

Power Analysis for SEM: A Few Basics

Overall Model Fit

Much of the literature on power analysis in SEM has focused on estimating power of chi-square to detect false models in the population (MacCallum, Browne, & Sugawara, 1996) or to detect significant differences between nested models (Satorra & Saris, 1985; Saris & Satorra, 1993). The original Satorra and Saris approach involved specifying hypothesized values for the parameters in the model and then estimating effect size from the resulting fit function values relative to a population value or the fit function for an alternative model.

MacCallum and colleagues suggested a simplified approach using the Root Mean Square Error of Approximation (RMSEA), which can be simply computed from the chi-square (or fit function) and degrees of freedom.¹ They showed that a noncentrality parameter, λ , which estimates the magnitude of the effect (falsity of the model) in the population, can be computed from the RMSEA value.

$$\lambda = (N - 1)(df)(RMSEA)^2$$

The noncentrality parameter can be used with the chi-square distribution to estimate power. Power for overall model fit is the estimate that the probability for this noncentrality parameter is greater than the null value (which is equal to the df for a model that fits perfectly) using the chi-square distribution (e.g., in a spreadsheet like Excel). The value of this approach is that one can more readily describe the fit of a model in practical terms, compare values of the RMSEA for alternative models, and consider power to detect close fit rather than exact fit (Lee, Cai, & MacCallum, 2012). The same process can be used for relative fit, by comparing the difference in probabilities for noncentrality parameter when RMSEA is less than .05 vs. greater than .08, for instance (e.g., Hermida, Luchman, Nicolaides, & Wilcox, 2015). Alternatively, Preacher and Coffman (<http://www.quantpsy.org/rmse/rmse.htm>) as well as Timo Gnamb (<https://timo.gnamb.at/research/power-for-sem>) have created online calculators or code generators which can be used for estimating power for a given RMSEA and sample size or estimating sample size for a given RMSEA and an acceptable power, such as .8.

Although this is important work, as a researcher who has needed power analysis for grant proposals, I have not found this type of power analysis for SEM models that useful. I have typically been more concerned with demonstrating adequate power to detect significance for particular path coefficients rather than sufficient power for rejecting false models overall. The methods described above can, of course, be used to estimate power for particular parameters, if the significance of the parameter is tested using a nested chi-square test (e.g., comparing the fit of a model with a path set to 0 to one in which the path is freely estimated). Quantifying the expected effect size in terms of either chi-square difference values or difference in RMSEA values for this purpose is cumbersome, however.

Specific Parameter Tests

A simple method that I have used is to approximate power for path coefficients in a structural equation model is to conduct a power analysis for the significance of coefficients in a regression analysis (see the "Sample Size and Power for Regression" handout for my multiple regression and multivariate quantitative methods course). This approach most likely provides reasonable approximations if the estimates of the expected effect sizes (e.g., standardized coefficients) are appropriate, because the statistical test of the path in a path model is essentially the same (Hancock, 2013). Using this proxy approach depends on whether you can make sensible decisions that take into account whether the analyses that the effect sizes used regression or SEM, whether

¹ $RMSEA = \sqrt{(\chi^2 - df) / [(df)(N - 1)]}$

the variables involved were measured variables or latent variables, how reliable the measured variables were, and other specifics about the model. In the end, this is a less than perfect approach.

Monte Carlo Approach

A more precise method is the one proposed by Muthén and Muthén (2002), which uses the MONTECARLO procedure in Mplus designed for simulations to conduct power analyses. The Monte Carlo approach uses simulated data, taking repeated samples under a particular population model specified with certain parameter values. The model is then tested in each of the multiple random samples that are drawn (replications) and the proportion of significant results estimates the power (under the condition that H_a is true in the population, power is equal to .8 when 80% of the samples find a significant result for the parameter). This approach to power estimation is perhaps the most flexible because of the ability to incorporate measurement models, tailor the power analysis to specific estimation approaches, take into account nonnormality, and estimate power under various missing data conditions (Davey & Salva, 2010).

