Small Sample Solutions for SEM

Yves Rosseel[1*]

[1] Department of Data Analysis, Ghent University, Belgium
* Presenting author

Suggested talk duration: 45min (I could do 60min too, if there is time)

Summary (max. 500 words)
In this presentation, an overview will be given of old and recent solutions to handle small samples in the framework of structural equation modeling (SEM). A distinction is made between (1) solutions for estimation (and avoiding non-convergence issues), and (2) solutions for small-sample inference.
For estimation, we will advocate the use of a divide-and-conquer approach. The general (and old) idea is to break down the model into smaller pieces, estimate the parameters of each piece in turn, and finally combine these pieces again to get the final result. For example, in a SEM with many (measured) latent variables, one can estimate the parameters of each measurement model, one at a time. Once all the measurement parts have been estimated, we can hold the parameters fixed to these estimates, and only estimate the parameters of the structural part in a second step. Another approach is to first generate factor scores for each latent variable, and then (after a suitable transformation, known as the Croon correction) use these factor scores as if they are observed variables in a path analysis. Remarkably, both approaches are able to obtain consistent estimates. The end result is that we can estimate fairly large models, with a relatively small sample, and still get stable results in a frequentist framework.
For inference, we will discuss several methods to obtain unbiased standard errors, if we use a divide-and-conquer approach. As expected, the price to pay for using a multiple-step approach (instead of a single-step approach) is a modest loss of efficiency. But even if we use a classic SEM to estimate all the parameters in a single step, corrections are needed to get standard errors (and confidence intervals) that lead to better small-sample behavior. For the goodness-of-fit test statistic, we will discuss a Bartlett correction, including extensions to the non-normal case, and the incomplete data case.
Finally, if time permits, we will briefly discuss extensions of these small-sample solutions to the multilevel SEM setting.

Relevance to conference theme
SEM is widely used, but suffers from the perception that large samples are always needed. This presentation will debunk this myth, as well as the other myth that the only way out is to go Bayesian.

Keywords (max. 3)
SEM, factor score regression, small-sample corrections