Contribution ID: **229**   Type: **Poster**

A Cautionary Note on Simulating Multilevel Data

Wednesday 23 July 2025 11:45 (15 minutes)

Poster

A Cautionary Note on Simulating Multilevel Data

Author

Diego Iglesias

Affiliation

Universidad Autónoma de Madrid

Abstract

Simulation studies are widely used in methodological research to assess the performance of statistical methods in scenarios relevant to the behavioral and social sciences. In some cases, the real-world scenario being simulated can be statistically complex. Specifically, in the case of multilevel data, where observations are nested within clusters, simulating level-2 variables and effects may not be straightforward. For example, when simulating random intercept variation, slopes, and level-2 predictors, a common approach is to sample observations from a distribution with a fixed mean and variance and then repeat the sampled values as many times as there are observations per cluster. However, this repetition alters the variance of the resulting variables, which no longer matches the population values previously specified. In this study, we illustrate the impact of this issue in two contexts: parameter estimation and power analysis in multilevel models. Through illustrative simulations, we show that when multilevel data are generated using this approach, having few clusters leads to simulated datasets in which the variance of intercepts, slopes, and level-2 predictors is 10% lower than the intended value. This discrepancy biases the results obtained when fitting statistical models to these simulated data, leading to an underestimation of the variance of intercepts and slopes, as well as the statistical power associated with level-2 effects. We present how the data generating process can be adjusted to correct this issue and discuss practical situations in which failing to account for it could have a greater impact.

Keywords

Multilevel Models, Simulation, Estimation, Power

Primary author: IGLESIAS, Diego (Universidad Autónoma de Madrid)

Co-authors: SORREL, Miguel A. (Universidad Autónoma de Madrid); OLMOS, Ricardo (Universidad Autónoma de Madrid)

Presenters: IGLESIAS, Diego (Universidad Autónoma de Madrid); SORREL, Miguel A. (Universidad Autónoma de Madrid); OLMOS, Ricardo (Universidad Autónoma de Madrid)

Session Classification: Poster Session 1

Track Classification: Statistical analyses: Statistical analyses