Power Analysis for Mediation Models: Methods & Results

Heqiao Ruan, Shubha Sankar Banerjee
Satish Iyengar (Advisor)

Department of Statistics
University of Pittsburgh

November 9, 2023

Contents

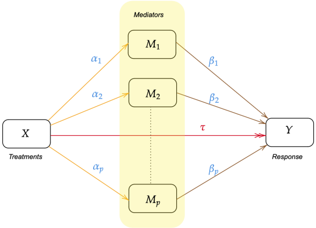
1. Introduction
2. Methods and Results
3. R code
4. Dealing with confounders
5. References

Introduction

Prior to implementing the study, researchers want to know the sample size needed to detect an effect of interest in complex experimental designs.

We assume the following structure for the mediation setup.

- Treatments $X \in \mathbb{R}^{n \times 1}$.
- Mediators: M
- Responses: Y
- Later,
  we will include confounders Z



Problems to solve

- Is the sample size large enough to detect mediation effect in psychiatric studies?
- We conduct **power analysis** under the hypothesis testing framework.
- We want to maximize the power with control of the type-I error.

Steps for conducting power calculation:

- Set up a hypothesis test.
- Fix power and type-I error.
- Search for the minimal sample size to achieve the desired power with the type-I error control.

Effect sizes

Hypothesis testing problem is:

$$H_0 : \text{No Mediation effect} \quad vs. \quad H_a : \text{There are mediation effects}$$

- Traditional Methods:
  - Sobel test Sobel (1982): z-test is $\frac{\hat{\alpha}\hat{\beta}}{\sqrt{\hat{\alpha}^2\hat{\sigma_\beta}^2+\hat{\beta}^2\hat{\sigma_\alpha}^2}}$. *Biased.*
  - Joint Testing Huang (2018): $H_{00} : \alpha = 0, \beta = 0$, $H_{01} : \alpha = 0, \beta \neq 0$, $H_{10} : \alpha \neq 0, \beta = 0$ and $H_0 : H_{00} \cup H_{01} \cup H_{10}$. $vs.$ $H_a : H_0^C$. *Very conservative for multiple comparisons.*
  - Both effect-sizes are biased and require multiple testing procedure.
- **The effect size we adopt is R^2** Fairchild et al. (2009)
  - $R^2 = R^2_{Y\sim M} + R^2_{Y\sim X} - R^2_{Y\sim MX}$ where the $R^2_{Y\sim M}$ is the R^2 for regression of Y onto M, $R^2_{Y\sim X}$ is the R^2 for the regression of Y onto X, $R^2_{Y\sim MX}$ is the R^2 for the regression of Y onto $[X, M]$.
  - R^2 can be interpreted as the amount of variance in Y that is explained by M, specific to the mediated effect if no other unmeasured confounders exist.
- R^2 is a unit measure can work comfortably in multiple mediators without multiple comparison procedures.

High-level pseudocode for power analysis

Assume sample size n, univariate treatments $X \in \mathbb{R}^{n\times 1}$, p_M mediators and p_Y-dimensional response $Y \in \mathbb{R}^{n\times p_Y}$

Algorithm Monte Carlo simulation of mediation effects and bootstrapped power

for $k = 1, \ldots, n_{mc}$ (Monte Carlo samples, 500 by default), $count_{pow} = 0$:

In Steps 1-3, the random numbers are generated n times.

- S1 Generate $X \sim N(0,1)$
- S2 Generate $M = X\alpha + \varepsilon_M,\ \varepsilon_M \sim N(0, \sigma_M^2)$
- S3 Generate $Y_j = M\beta_j + \tau' X + \varepsilon_Y,\ j = 1, \ldots, p_Y,\ \varepsilon_Y \sim N(0, \sigma_Y^2)$, where, $\mathbf{Y} = (Y_1, \ldots, Y_{p_Y})$.
- S4 **for** $b = 1, \ldots, B$ (Bootstrapping; typically $B = 400$):
  - S4.1 Calculate the value of the chosen effect size.
- S5 Compute the 95% confidence interval (C_B) for the chosen effect size using the B bootstrapped estimates obtained in $S4$.
- S6 $count_{pow} = count_{pow} + \mathbf{1}\{0 \notin C_B\}$