The MONTECARLO feature in Mplus makes the model specification process relatively simple.² Below, I provide a very simple analysis example in Mplus, with a latent variable predicting a measured outcome. The example could be varied to use different loading values, adjust the effect for unreliability of the measured dependent (e.g., by setting up a single indicator latent variable with measure residual variance equal to $1 - \text{reliability} \times \text{variance}$), or incorporate covariates.

I also illustrate with an example described by Jobst and colleagues (2023) using the semPower package in R that can be used withlavaan code. The semPower (Moshagen & Erdfelder, 2016) package uses a likelihood ratio testing approach to power for overall fit or single or multiple parameters.

The general approach is to simulate data, estimating power for various samples sizes and effect sizes. As with any a priori power analysis, you can base the estimates on specific effect sizes found in the related literature or use a range of effect sizes based on conventions. It is possible to work with unstandardized values from prior research where the variances are known, but the process can also be simplified by using standardized values and working with general effect size standards (Cohen, 1988). In the example below, I use the .1, .3, and .5 values associated with Cohen's conventions for small, medium, and large effects for my standardized path estimate.

Further examples for nonnormal data, growth curve models, and more are provide in the Muthén and Muthén (2002) paper, at the Mplus website (<https://www.statmodel.com/power.shtml>), and at Paul Tremblay's website (<http://publish.uwo.ca/~ptrembla/lecture12-fall2013.pdf>).

Mplus Monte Carlo Example

A simple example below illustrates how the Mplus MONTECARLO feature can be used to estimate power for different models. This is done empirically by generating many samples of a certain size for specific models with parameters set to desired values. The sample sizes and the parameter values can be varied as I did for multiple variations of the model below. I used complete normal data, but nonnormal data, different estimators, and missing data could be incorporated. Fit information and parameter estimates can be examined. Below, the % sig coeff column gives the percentage of samples that had a significant parameter estimate (for the ON statement). The desired power is usually .8 and this particular run with $N = 200$ and effect size equal to .3 had power equal to .985.

```
title: simple illustration of power analysis with  
normal complete data (see also Muthen & Muthen, 2012);
```

² See also the simsem package in R and Davey and Savla (2010) for a similar approach with Amos software

```
montecarlo:
names = y1 x1-x4; !one outcome and one latent variable;
nobservations = 200; !choose desired sample size (vary);
nreps = 1000; !use at least 500 samples for stability of estimates;
seed = 060418; !can reproduce the data if same seed is used;

analysis: estimator = ml;
model=nomeanstructure; information=expected;
model population:
etal by x1-x4*.8; !gives standardized loadings of .8 here;
etal@1; !set factor variance to one (standardize);
x1-x4*.36; !using 1-loading squared for measurement residuals;

! can vary particular standardized parameter value;
! such as .1, .3, and .5 based on Cohen's sm, med, lg;
y1 on etal*.3;
y1@.91; !set disturbance to 1 - sqrt(beta);

model: !generally has same specifications as above;
etal by x1-x4*.8; !gives standardized loadings of .8 here;
etal@1; !set factor variance to one (standardize);
x1-x4*.36; !using 1-loading squared for measurement residuals;

! can vary particular standardized parameter value;
! such as .1, .3, and .5 based on Cohen's sm, med, lg;
y1 on etal*.3;
y1@.91; !set disturbance to 1 - sqrt(beta);

output: tech9;
```

Output (some omitted)

INPUT READING TERMINATED NORMALLY
simple illustration of power analysis with
normal complete data (see also Muthen & Muthen, 2012);

SUMMARY OF ANALYSIS

Number of groups	1
Number of observations	200
Number of replications	
Requested	1000
Completed	1000
Value of seed	60418
Number of dependent variables	5
Number of independent variables	0
Number of continuous latent variables	1

Estimator	ML
Information matrix	EXPECTED
Maximum number of iterations	1000
Convergence criterion	0.500D-04
Maximum number of steepest descent iterations	20

Chi-Square Test of Model Fit	
Degrees of freedom	6
Mean	6.103
Std Dev	3.480
Number of successful computations	1000

