power	Sample Size N	Reg Coef of X β_X	Reg Coef of M β_M	Corr of X and M ρ_{XM}	Std Dev of X σ_X	Prob M=1 P(M=1)	Marg Mean of Y E(Y)	Alpha
0.8001	1037	0.337	0.300	0.500	1.00	0.250	0.50	0.025

PASS matches the calculation of N = 1037.