Simple Second Order Chi-Square Correction

Tihomir Asparouhov and Bengt Muthén

May 3, 2010

1   Introduction

In this note we describe the second order correction for the chi-square statistic implemented in Mplus Version 6 with the estimators WLSMV/ULSMV and MLMV. Prior to Version 6 the second order correction in Mplus for the chi-square statistics is based on a Satterthwaite (1941) type correction, see also Satorra and Bentler(1994) and Muthén et al. (1997). This type of correction scales appropriately the chi-square statistic and estimates the optimal degrees of freedom to obtain the best approximation possible. Here we introduce an alternative to the Satterthwaite style correction which has the advantage that it does not need to estimate the degrees of freedom and instead uses the usual degrees of freedom, which is simply the difference between the number of parameters in the two models.

In Mplus Version 6 both the new second order correction as well as the Satterthwaite second order correction can now be performed. The new second order correction is now the Mplus default and the Satterthwaite correction can be obtained using the command *Satterthwaite=ON*.

In Section 2 we describe the three types of scaled test statistics available in Mplus Version 6 with the weighted least squares estimation. In Section 3 we present the results of a simulation study to evaluate the performance of the three scaled test statistics. In Section 4 we also include a brief discussions on how the new second order correction is constructed for the Mplus DIFFTEST command used to test two nested models. In Section 5 we discuss the construction of modification indices for scaled chi-square statistics.

2   Scaled Chi-Square Statistics

Consider the weighted least squares estimation in Mplus and denote by F the fit function used in the estimation. Denote by T the minimal value of F obtained in the estimation. The value T is used to construct the test statistics. The asymptotic distribution of T is not a chi-square distribution. This distribution can be represented as a mixture (weighted sum) of chi-square distributions of 1 degree of freedom

$$\sum_{d=1}^{D} w_i \chi^2(1) \tag{1}$$

where the weights w_i are the eigenvalues of a certain matrix M, see Satorra and Bentler (2001). In the above expression D is the difference between the number of parameters in the unrestricted model and the number of parameters in the estimated model.

Since it is not as straight forward to compute P-values for distribution (1) we use approximations based on the chi-square distribution. The first order approximation is to use instead of T the statistic

$$T_1 = T\frac{D}{\sum w_i} = T\frac{D}{Tr(M)}$$

and assume that T_1 has a chi-square distribution with D degrees of freedom. By definition $E(T_1) = D$, i.e., using the scale factor $D/\sum w_i$ results in the fact that the distribution of T_1 and the chi-square distribution used as an approximation have the same mean.

The second order approximation by Satterthwaite (1941) is designed to match not only the mean of the test statistic distribution with the mean of the chi-square distribution but also the variance. This is done by estimating the degrees of freedom $\hat{D}$, see Muthén (1998-2004) and Satorra and Bentler (1994), and using the test statistic

$$T_2 = T\frac{\hat{D}}{\sum w_i} = T\frac{\hat{D}}{Tr(M)} \tag{2}$$

where the degrees of freedom $\hat{D}$ is estimated as the integer closest to

$$\frac{(Tr(M))^2}{Tr(M^2)}. \tag{3}$$

Note now that

$$E(T_2) = \frac{\hat{D}}{\sum w_i}E(T) = \frac{\hat{D}}{\sum w_i}\sum w_i = \hat{D}$$

$$Var(T_2) = \frac{\hat{D}^2}{(\sum w_i)^2}Var(T) = 2\frac{\hat{D}^2}{(Tr(M))^2}\sum w_i^2 = 2\frac{\hat{D}^2}{(Tr(M))^2}Tr(M) \approx 2\hat{D}.$$

Thus the mean and the variance of T_2 match those of the chi-square distribution with $\hat{D}$ degrees of freedom used as an approximation.

In Muthén et al. (1997) it is shown in a simulation study that the second order correction T_2 performs much better than the first order correction

T_1. Here we propose a new second order correction statistic T_3 which has approximately a chi-square distribution with D degrees of freedom, i.e., the degrees of freedom is unchanged. We will also show with simulations that the new second order approximation statistic T_3 performs as well as T_2.

To match the mean and the variance of the chi-square distribution with D degrees of freedom in the construction of T_3 we use not just a scale factor but also a shift parameter, that is

$$T_3 = aT + b$$

where a and b are chosen so that $E(T_3) = D$ and $Var(T_3) = 2D$. Essentially to obtain a and b we solve the following system of linear equations

$$D = E(T_3) = aE(T) + b = a\ Tr(M) + b$$

$$2D = Var(T_3) = a^2 var(T) = 2a^2\ Tr(M^2).$$

The solution of that system is given by

$$a = \sqrt{\frac{D}{Tr(M^2)}}$$

and

$$b = D - \sqrt{\frac{D\ Tr(M)^2}{Tr(M^2)}}.$$

Thus the proposed T_3 is

$$T_3 = T\sqrt{\frac{D}{Tr(M^2)}} + D - \sqrt{\frac{D\ Tr(M)^2}{Tr(M^2)}}. \tag{4}$$

Because of the above construction T_3 has the same mean and variance as the chi-square distribution with D degrees of freedom.