Set $pow = count_{pow}/n_{mc}$

Choosing path coefficients

- Following Fritz and Mackinnon (2007), we consider different *levels* for α and β generation in Algorithm 1.
- For each path coefficient, we specify the levels *Small* (S = 0.14), *Medium* (M = 0.26), *Large* (L = 0.59) and the level 0.
- We generate α and β from each level (S, M or L) from a uniform distribution around a small neighborhood of the levels.
  eg. for small α, generate $\alpha \sim Unif(S-\varepsilon, S+\varepsilon)$. We have typically chosen $\varepsilon = 0.02$.
- To allow for more flexibility in our analysis in model and effect size specification, we split the analysis in three cases:
  - Full Mediation: $\alpha \neq 0,\ \beta \neq 0,\ \tau = 0$.
  - Partial mediation: $\alpha > 0,\ \beta > 0,\ \tau \neq 0$.
  - No Mediation: $\alpha = 0$ or $\beta = 0,\ \tau \neq 0$.

Type-I Error

When at least one of α, β is 0, there's no indirect effect, type I error is evaluated.

Linear Model	**n = 100**	**n = 200**	**n = 500**
Zero α, Small β	0.036	0.032	0.044
Zero α, Medium β	0.048	0.028	0.048
Small α, Zero β	0.044	0.040	0.016
Medium α, Zero β	0.028	0.024	0.036

Type-I error Analysis for simulation of linear model

Power	**n = 100**	**n = 200**	**n = 500**
Zero α, Small β	0	0	0.0025
Zero α, Medium β	0.005	0.010	0.012
Small α, Zero β	0.028	0.024	0.042
Medium α, Zero β	0.032	0.038	0.048

Type-I Error Analysis for simulation of generalized linear model(GLM)

Good control of type-I Error under 5%.

Power results for R^2 effect-size

- 5 mediators in partial mediation setting, 3-dimensional response Y.
- Desirable power for not-too-small effect size.

Effects	**n = 100**	**n = 200**	**n = 500**
SS	[0.068, 0.052, 0.088]	[0.124, 0.132, 0.164]	[0.240, 0.248, 0.460]
SM	[0.164, 0.128, 0.124]	[0.152, 0.148, 0.362]	[0.700, 0.392, 0.348]
SL	[0.480, 0.492, 0.416]	[0.720, 0.716, 0.584]	[0.988,0.980,0.968]
ML	[0.936, 0.904, 0.908]	[0.996, 0.982, 0.992]	[1,1,1]
LL	[1,1,0.998]	[1,1,1]	[1,1,1]

Partial Mediation analysis for R^2 effect size. S: Small, M: Medium, L: Large. ML stands for medium α, large β.

The required sample size to achieve desired power decreased as effect size increase.

Power Result compared to current methods

Compare our power on the sample-size that reached desired power in existing methods (Fritz and Mackinnon (2007)).

Sample Size	Effect Size	Power	Sample Size	Effect Size	Power
n = 530	SSS	0.832	n = 398	MSM	0.828
n = 402	SHS	0.804	n = 116	MHM	0.823
n = 400	SMS	0.724	n = 71	MMM	0.794
n = 413	SLS	0.705	n = 53	MLM	0.787
n = 368	HSH	0.813	n = 396	LSL	0.886
n = 158	HHH	0.872	n = 115	LHL	0.867
n = 148	HMH	0.856	n = 54	LML	0.841
n = 120	HLH	0.781	n = 32	LLL	0.852

Power Comparison with other methods - Linear Models with the **baseline power 0.8** for the minimal sample-size across the test in that condition. S: 0.14, H: 0.26, M: 0.39, L: 0.59, SHS means small α, high β and Small τ.

We achieve higher power than the current methods (Sobel test, Joint test) when the effect sizes are not too small.

