Psychology 454: Latent Variable Modeling How do you know if a model works?

William Revelle

Department of Psychology
Northwestern University
Evanston, Illinois USA

November, 2012

Outline

A number of tests of fit taken from Marsh et al. (2005)

1. Marsh, Hau & Grayson (2005) lists 40 different proposed measures of goodness of fit
2. Measures of absolute fit
   - I_o = index of fit for original or baseline model
   - I_t = index of fit for target or "true" model
3. Measures of incremental fit Type I
   - $\frac{|I_t - I_o|}{Max(I_o, I_t)}$ which is either
     - $\frac{I_o - I_t}{I_o}$
     - or $\frac{I_t - I_o}{I_t}$
4. Measures of incremental fit Type II
   - $\frac{|I_t - I_o|}{E(I_t - I_o)}$ which is either
     - $\frac{I_o - I_t}{I_o - E(I_t)}$
     - or $\frac{I_t - I_o}{E(I_t) - I_o}$

Fit functions from Jöreskog

1. Ordinary least squares $F = \frac{1}{2} tr(S - \Sigma)^2$
   - The squared difference between the observed (S) and model (Σ) covariance matrices
   - tr means trace of the sum of the diagonal values of the matrix of squared deviations
2. Generalized least squares $F = \frac{1}{2} tr(I - S^{-1}\Sigma)^2$
   - I is the identity matrix
   - if the model = data, then $S^{-1}\Sigma$ should be I
   - weight the fit by the inverse of the observed covariances
3. Maximum Likelihood $F = log|\Sigma| + tr(S\Sigma^{-1}) - log|S| - p$
   - weight the fit by the inverse of the modeled covariance
   - p is the number of variables
   - tr (I) = p, and thus the ML should be 0 if the model fits the data

Fit-function based indices

1. Fit Function Minimum fit function (FF)
   - $FF = \frac{\chi^2}{(N-1)}$
2. Likelihood ratio $LHR = e^{-1/2FF}$
3. χ^2 (minimum fit function chi square)
   - $\chi^2 = tr(\Sigma^{-1}S - I) - log|\Sigma^{-1}S| = (N-1)FF$
4. $p(\chi^2)$ probability of observing a χ^2 this larger or larger given that the model fits
5. $\frac{\chi^2}{df}$ has expected value of 1

Non-centrality based indices

1. Dk: Rescaled noncentrality paramter (McDonald & Marsh, 1990)
   - $Dk = FF - df/(N-1) = \frac{\chi^2 - df}{N-1}$
2. PDF (population discrepancy function = DK normed to be non-negative)
   - $PDF = max(\frac{\chi^2 - df}{N-1}, 0)$
3. Mc: Measure of centrality (CENTRA, MacDonald Fit Index (MFI)
   - $Mc = e^{\frac{-(\chi^2 - df)}{2(N-1)}}$
4. Non-centrality parameter
   - $NCP = \chi^2 - df$

Error of approximation indices

How large are the residuals, estimated several different ways

1. RMSEA (root mean square error of approximation)
   - $RMSEA = \sqrt{PDF/df} = \sqrt{\frac{max(\frac{\chi^2 - df}{N-1}, 0)}{df}}$
   - based upon the non-central χ^2 distribution to find the error of fit
2. MSEA (mean square error of approximation – unnormed version of RMSEA)
   - $MSEA = \frac{Dk}{df} = \frac{\chi^2 - df}{(N-1)df}$
3. RMSEAP (root mean square error of approximation of close fit)
   - RMSEA $< .05$
4. RMR Root mean square residual (or, if S and Σ are standardized, the SRMR). Just
   - square root of the average squared residual
   - $\sqrt{\frac{2\sum (S-\Sigma)^2}{p*(p+1)}}$

Information indices

Compare the information of a model with the number of parameters used for the model. These allow for comparisons of different models with different degrees of freedom.

