

Nested Models, Model Modifications, and Correlated Errors

Nested Models

It is often recommended that researchers compare the fit of their model to alternative models. A chi-square difference test can be conducted using chi-square values and degrees of freedom from any two *nested models*. A *nested model* is a model that uses the same variables (and cases!) as another model but specifies at least one additional parameter to be estimated. The model with fewer restrictions or more free parameters (i.e., fewer degrees of freedom), which could be called a *reduced model*, is nested within the more restricted model, which could be called the *full model*. There are several ways this can be done, such as comparing two different models with the same variables but one more parameters added, comparing a one-factor to a two-factor model with the same variables, imposing equality constraints (two parameters are set to be equal), or setting parameters to a particular value compared with allow them to be freely estimated.

The chi-square test is simply the difference between the full model and the reduced model, using the difference in degrees of freedom as the degrees of freedom for the test. Tests are typically computed by hand and compared to a standard chi-square table to determine significance.

$$\chi^2_{diff} = \chi^2_{full} - \chi^2_{reduced}$$
$$df_{diff} = df_{full} - df_{reduced}$$

Not all nested models are obvious and not all obvious model comparisons are nested, so researchers should exercise some caution (Rigdon, 1999). Tests are also sensitive to sample size. With large sample sizes, even small magnitude differences in fit will be significant. Researchers must decide whether differences are of practical magnitude or not.

Modification Indices

Modification indices, which involve a series of simple nested model tests, can be requested from most computer packages. A modification index is a one degree of freedom chi-square test of the addition of a new parameter or the deletion of a parameter. Each modification index represents the change in the overall chi-square for the fit of the model if that particular parameter is changed. Thus, a significant chi-square value (greater than 3.84, because this is the chi-square critical value for $df = 1$) will significantly improve the fit of the model. This type of chi-square difference test, known as a likelihood ratio test, is the type of modification index used in Lisrel, Mplus, and Amos. EQS provides two slightly different types of modification indices. EQS uses similar asymptotically equivalent tests, called the Lagrange multiplier for adding parameters to the model and Wald tests for eliminating parameters from the model. A similar test called the Score test is occasionally mentioned. Some packages also will print the expected change in a parameter, representing what the unstandardized or standardized value of the added path would be. Another way to roughly gauge the size of the modification index is to compare the chi-square difference relative to the overall chi-square value (i.e., by what percentage will the chi-square be improved by the addition of the new parameter to the model).

I suggest that you make one modification at a time, because each change may affect other parameters in the model and may, therefore, alter the relative importance of modification indices in a revised model. Keep in mind that many modifications to a model are considered by reviewers to be exploratory. Modifications can lead to development of incorrect models (MacCallum, 1986). Because a large number of tests are conducted when modification indices are requested, the group of tests are subject to familwise error (alpha inflation). I suggest that when changes are made they are reported and that you make sure they can be theoretically justified.

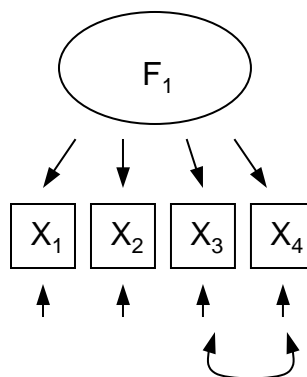
Chi-square Difference Tests with Non-normal Data

With non-normal data, it is not appropriate to use a simple difference in chi-square values, and a weighted correction is required (Satorra & Bentler, 1999). Mplus allows the user to specify difference tests

(DIFFTEST) and saves out information from the first model and compares it to the second model. I have an Excel sheet that I use to compute does the correct difference test.

Correlated Errors

One nested model modification that is fairly commonly mentioned by researchers, is an addition to the measurement model that allows measurement errors to be correlated.¹ Typically, researchers begin testing a confirmatory factor model by assuming that measurement errors are independent of one another. This assumption implies that the variance of a particular item not caused by the factor has a source that is unique to that particular variable. This assumption is not always valid, so researchers may incorporate correlated errors in a factor model. Inclusion of the correlation will decrease the loadings for the items involved if the correlation is positive. The following figure graphically illustrates a correlation between errors for items 3 and 4.



Correlated errors are often due to similar item wording or content (e.g., "I feel blue whenever my spouse is around" and "I feel sad whenever my spouse is around"). The basis for including a correlation between measurement errors can be data driven (e.g., suggested by modification indices) or theoretically driven. A common, and theoretically motivated, application is in longitudinal research in which the measurement error for the same variable is correlated over time. Many additional correlated errors will likely appear exploratory to reviewers, but my experience is that one or two theoretically justified correlated errors are not met with too much resistance. Correlated errors also may be used to account for method effects (e.g., telephone interviews and face-to-face interviews, observation vs. self-report). One common type of method effect that you may encounter has to do with items that are positively or negatively worded (e.g., "I find statistics dreadful" vs. "I find statistics most interesting"; see Tomás & Oliver, 1999, for detailed discussion). In such cases, model misfit is indicated for correlations among measurement errors for items worded in the same direction. Modifications to the model involving correlated errors are subject to the same precautions as mentioned above.

References

- Rigdon, E.E. (1999). Using the Friedman method of ranks for model comparison in structural equation modeling. *Structural equation modeling*, 6(3), 219-232
- MacCallum, R. (1986). Specification searches in covariance structure modeling. *Psychological Bulletin*, 100, 107-120.
- Satorra, S., & Bentler, P.M. (1999). *A scaled difference chi-square test statistic for moment structure analysis*. Unpublished technical report, Universitat Pompeu Fabra, Barcelona, Spain.
- Tomás, J. M., & Oliver, A. (1999). Rosenberg's self-esteem scale: Two factors or method effects. *Structural Equation Modeling*, 6, 84-98.

¹ As elsewhere, I am overly general in that I describe errors associated with indicators as "measurement error" when they really represent other sources of error, something better termed a *unique error* or a *unique component*. Essentially the unique error is any source of variation in the indicator that is not accounted for by the latent variable and may include systematic or meaningful causes. In other words, we should think about unique errors as any variation in the indicator due to unknown sources (i.e., not accounted for by the latent variable).