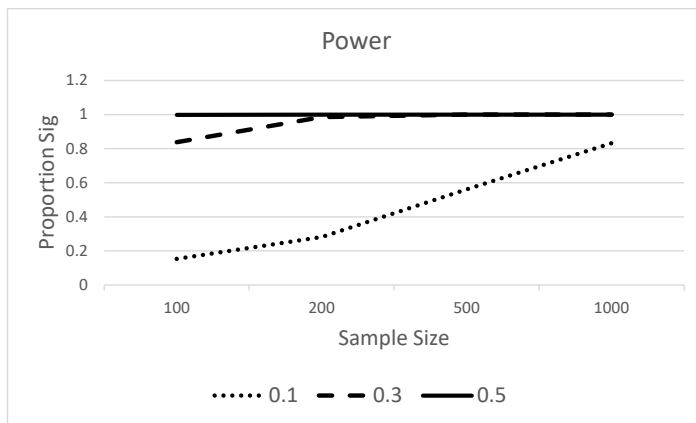
SRMR (Standardized Root Mean Square Residual)

Mean	0.030
Std Dev	0.013
Number of successful computations	1000

MODEL RESULTS

		Population	ESTIMATES		S. E.	M. S. E.	95% Cover	% Sig Coeff
			Average	Std. Dev.	Average			
ETA1	BY							
X1		0.800	0.7960	0.0629	0.0613	0.0040	0.947	1.000
X2		0.800	0.7972	0.0623	0.0614	0.0039	0.951	1.000
X3		0.800	0.7947	0.0642	0.0613	0.0041	0.937	1.000
X4		0.800	0.7953	0.0657	0.0615	0.0043	0.922	1.000
Y1	ON							
ETA1		0.300	0.2989	0.0753	0.0735	0.0057	0.946	0.985
Variances								
ETA1		1.000	1.0000	0.0000	0.0000	0.0000	1.000	0.000
Residual Variances								
Y1		0.910	0.9100	0.0000	0.0000	0.0000	1.000	0.000
X1		0.360	0.3538	0.0462	0.0477	0.0022	0.939	1.000
X2		0.360	0.3552	0.0496	0.0479	0.0025	0.934	1.000
X3		0.360	0.3550	0.0484	0.0477	0.0024	0.937	1.000
X4		0.360	0.3580	0.0499	0.0480	0.0025	0.942	1.000

The results above are from one run, in which N = 200 and the effect size equaled .3. The figure below was generated in Excel using the power estimates I obtained for all of the Monte Carlo runs using all the sample sizes (N=100, 200, 500, and 1000) and all of the effect sizes ($\beta = .1, .3, \text{ and } .5$).



R semPower Example

This example runs one of the several *a priori* power examples described by Jobst and colleagues (2023). The semPower package (Moshagen & Erdfelder, 2016) uses a likelihood ratio (nested model comparison) approach to power. A population model is constructed with the parameter values that are expected based on prior research, theory, or a range of possible arbitrary values. The example is their “local hypothesis” is an estimate of the sample size needed to detect significance one path from intelligence to financial well-being for power equal to .80. (note that the sample.nobs = 1000 is only an arbitrary value that is not related to the sample size estimate here and does not need to be changed).

```
> library(lavaan)
> library(semPower)
>
> #parameter values expected in the planned study are inserted into the hypothesized model
> #test of power to detect significance for path of intelligence predicting financial well-being

> population.model = '
+ # define factors and set all loadings to .7
+ intelligence =~ .7*x1+ .7*x2+ .7*x3
+ social.status =~ .7*x4+ .7*x5+ .7*x6
+ well.being =~ .7*y1+ .7*y2+ .7*y3
+ # set all item residual variances to .51 (=1-.7*.7)
+ x1~~ .51*x1
```

```

+ x2~~ .51*x2
+ x3~~ .51*x3
+ x4~~ .51*x4
+ x5~~ .51*x5
+ x6~~ .51*x6
+ y1~~ .51*y1
+ y2~~ .51*y2
+ y3~~ .51*y3
+ # set factor variances to 1
+ intelligence~~ 1*intelligence
+ social.status~~ 1*social.status
+ # define orthogonal exogenous factors
+ intelligence~~ 0*social.status
+ # define regression relationship
+ well.being ~ .3*intelligence + .6*social.status
+ # define residual variance of well.being (=1-(.3^2+.6^2))
+ well.being~~ .55*well.being'
>
> #then model is specified without parameter values to estimate power
> hypothesized.model = '
+ intelligence =~ x1+ x2+ x3
+ social.status =~ x4+ x5+ x6
+ well.being =~ y1+ y2+ y3
+ # fix slope of intelligence to zero
+ well.being ~ 0*intelligence + social.status'
> cov.population.model = fitted(sem(population.model))$cov
> fit = sem(hypothesized.model, sample.cov = cov.population.model,
+         sample.nobs = 1000, sample.cov.rescale = FALSE)
>
>
> cov.hypothesized.model = fitted(fit)$cov
> apriori = semPower.apriori(SigmaHat = cov.hypothesized.model,
+   Sigma = cov.population.model, alpha = .05, power = .80, df = 1)
> summary(apriori)

