

Powerful SEM: An Easy and Powerful Auxiliary for Structural Equation Modelling Analysis

Zhenwei Dai*

School of Population Medicine and Public Health, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China.

*Corresponding Author: Zhenwei Dai

Email: Huggedata123@163.com

Abstract

Powerful SEM is a free R package that addresses some important issues in structural equation modelling. It can calculate sample size and statistical power of structural equation modeling (SEM) more easily, and provide adjusted model fit indices (Bollen-Stine P correction), to make it more convenient to conduct SEM analysis. Currently, it has 2 versions: v0.1.0, and v0.2.0. The latest version (v0.2.0) has three functions: `semsample()`, `sempower()`, and `BSfit()`. This paper will introduce the functions of `powerfulSEM` to you.

V0.1.0

`PowerfulSEM` v0.1.0 is a free R package that makes calculating sample size and the power of structural equation modeling (SEM) more easy. The root mean square error of approximation (RMSEA) is recommended to be used for power analysis and sample size estimation in SEM [5]. Therefore, `powerfulSEM` uses the hypothesis-testing framework for RMSEA as a vehicle for defining a procedure for statistical power analysis and determination of minimum sample size for SEM [4].

Functions

`PowerfulSEM` 0.1.0 had two functions, namely `"semsample()"` and `"sempower()"`, respectively. `"semsample()"` function is used for calculating the minimum sample size requirement in SEM, and `"sempower()"` function is used for calculating the statistical power of SEM.

In `"semsample()"` function, you need to input 4 parameters: `"rmsea"`, `"df"`, `"a"`, and `"p"`, in which `"rmsea"` means alternative hypothesized RMSEA (RMSEAA); `"df"` means degrees of freedom of the SEM; `"a"` means alpha level, normally 0.05; and `"p"` means desired power, normally 0.8, 0.9, etc. For more information on this function, please input `help(semsample)` in R.

In `"sempower()"` function, you also need to input 4 parameters: `"a"`, `"df"`, `"s"`, and `"rmsea"`, where `"a"` means alpha level,

normally 0.05; `"df"` means degrees of freedom of the SEM; `"s"` means sample size; and `"rmsea"` means alternative hypothesized RMSEA (RMSEAA). For more information on this function, please input `help(sempower)` in R.

`PowerfulSEM` set the null hypothesized RMSEA (RMSEA0) as 0.05 by default. Previous research suggested that, for the test of close fit, the effect size is represented by the pair of values `RMSEA0 = 0.05` and `RMSEAA = 0.08`; for the test of not-close fit, the effect size is reflected by the pair of values `RMSEA0 = 0.05` and `RMSEAA = 0.01` [4]. That is, the parameter `"rmsea"` in both functions should be set as 0.01 or 0.08. The choice of result should be based on the conservative principle.

Examples of calculating sample size

Assuming we need to estimate the sample size of a model with degrees of freedom of 120, alpha of 0.05, and power of 0.8. For the test of close fit, we use `semsample(0.08, 120, 0.05, 0.8)`, and we can get the result: 116.4062. For the test of not-close fit, we use: `semsample(0.01, 120, 0.05, 0.8)`, and get the result: 159.7656. In this case, considering the conservative principle, we should use the result of a not-close fit of 160 as our minimum sample size instead of 116. However, assuming we need to estimate the sample size of a model with degrees of freedom of 400, alpha of 0.05, and power of 0.8. For the test of close fit, we use `semsample(0.08, 400, 0.05, 0.8)`, and we

get the result of 55.46875. For the test of not-close fit, we use `semsample(0.01, 400, 0.05, 0.8)`, and the result is 81.83594. While in covariance-based SEM, we know the sample size might not be lower than 200 [3]. Therefore, we may not use the result of `powerfulSEM` but 200 or other rules of thumb such as the “N:q rule” instead, to determine our minimum sample size [2]. Of note, the “`rmsea`” cannot be set as 0.05 to avoid non-convergence since the null hypothesized RMSEA (RMSEA0) has been set as 0.05 by default in `powerfulSEM`.