3   Simulation Study

We replicate the simulation study conducted in Muthén et al. (1997) and we include the new test statistics T_3. The goal of the simulation study is to evaluate the type I error of the test statistics. The data is generated according

Table 1: Rejection rates under correct null hypothesis

Estimator	WLSM	WLSMV	WLSMV
Satterthwaite	-	ON	OFF
Test Statistic	T_1	T_2	T_3
Table 1	16%	4.8%	5.2%
Table 3	25.2%	9.2%	10%
Table 5	17.6%	7.6%	8.6%
Table 7	14.4%	7.2%	7%
Table 9	10.2%	6%	6.4%
Table 11	10%	6.2%	6.6%
Table 13	23%	10.6%	11.2%
Table 15	12.6%	6%	6.8%

to a model and analyzed according to the same model. Each test statistic is computed and the model is rejected if the test statistic is larger than the 95% percentile of the corresponding chi-square distribution, i.e., if the P-value is less than 5%. In Table 1 we report the percentage of rejected models, i.e., the rejection rates for each of the three test statistics. Rejection rates near 5% indicate correct performance and higher values indicate inflated type I error which is undesirable. All of the simulation studies from Muthén et al. (1997) are conducted. A detailed account for these simulations can be obtained in Muthén et al. (1997). In the table below we use the Table numbering from Muthén et al. (1997) to refer to a particular Montecarlo setup, i.e., Table 1 refers to the simulation study presented in Muthén et al. (1997) Table 1.

The results in Table 1 show that T_1 indeed has inflated rejection rates and that both T_2 and T_3 perform substantially better than T_1. The difference between the rejection rates of T_2 and T_3 is negligible and both statistics perform quite well in all cases.

4 T_3 Style Second Order Correction for Chi-Square Difference Testing

Consider the case when there are two nested models A and B and we need to test the more restricted model A against the less restricted model B using a second order adjusted test statistic. In Mplus this is done using the DIFFTEST command, see Asparouhov and Muthén (2006). Suppose that the fit functions for the two models are T_A and T_B. We use

$$T = T_A - T_B$$

as the base of the test statistic. The distribution of T is not a chi-square distribution but it is a distribution such as the one given in (1). Prior to Version 6 the second order adjustment of T in Mplus is computed using Satterthwaite's approximation as described in (2) and (3) for a particular matrix M given in formula (9) in Asparouhov and Muthén (2006). To construct a T_3 style difference we use again formula (4) using the corresponding matrix M. The advantage of T_3 over T_2 is again that the degrees of freedom for this chi-square approximation matches the difference in the number of parameters of the two nested models.

5 Modification Indices for Scaled Chi-Square Statistics

Let F be a fit function such as the log-likelihood, the log-likelihood difference between a restricted and unrestricted model or a weighted least squares fit function. In all of these cases F is used to construct a chi-square test of fit. Modification indices is a technique developed by Sörbom (1989) that allows us to obtain an approximate estimate for the change in the chi-square test when an additional parameter is estimated in the model. This is usually done for a large number of additional parameters and only using the first and the second order derivatives of F. When the test statistic is a scaled test statistic, it is obtained by using a formula of this form $aF+b$ where a and b are estimated quantities and F is the fit function value after the minimization. Using Sörbom (1989) method we can again make an approximate estimate for the change in the fit function value using first and second order derivatives of F. Let's assume that a and b are approximately constants or that they will

not change dramatically when an additional parameter is estimated. This is not a unreasonable assumption. Practical experience has shown that this is the case in most situations. Usually a and b are functions of averages of eigenvalues. When an additional parameter is estimated, essentially that amounts to eliminating a corresponding eigenvalue. These averages however will not be affected dramatically by removing a single eigenvalue. Under the assumption of approximately constant a and b we can use as the modification index the approximate change in F multiplied by a. This is the method used in Mplus for computing modification indices for all scaled chi-square tests of fit.

In certain situations the assumption of approximately constant a and b may not be appropriate, such as when the degrees of freedom is small. The modification indices technique, even for unscaled chi-square statistics, does not provide a precise value nor does it guarantee a small enough error in the estimation and therefore its use should be limited to an exploratory phase in the model building rather than used as evidence for model fit.

References

[1] Asparouhov, T & Muthén , B. (2006) Robust Chi Square Difference Testing with Mean and Variance Adjusted Test Statistics. Mplus Web Notes: No. 10. http://statmodel.com/download/webnotes/webnote10.pdf

[2] Muthén, B., du Toit, S. H. C., & Spisic, D. (1997). Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes.

[3] Muthén, B.O. (1998-2004). Mplus Technical Appendices. Los Angeles, CA: Muthén & Muthén Copyright 1998-2004 Muthén & Muthén Program Copyright 1998-2004 Muthén & Muthén Version 3.

[4] Satorra, A., and Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye and C. C. Clogg (Eds.), Latent variables analysis: Applications for developmental research. 399-419. Thousand Oaks, CA: Sage.

[5] Satorra, A., and Bentler, P. M. (2001). A Scaled Difference Chi-square Test Statistic for Moment Structure Analysis. Psychometrika 66, 507-514.

[6] Satterthwaite, F. E. (1941). Synthesis of variance. Psychometrika, 6, 309-316.

[7] Sörbom, D. (1989) Model modification. Psychometrika, 54, 371-384.