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

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

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

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

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

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

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

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

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

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

$$B = p_{00}$$

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

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

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

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

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

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

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

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

Notes

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

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

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**
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**

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 of M σ_M	Marginal Mean of Y E(Y)	Alpha
0.90003	278	0.3	0.4	1	0.5	0.05
0.90043	174	0.3	0.4	1	0.8	0.05
0.90116	157	0.4	0.4	1	0.5	0.05
0.90080	98	0.4	0.4	1	0.8	0.05
0.90261	101	0.5	0.4	1	0.5	0.05
0.90205	63	0.5	0.4	1	0.8	0.05

Model $\log(Y) = \beta_0 + \beta_X(X) + \beta_M(M)$.
X The primary predictor. It is a continuous 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.
E(Y) The marginal mean of the outcome, Y.
Alpha The probability of rejecting a true null hypothesis.

Summary Statements

A mediation effect (single group, count response 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 Poisson 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 1. The marginal mean of the outcome, E(Y), is assumed to be 0.5. The correlation between X (primary predictor) and M (mediator) is assumed to be 0.4. To detect a mediation effect (mediator regression coefficient, β_M) of 0.3 with 90% power, the number of needed subjects will be 278.

Dropout-Inflated Sample Size

Dropout Rate	Sample Size N	Dropout-Inflated Enrollment Sample Size N'	Expected Number of Dropouts D
20%	278	348	70
20%	174	218	44
20%	157	197	40
20%	98	123	25
20%	101	127	26
20%	63	79	16

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, 348 subjects should be enrolled to obtain a final sample size of 278 subjects.

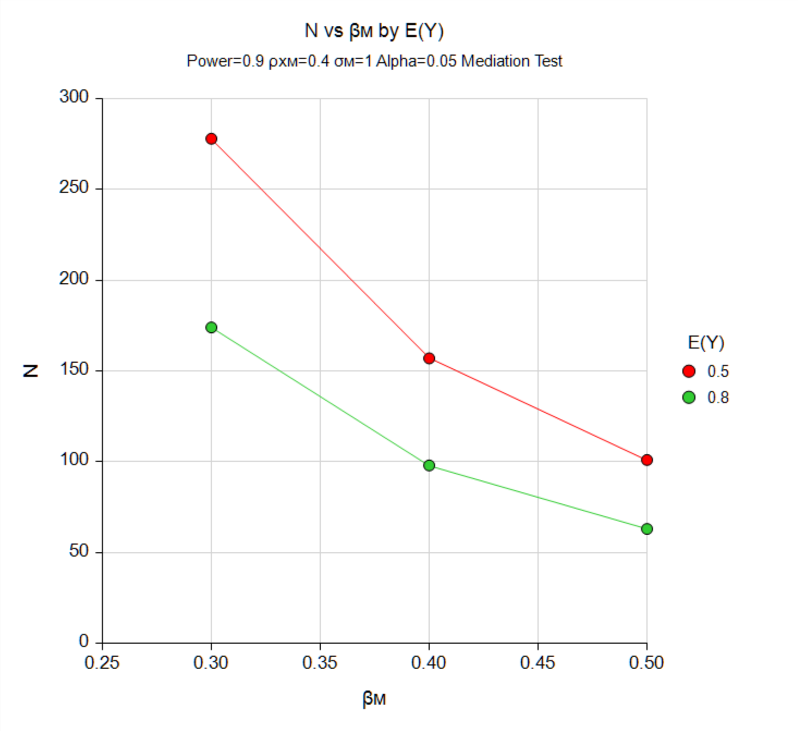
References

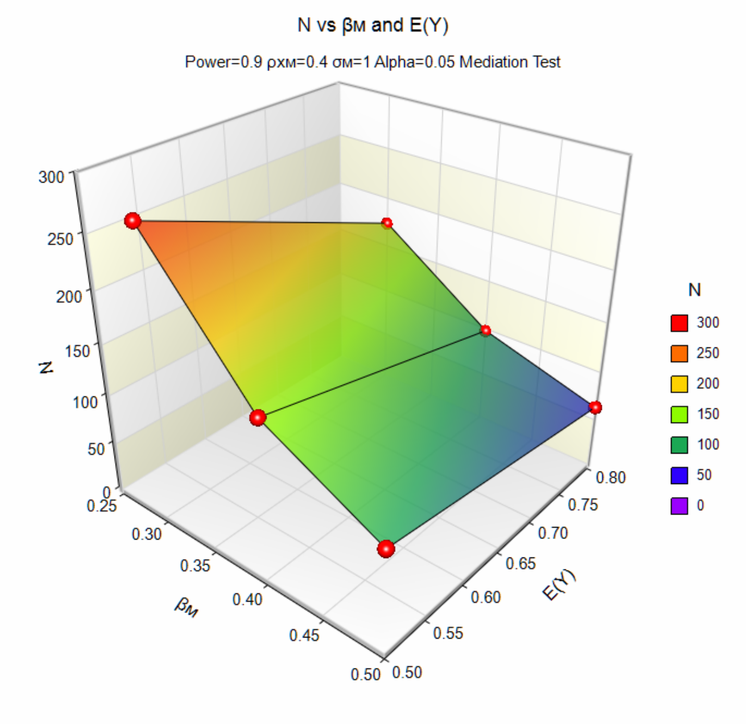
Vittinghoff, E., Sen, S., and McCulloch, C.E. 2009. 'Sample size calculations for evaluating mediation.' Statistics in Medicine, Vol. 28, Pages 541-557.

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

Plots Section

Plots





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

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**
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

Solve For: N (Sample Size)
Alternative Hypothesis: One-Sided
Hypotheses: H0: H0: $\beta_M \le 0$ versus H1: $\beta_M > 0$ or H0: $\beta_M \ge 0$ versus H1: $\beta_M < 0$

Power	Sample Size N	Regression Coefficient X β_X	Regression Coefficient M β_M	Correlation of X and M ρ_{XM}	Standard Deviation of X σ_X	Probability M = 1 P(M = 1)	Marginal Mean of Y E(Y)	Alpha
0.8001	1037	0.337	0.3	0.5	1	0.25	0.5	0.025

PASS matches the calculation of N = 1037.