Extension to Non-normal Response

For non-normal response:

- In simulation we impose a link function to account for non-normal response. $\tilde{Y} = \mathbf{g}^{-1}(Y)$. For normal response, link $\mathbf{g}$ is identity. For binary response, link $\mathbf{g}$ is logit function.
- We adapt the R^2 to MacFadden pseudo-R^2 (McFadden et al. (1974)) for generalized linear Model(GLM): $pR^2 = pR^2_{Y\sim M} + pR^2_{Y\sim X} - pR^2_{Y\sim MX}$.

Power	**n = 100**	**n = 200**	**n = 500**
Small α, β	0.146	0.228	0.544
Medium α, β	0.678	0.752	0.840
Large α, β	0.968	0.996	1
Small α, medium β	0.206	0.416	0.630
Small α, large β	0.402	0.614	0.862
Medium α, large β	0.884	0.998	1

Partial-Mediation Power Analysis for Logistic Regression

For binary response the required sample-size are larger than the linear model.

Mediation for longitudinal studies

For study planning on multiple time points, consider also temporal dependence an effects. We incorporate a random intercept for each subject. We follow the model in Pan et al. (2018).

$i = 1, 2, ..., n$: indexes the n subjects, $j = 1, 2, ..., p_M$ indexes the dimension of mediators, $k = 1, 2, ..., T$ indexes the time points

$$M_{ijk} = \tau_{jk,1} + \alpha_j X_i + \xi_{M,ijk},\ \xi_{M,ijk} \sim N(0, \sigma^2)$$

$$\tau_{jk,1} = \mu_1 + \rho_{jk,1},\ \rho_{jk,1} \sim N(0, \sigma_\tau^2)$$

$$Y_{ik} = g^{-1}(\tau_{k,2} + \beta_\tau X_i + \sum_{j=1}^{p_M} M_{ijk}\beta_j + \xi_{Y,ik})$$

$$\xi_{Y,ik} \sim N(0, \sigma^2),\ \tau_{k,2} = \mu_2 + \rho_{k,2},\ \rho_{k,2} \sim N(0, \sigma_\tau^2)$$

$\tau_{jk,1}$ represent the random intercept, $\tau_{jk,2}$ represent the random intercept, g is the link function specific for the type of response.
We assume that both the mean of random intercept μ_1 and μ_2 is 0.

Account for longitudinal effects

In practice, researchers need to specify the intra-class correlation (ICC) for the model: $\frac{\sigma_\tau^2}{\sigma_\tau^2+\sigma^2}$ is the proportion of variance attributed to the longitudinal effect.

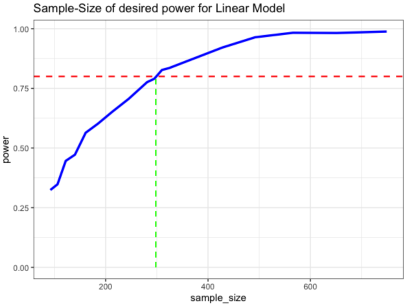
Effect-Size	**ICC**	**100**	**200**	**300**	**400**	**500**	**600**
LML	0.2	0.650	0.796	0.876	0.914	0.960	0.980
LML	0.4	0.532	0.672	0.812	0.856	0.868	0.926
LML	0.6	0.436	0.552	0.664	0.702	0.744	0.802
LML	0.8	0.300	0.420	0.516	0.604	0.664	0.698

Partial Mediation Power Analysis table for mixed-effects model with binary response; α: Large, β: Medium, τ: Large. LML means large α, medium β and large τ

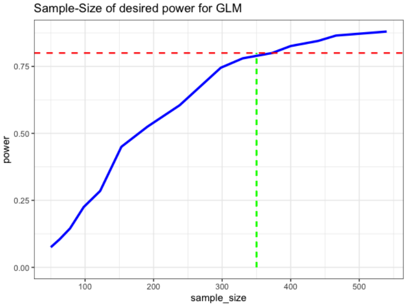
Typically we will need a larger sample-size to achieve desirable power after accounting for the longitudinal effect.