1. AIC (Akaike Information Criterion for a model penalizes for using up df)
   - $AIC = \chi^2 + p * (p + 1) - 2df = \chi^2 + 2K$
   - where $K = \frac{p*(p+1)}{2} - df$
2. Baysian Information Criterion
   - $-2Log(L) + plog(N) = \chi^2 - Klog(N(.5(p(p + 1))$

Goodness of fit indices

1. GFI from LISREL
   - $GFI = 1 - \frac{tr(\Sigma^{-1}S - I)^2}{tr(\Sigma^{-1}S)^2}$
2. Adjusted Goodness of Fit (Lisrel)
   - $AGFI = 1 - \frac{p(p+1)}{2df}(1 - GFI)$
3. Unbiased GFI (from Steiger)
   - $\text{GFI} = \frac{p}{2\frac{(\chi^2 - df)}{(N-1)} + p}$

Comparing solutions to solutions

1. Incremental fit indices without correction for model complexity
   - RNI (relative non-centrality) McDonald and Marsh
   - CFI Comparative fit index (normed version of RNI) Bentler
   - Normed Fit index (Bentler and Bonett)
2. Incremental fit indices correcting for model complexity
   - Tucker - Lewis Index
   - Normed Tucker Lewis
   - Incremental Fit index
   - Relative Fit Index
3. Parsimony indices

Incremental fit indices without correction for model complexity

1. RNI (relative non-centrality) McDonald and Marsh
   - $RNI = 1 - \frac{Dk_t}{Dk_n}$
   - where $DK = \frac{\chi^2 - df}{N-1}$ for either the null or the tested model
2. CFI Comparative fit index (normed version of RNI) Bentler
   - Just norm the RNI to be greater than 0.
   - $CFI = 1 - \frac{MAX(NCP_t,0)}{MAX(NCP_n,0)}$
3. Normed Fit index (Bentler and Bonett)

Fitting functions from Loehlin

1. Let S be the "strung out" data matrix
2. Let Σ be the "strung out" model matrix
3. $Fit = (S - \Sigma)' W^{-1} (S - \Sigma)$
4. Where W =
   - Ordinary Least Squares $W = I$
   - Generalized Least Squares $W = SS'$
   - Maximum likelihood $W = \Sigma\Sigma'$

Practical advice

1. Taken from Kenny
   - http://davidkenny.net/cf/fit.htm
2. Bentler and Bonnet Normed Fit Index
   - $\frac{\chi^2_{Null} - \chi^2_{Model}}{\chi^2_{Null}}$
   - Between .90 and .95 is acceptable
   - $>$ .95 is "good"
3. RMSEA
   - if $\chi^2 < df$, then RMSEA $= 0$
   - "good" models have $RMSEA < .05$
   - "poor" models have$RMSEA > .10$
4. p of close fit
   - Null hypothesis is that RMSEA is .05
   - test if $RMSEA > .05$
   - Claim good fit if $p(RMSEA > .05) > .05$

Fits and sample size

1. See associated simulation results
   -
2.
   -

Considering rules of thumb and fit

1. Fit functions have distributions and thus are susceptible to problems of type I and type II error.
   - Compare the fits for correct model as well as those for a simple incorrect
2. Should we just use chi square and reject models that don't fit, or should we reason about why they don't fit

What does it mean if the model does not fit

1. Model is wrong
2. Measurement is wrong
3. Structure is wrong
4. Assumptions are wrong
5. at least one of above, but which one?

Specification & Respecification

1. Is the measurement model consistent
   - revise it
     - evaluate loadings
     - evaluate error variances
     - more or fewer factors
     - correlated errors?
2. Structural model:
   - adjust paths
   - drop paths
   - add paths
3. Equivalent models
   - What models are equivalent
   - Do they make equally good sense

Conclusion

1. Latent variable models are a powerful theoretical aid but do not replace theory
2. Nor do latent modeling algorithms replace the need for good scale development
3. Latent variable models are a supplement to the conventional regression models of observed scores.
4. Other latent models (not considered) include
   - item Response Theory
   - Latent Class Analysis
   - Latent Growth Curve analysis

Marsh, H. W., Hau, K.-T., & Grayson, D. (2005). Goodness of Fit in Structural Equation Models. In A. Maydeu-Olivares & J. J. McArdle (Eds.), *Contemporary Psychometrics* chapter 10, (pp. 275–340). New York: Routledge.

McDonald, R. & Marsh, H. (1990). Choosing a multivariate model: Noncentrality and goodness of fit. *Psychological Bulletin*, *107*(2), 247–255.