Determining Power and Sample size for Mediation Models

Alexander M. Schoemann [1] Aaron J. Boulton [2]

[1]East Carolina University

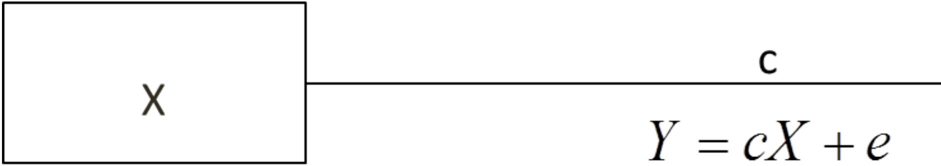
[2]University of North Carolina - Chapel Hill

SSSP 2015

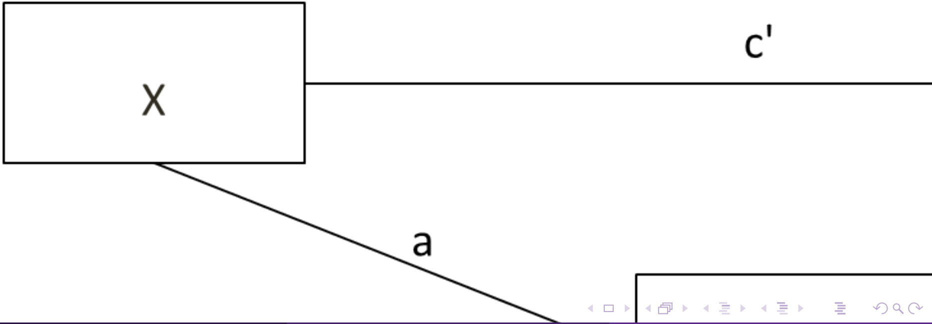
Basic Mediation

- Three variable system
  - Predictor variable - X
  - Outcome variable - Y
  - Mediator variable - M
- M (partially) explains the relationship between X and Y

Mediation Vocabulary



$$Y = cX + e$$



How do we get estimates of a, b, and c?

- We get estimates of the slopes through a series of separate regression models:
  - $Y = cX + e$
  - $Y = bM + c'X + e$
  - $M = aX + e$

what do that paths tell us about mediation?

- When investigating mediation we are interested in if strength of the relationship between X and Y is reduced when M is included in the regression equation
  - Is the direct effect (c') less than the total effect (c)?
  - Does $c - c' = 0$?
- c - c' is also called the indirect effect
- The indirect effect can also be estimated by multiplying a*b
- a*b and c - c' should provide the same value

Testing the indirect effect

- When we test for mediation we test if the indirect effect is different from zero.
  - If the indirect effect does not differ from zero there is no evidence of mediation
  - If the indirect effect differs from zero there is evidence of mediation

Testing the indirect effect

- There have been many methods developed to assess the indirect effect:
  - Difference in coefficients (Baron & Kenny, 1984)
  - Sobel test
  - Distribution of the product method
  - Bootstrap confidence intervals
  - Monte Carlo confidence intervals

Testing the indirect effect: power

- Methods of testing the indirect effect have different power for the same data (e.g. Fritz & MacKinnon, 2007, MacKinnon, Lockwood, & Williams, 2004)
    - Difference in coefficients method has extremely low power (especially when c is small)
    - Distribution of the product, bootstrap CI and Monte Carlo CI have higher power
        - Power for these methods is generally comparable

Testing the indirect effect: power

- What else affects power to detect an indirect effect?
  - Sample size
  - Effect size
    - Effect size of both the a and b path

Determing power for an indirect effect

- Methods of power analysis should match methods of analysis
- No analytic method of power analysis for distribution of the product, bootstrap CI or Monte Carlo CI!
- We need to use a Monte Carlo power analysis

Monte Carlo Power Analysis

- General steps in a Monte Carlo Power Analysis
  1. Specify all population parameters
  2. Create a sample of size N, based on population parameters
  3. Analyze sample data from Step 2 with chosen statistical method(s)
  4. Repeat steps 2 and 3 for each of r replications (often $r>1000$)
  5. The proportion of replications with a significant parameter is an estimate of power (for all parameters not equal to 0)

