

Simulation and Power



Brad Verhulst

International Statistical Genetics Workshop

March 8th, 2024

Motivations for Simulating Data



- Model Testing
 - Can I generate a data to test whether my model works as advertised?
- Bias Testing
 - What happens if the data I generate doesn't match the model that I want to fit?
- Power Analysis
 - How many people would I need to have to be reasonably confident that I will accurately reject the null hypothesis?

Motivations for Simulating Data



Simulations for Model Testing

- Often used in methodological development.
- Does the model that I am building recover the “true” estimates
- What does the distribution of the parameters look like

Motivations for Simulating Data



Simulations for Bias Testing

- This allows you to test how violations in the assumptions will bias the parameter estimates.
- You can simulate models that cannot be identified (and therefore could not be fit)
- Only some parameters will be biased.
 - Knowing where the bias is allows you to think about how your model will be affected by unmodeled (and potentially unmodel-able) parameters

Motivations for Simulating Data



Simulations for Power Analysis

- Simulating data that corresponds with the way you think the world work will allow you to estimate how many people you will need to reasonably identify a significant effect.
- Two primary methods for power analysis in twin studies:
 - Monte Carlo Power analysis.
 - Simulate and evaluate the model. The proportion of times your parameter is significant is your power.
 - Non-Centrality Parameter approach:
 - Identify how much each observation contributes to the effect (on average).
 - The NCP scales linearly with sample size

The NCP approach is implemented for a variety of examples in this paper

Behav Genet (2017) 47:255–261
DOI 10.1007/s10519-016-9828-9

BRIEF COMMUNICATION

A Power Calculator for the Classical Twin Design

Brad Verhulst¹ 

This approach does not work if you want to include definition variables

Simulations in the Alt Script



General Simulations

- Linear Regression
- Latent Factor
- ACE

Simulating Bias

- ACE model with rAC and rAE

Simulations elsewhere in the workshop

- Simulations for the CPM and IPM model on Wednesday (Me)
- Simulations for the prs-rGE model (Dan)
- Simulations for Doc Model (Madhur)