```

semPower: A priori power analysis

F0	0.062066
RMSEA	0.249130
Mc	0.969444
df	1
Required Num Observations	128
Critical Chi-Square	3.841459
NCP	7.882380
Alpha	0.050000
Beta	0.198331
Power (1 - Beta)	0.801669
Implied Alpha/Beta Ratio	0.252104

Results include model fit estimates for chi-square as well as RMSEA and McDonald's NCI for relative fit which can be of use for overall model fit. But these are not of interest in this particular power analysis. The sample size needed for power equal to .80 is listed under "Required num Observations" and indicates 128 is needed. One could then alter the parameter values in the population model to test different effect sizes to obtain a range of sample sizes depending on the effect size the way I did above with Mplus.

Other Software

The Saris and Satorra methods (Satorra & Saris, 1985; Saris & Satorra, 1993) for power analyses can be conducted with the semTools package in R. A shiny app in R developed by Jak and colleagues (Jak et al., 2021) implements the approach by MacCallum and colleagues. There is an online calculator from Daniel Soper, but there is not much detail on how these values are computed, so I would do a little more investigation before using or citing this resource <https://www.danielsoper.com/statcalc/calculator.aspx?id=89>.

References and Further Reading

- Davey, A., & Savla, J. (2010). *Statistical power analysis with missing data: A structural equation modeling approach*. NY: Routledge/Taylor & Francis.
- Hancock, G.R. (2013). Power analysis in covariance structural modeling. In G.R. Hancock & R.O. Mueller (Eds.), *Structural equation modeling: A second course, 2nd Edition* (pp. 69-114). Charlotte, NC: Information Age Publishing.
- Hermida, R., Luchman, J. N., Nicolaidis, V., & Wilcox, C. (2015). The issue of statistical power for overall model fit in evaluating structural equation models. *Computational Methods in Social Sciences*, 3, 25.

- Jak, S., Jorgensen, T. D., Verdam, M. G., Oort, F. J., & Elffers, L. (2021). Analytical power calculations for structural equation modeling: A tutorial and Shiny app. *Behavior Research Methods*, 53, 1385-1406.
- Jobst, L. J., Bader, M., & Moshagen, M. (2023). A tutorial on assessing statistical power and determining sample size for structural equation models. *Psychological Methods*, 28(1), 207-221.
- Lee, T., Cai, L., & MacCallum, R.C. (2012). Power analysis for tests of structural equation modeling. In Hoyle, R. H. (Ed.). *Handbook of structural equation modeling* (pp. 181-194). New York: Guilford Press.
- Maccallum, R. C., Browne, M. W., & Sugawara, H. H. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, 1(2), 130-149.
- Moshagen, M., & Erdfelder, E. (2016). A new strategy for testing structural equation models. *Structural Equation Modeling*, 23(1), 54-60.
- Muthén, L.K. and Muthén, B.O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling*, 4, 599-620.
- Saris, W. E., & Satorra, A. (1993). Power evaluations in structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 181-204). Newbury Park, CA: Sage.
- Satorra, A., & Saris, W. E. (1985). The power of the likelihood ratio test in covariance structure analysis. *Psychometrika*, 50, 83-90.
- Yi, F., & Hancock, G.R. Power analysis with a structural equation modeling framework. In R.H. Hoyle, *Handbook of structural equation modeling, second edition* (pp. 163-183). Guilford.