Examples of calculating power

Assuming we need to estimate the statistical power of a model with degrees of freedom of 70, alpha level of 0.05, and sample size of 200. For the test of close fit, we use: `sempower(0.05, 70, 200, 0.08)`, and get the result of 0.7427872. For the test of not-close fit, we use `sempower(0.05, 70, 200, 0.01)`, and get the result of 0.876987. Considering the conservative principle, we could only conclude that the power of the model is 0.743 instead of 0.877.

V0.2.0

Based on v0.1.0, `powerfulSEM` v0.2.0 added a new function named “`BSfit()`”, which could help to get the Bollen-Stine P correction model fit indices of SEM. This function could be used along with AMOS software since AMOS is a professional software for SEM analysis and provide the Bollen-Stine P correction function [1]. In “`BSfit()`” function, you need to input 6 parameters: “`BSC`”, “`IMC`”, “`DMDF`”, “`NPAR`”, “`IMDF`”, and “`S`”, where “`BSC`” means Bollen-Stine chi-square, “`IMC`” means independence model chi-square, “`DMDF`” means default model degree of freedom, “`NPAR`” means the number of default model parameters, “`IMDF`” means independence model degree of freedom, and “`S`” means sample size. `PowerfulSEM` 0.2.0 can provide the following adjusted model fit indices: Normed Chi-square (Chi-square/df), Goodness of Fit (GFI), Adjust Goodness of Fit (AGFI), Non-Normed Fit Index (NNFI)/Tucker-Lewis Index (TLI), Comparative Fit Index (CFI), Root Mean Square Error of Approximation (RMSEA), Akaike information criterion (AIC), Bayes information criterion (BIC), and Expected cross-validation index (ECVI). For more information on this function, please input `help(BSfit)` in R.

Example of BSfit()

Assuming our model had a Bollen-Stine chi-square of 159.3, an independence model chi-square of 2748.4, the number of default model parameters 126, an independence model degree of freedom of 153, and a sample size of 308, we can input: `BS(159.3,2748.4,126,45,153,308)`, the results are:

[1] “`CMIN/df: 1.26428571428571`”

[1] “`GFI: 0.942039004511716`”

[1] “`AGFI: 0.909838451462669`”

[1] “`TLI/NNFI: 0.984420238003501`”

[1] “`CFI: 0.987169607767589`”

[1] “`RMSEA: 0.0293405098600208`”

[1] “`AIC: 249.3`” [1] “`BIC: 417.154490233811`”

[1] “`ECVI: 0.809415584415584`”

More information

The “`powerfulSEM`” package files were of binary form. “`powerfulSEM` v0.1.0” has both a Windows version (`powerfulSEM_0.1.0.zip`) and a Mac version (`powerfulSEM_0.1.0.tgz`). “`powerfulSEM` v0.2.0” only has a Windows version (`powerfulSEM_0.2.0.zip`).

You can download the packages from ResearchGate.

DOI: 10.13140/RG.2.2.16085.70883.

References

1. Bollen KA, Stine RA. Bootstrapping Goodness-of-Fit Measures in Structural Equation Models [10.1177/0049124192021002004]. *Sociological Methods & Research*. 1992; 21: 205-229.
2. Jackson DL. Revisiting sample size and number of parameter estimates: Some support for the N : q hypothesis. *Structural Equation Modeling-a Multidisciplinary Journal*. 2003; 10: 128-141.
3. Kline, R. B. *Principles and Practice of Structural Equation Modeling* (Vol. Fourth Edition). Guilford Press. 2016.
4. MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*. 1996; 1: 130-149.
5. MacCallum RC, Hong S. Power Analysis in Covariance Structure Modeling Using GFI and AGFI. *Multivariate Behav Res*. 1997; 32: 193-210.