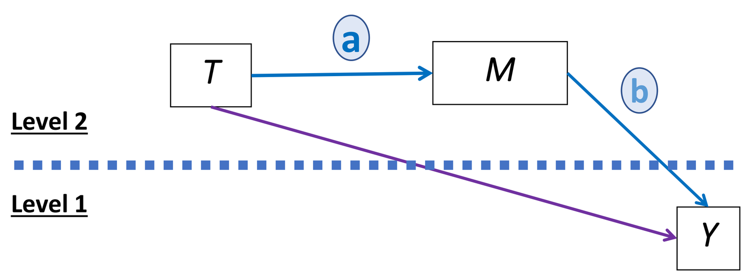
Optimal Sample Allocation in Experimental Studies Probing Mediation Effects

Zuchao Shen, Ph.D. & Feng Liu, Ph.D.
University of Florida & Western Illinois University

Introduction

- Mediation analyses often represent a cornerstone in the development of scientific theories because they investigate the underlying theory of action (Kelcey & Shen, 2020).
- A key consideration in designing mediation studies in cluster-randomized trials is to identify the sample sizes that provide a sufficient level of power for detecting mediation effects while efficiently using resources (Cox & Kelcey, 2019; Kelcey & Shen, 2020).



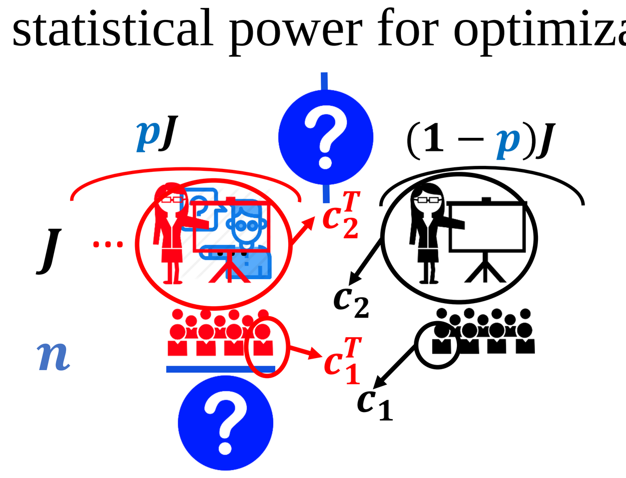
- The conventional optimal design framework has been developed for main and mediation effects for multilevel experiments (e.g., Raudenbush, 1997). However, it has assumed equal sampling costs between treatment conditions and constrained design with balanced ones (i.e., equal number of groups in each condition).
- Recent literature has developed a flexible optimal design framework for main treatment effects in multilevel experiments (Shen & Kelcey, 2020, in press a, in press b).
- The purpose of this study is to extend this flexible optimal sampling framework to cluster-randomized trials targeting mediation effects with cluster-level mediators.

Optimal Sample Allocation

- Sample allocation is the sample sizes (to be) allocated at different levels and treatment conditions in (group-randomized) experiments.
- Optimal sample allocation is the sample allocation that achieves the maximum statistical power under a fixed budget, or uses the minimum resources to gain adequate statistical power.
- Why optimal sample allocation? It can be used to improve design efficiency and/or statistical precision.
  - More power, smaller effect sizes under a fixed budget.
  - Less budget to achieve a fixed statistical power.

Methods

- **Sobel Test**
  - $Z_{ab} = ab/\sqrt{\sigma^2_{ab}}$ with $\sigma^2_{ab} = a^2\sigma^2_b + b^2\sigma^2_a$ with $\sigma^2_a = \frac{1}{p(1-p)J}$ and $\sigma^2_b = \frac{\rho - b^2 + (1-\rho)/n}{J}$. Power is calculated in a standard normal distribution of Z_{ab}.
  - n: number of individuals per group;
  - J: total number of groups;
  - p: proportion of groups in the treatment condition;
  - ρ: intraclass correlation coefficient;
  - b: correlation between the mediator and the outcome.
- **Joint Significance Test**
  - Power(ab) = power(a)×power(b).
- **Cost Function**
  - $m = (nc_1^T + c_2^T)pJ + (nc_1 + c_2)(1-p)J$ with subscripts T denote costs in the treatment condition.
  - $J = \frac{m}{p(nc_1^T + c_2^T) + (1-p)(nc_1 + c_2)}$.
  - Then plug J in statistical power for optimization.