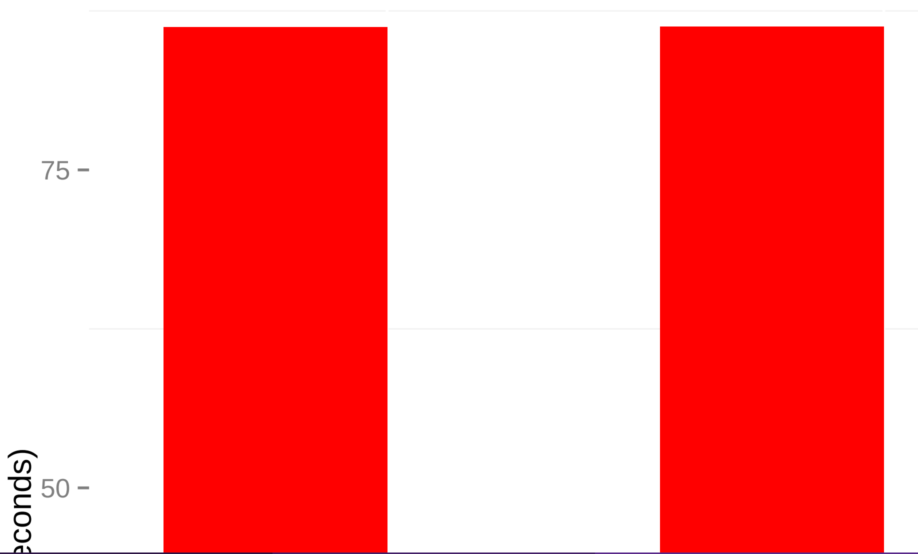
Monte Carlo Power Analysis

- Best practice approach for assessing power in mediation models (Zhang, 2014)
  - Can utilize bootstrapping, distribution of the product method, or Monte Carlo CI
- Two main limitations to Monte Carlo Power Analysis for mediation models
  - Limited, user friendly software (e.g. Mplus, bmem, simsem)
  - Very time consuming (especially with boostrapping!)

Monte Carlo Power Analysis

- Overcoming limitations
  - Use Monte Carlo CI to test indirect effects
    - Faster than bootstrapping
    - Estimates of power are extremely similar ($r = .99$)

Monte Carlo Power Analysis



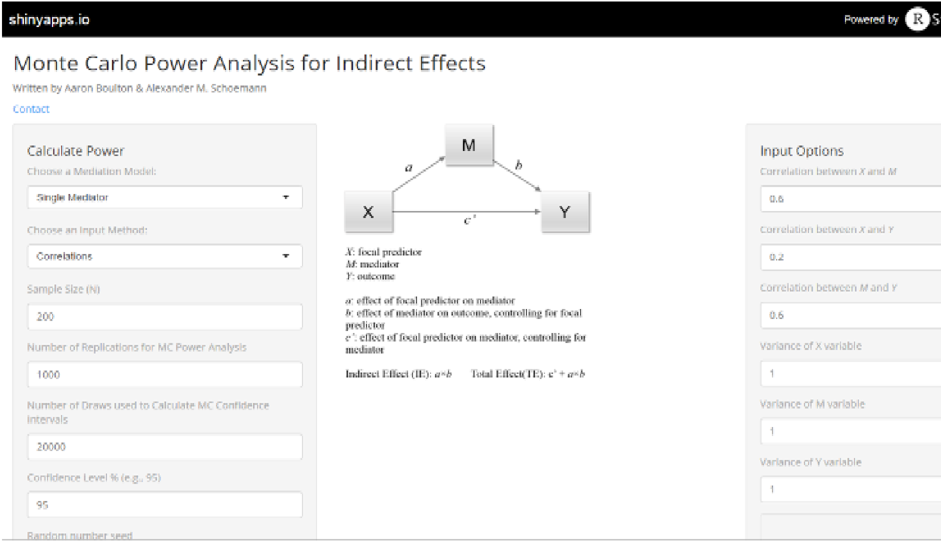
Monte Carlo Power Analysis

- Overcoming limitations
  - Use Monte Carlo CI to test indirect effects
    - Faster than bootstrapping
    - Estimates of power are extremely similar ($r = .99$)
  - Create user friendly software
    - Shiny app for simple mediation models

Power estimation app

- Available from MARIab.org
- Web based or run locally on your computer
    - Requires R is installed on a computer (it helps if RStudio is too)
- Population parameters can be entered as correlations or r^2 for each path

Power estimation app



Future directions

- Expansion of models available in app
  - 2 and 3 mediator models
  - Conditional process models
- Integration of other methods of sample size determination
  - Accuracy in parameter estimation (Kelly & Maxwell, 2003)
  - Varying N Monte Carlo simulations (Schoemann, Pornprasermanit, Miller, & Wu, 2014)

Thank you!

- Questions?
- Slides and shiny app from today at: `http://http://MARlab.org/Supplemental_Materials/`
- email: schoemanna@ecu.edu

References

- Fritz, M. S., & MacKinnon, D. P. (2007). Required Sample Size to Detect the Mediated Effect. *Psychological Science, 18*, 233-239.
- Kelley K., & Maxwell S. E. (2003). Sample size for multiple regression: Obtaining regression coefficients that are accurate, not simply significant. *Psychological Methods 8*, 305-321.
- MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence Limits for the Indirect Effect: Distribution of the Product and Resampling Methods. *Multivariate Behavioral Research, 39*, 99-128.
- Schoemann, A. M., Miller, P. R., Pornprasertmanit, S., & Wu, W. (2014). Using Monte Carlo simulations to determine power and sample size in planned missing data designs. *International Journal of Behavioral Development, 38*, 471-479.
- Zhang, Z. (2014). Monte Carlo based statistical power analysis for mediation models: Methods and software. *Behavior Research Methods*, 46, 1184-1198.