How to conduct power calculation in practice?

- For the specified level effect-size and mediation structure, find the sample-size achieving desired power (by default 0.8).



Power Calculation for linear model with Small α, Medium β



Power Calculation for generalized linear model with Small α, Medium β

Practical Usage of the software

```
mediation_power_analysis_r2 <- function(n, p_M, p_Y, B, errorM, errorY, mean_tau, error_tau,
                                        n_mc, mag_XM,
                                        mag_MY, alpha_level = 0.05, covariate_adjustment = FALSE,
                                        n_z, n_cont, snr)
```

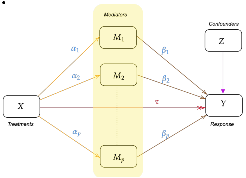
- Software in **R**
- Parameters to specify:
  - p_M: Number of mediators; n: Number of samples
  - mag_{XM}: Magnitude of the path between X (Treatment) and M (Mediator), mag_{MY}: Magnitude of the path between M (Mediator) and Y (Response). 1: Small, 2: Medium, 3: Large.
- Output: Power value between 0 and 1.
- Default Parameter:
  - errorM: random error for mediators, by default 1.
  - errorY: random error for response, by default 1.
  - error_tau: random error for the direct effect, by default 0.05.
  - B: number of bootstrap samples, by default 400.
  - n_mc: number of monte-carlo simulations, by default 500.

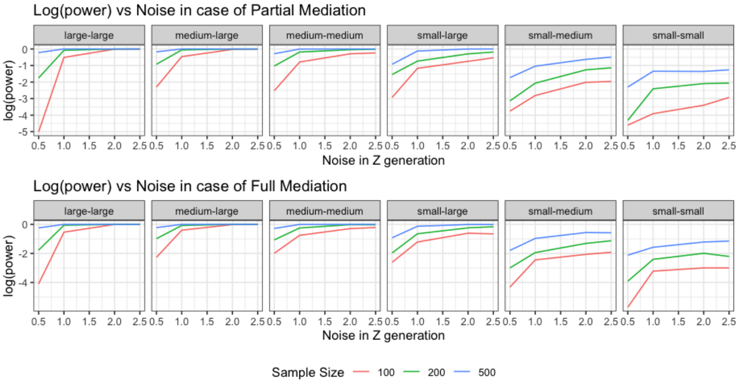
The typical computation time for a fixed effect size combination is around 8-10 minutes.

Studying effects of potential confounders

We were interested to see how robust the analysis was in presence of confounders with varying degrees of association with the response.

We simulate Z, the confounder, by adding progressively increasing Gaussian Noise to the response Y.





Innovation and Future Directions

- Innovation:
  - We adopted the R^2 effect-size for testing the mediation effect to conduct power calculation.
  - We developed a way to do power calculation for covariate-adjusted mediation analysis and demonstrated its efficiency and robustness.
  - We extended the R^2 effect-size for testing mediation effects to non-normal responses and account for the longitudinal effects.
- Future Directions:
  - Compare the R^2 effect-size to sobel test and joint test.
  - Extend the R^2 effect-size to more complex mediation structures.

References

Fairchild, A. J., MacKinnon, D. P., Taborga, M. P., and Taylor, A. B. (2009). R 2 effect-size measures for mediation analysis. *Behavior research methods*, 41(2):486–498.

Fritz, M. S. and Mackinnon, D. P. (2007). Required sample size to detect the mediated effect. *Psychol Sci*, 18(3):233–239.

Huang, Y.-T. (2018). Joint significance tests for mediation effects of socioeconomic adversity on adiposity via epigenetics. *Annal of Applied Statistics*, 12(3):1535–1557.

McFadden, D. et al. (1974). Conditional logit analysis of qualitative choice behavior. *Frontiers in econometrics*, pages 105 – 142.

Pan, H., Liu, S., Miao, D., and Yuan, Y. (2018). Sample size determination for mediation analysis of longitudinal data. *BMC Medical Research Methodology*, 18(1):1–11.

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological methodology*, 13:290–312.