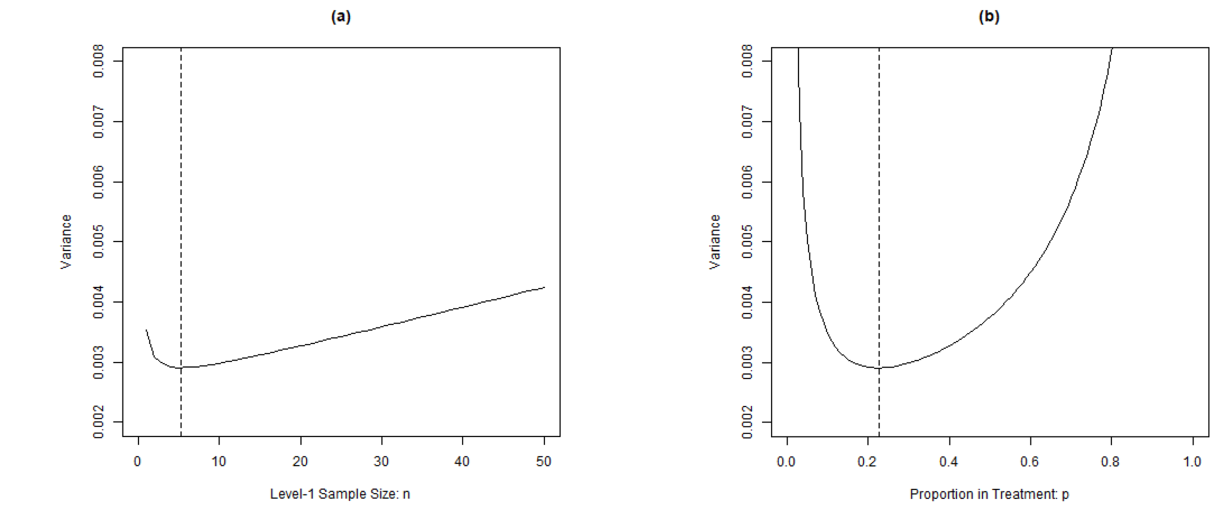


- **Optimization Method**
  - Sobel test: Minimize the variance of a mediation effect.
  - Joint significance test: Using ant colony optimization (ACO; Socha & Dorigo, 2008) to identify the optimal sample allocation.

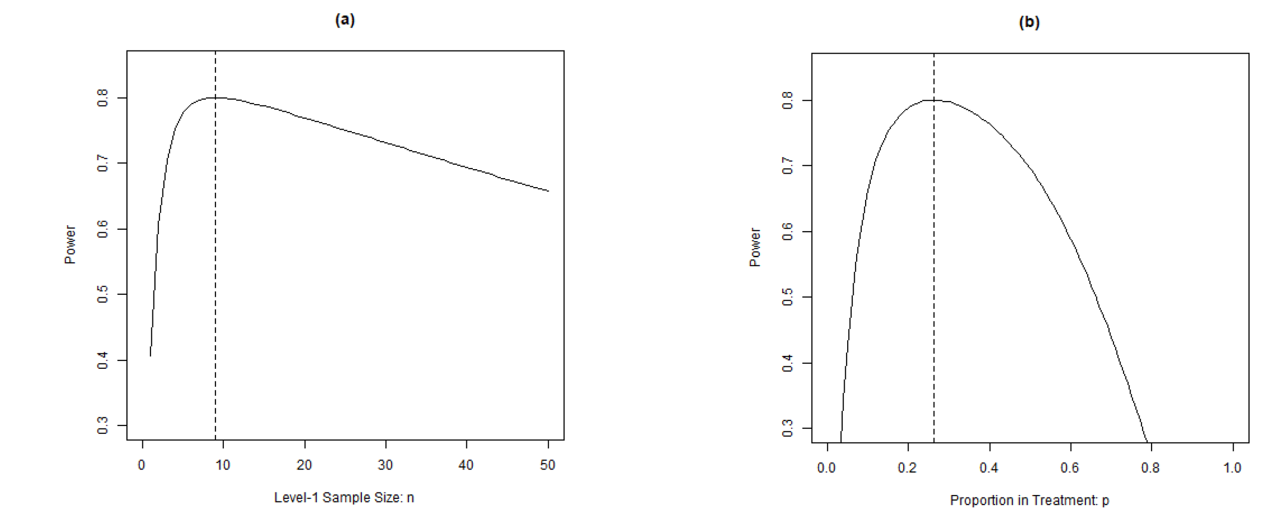
Illustration

- **An Example**
  - Effect size: a= .50, b= .30
  - ICC: $\rho = 0.20$
  - Costs: $c_1 = c_1^T = \$20, c_2 = \$500, c_2^T = \$5{,}000$.

Sobel Test (y-axis is σ^2_{ab})



Joint Significance Test (y-axis is power)



Note: dashed vertical lines are identified optimal sample allocation.

Conclusion

- The methods are implemented in the R package odr (Shen & Kelcey, 2021). They can also help identify constrained optimal sample allocation (e.g., with a constraint of $n = 20$).
- The conventional framework is a special case (with $p = 0.50$).

References

- Cox, K., & Kelcey, B. (2019). Optimal sample allocation in group-randomized mediation studies with a group-level mediator. The Journal of Experimental Education, 87(4), 616-640. https://doi.org/10.1080/00220973.2018.1496060
- Kelcey, B., & Shen, Z. (2020). Strategies for efficient experimental design in studies probing 2-1-1 mediation. The Journal of Experimental Education, 88(2), 311–334. https://doi.org/10.1080/00220973.2018.1533796
- Kelcey, B., Dong, N., Spybrook, J., & Shen, Z. (2017). Experimental power for indirect effects in group-randomized studies with group-level mediators. Multivariate Behavioral Research, 52(6), 699–719. https://doi.org/10.1080/00273171.2017.1356212
- Shen, Z., & Kelcey, B. (in press)). Optimal sampling ratios in three-level multisite experiments. Journal of Research on Educational Effectiveness.
- Shen, Z., & Kelcey, B. (in press). Optimal sample allocation in multisite randomized trials. The Journal of Experimental Education. https://doi.org/10.1080/00220973.2020.1830361
- Shen, Z., & Kelcey, B. (2020). Optimal sample allocation under unequal costs in cluster-randomized trials. Journal of Educational and Behavioral Statistics, 45(4), 446–474. https://doi.org/10.3102/1076998620912418
- Socha, K., & Dorigo, M. (2008). Ant colony optimization for continuous domains. European Journal of Operational Research, 185(3), 1155-1173. https://doi.org/10.1016/j.ejor.2006.06.046

UF | College of Education
UNIVERSITY of FLORIDA

Western Illinois University

Contact:
Feng Liu
Assistant Professor
Western Illinois University, Macomb, IL 61455
Email: F-Liu@wiu.edu