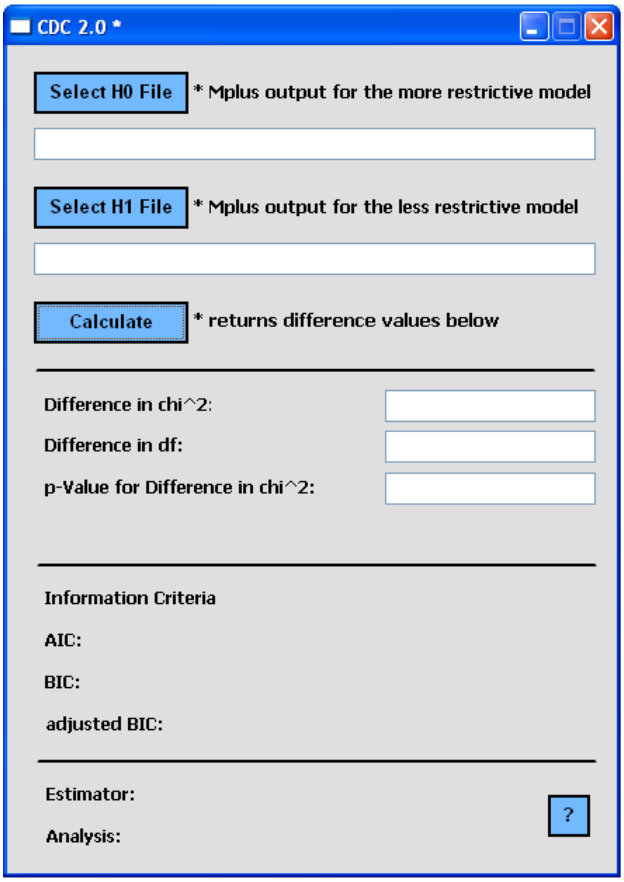
Little Helper #4:
The Chi-Square Difference Calculator



Claudia Crayen, crayen@gmail.com

Introduction

So-called likelihood-ratio-difference tests are often used to compare the fit of two nested models (see for example Bollen & Long, 1993; Schermelleh-Engel, Moosbrugger & Müller, 2003). The present software *Chi-square Difference Calculator* (CDC) facilitates this procedure for structural equation models (SEM) and latent class analysis models (LCA) that were estimated in the software Mplus (Muthén & Muthén, 1998-2007). Information criteria reported in the Mplus output files are also compared.

Theoretical Foundations of Chi-square Difference Testing

Likelihood-ratio tests are used for SEM and LCA to test the goodness of fit of a model to the data by means of a chi-square distributed statistic. A significant result of the likelihood-ratio test leads to rejection of the null hypothesis, that the model at hand fits the population perfectly.

You can also compare two hierarchically nested models by means of their likelihood-ratio statistic. Hierarchically nested implies that a so-called H0 model is a special case of a more general H1 model. This test based on the difference of the likelihood-ratio is often referred to as the chi-square difference test. The more restrictive H0 model has more degrees of freedom (df) due to additional restrictions compared to the H1 model. Because of those restrictions, the H0 model usually fits the data less well than the H1 model and exhibits a greater likelihood-ratio value. One can now test whether this difference in likelihood-ratio values is significant (whether the H0 model fits the data significantly worse that the H1 model)

The null hypothesis in this case is that the H0 model does NOT fit the data worse than the H1 model. The alternative hypothesis is that the H0 model does fit worse than the H1 model and that therefore the additional restrictions in the H0 model have been proven untenable.

For running the test, we form the difference of the likelihood-ratio values (H0 minus H1) and of the df (H0 minus H1). The chi-square difference value (Δ*chi-square*) follows a theoretical chi-square distribution with Δ*df* if the assumptions are met. If Δchi-square is significant, we reject the null hypothesis and assume that the H0 model does in fact fit the data worse than the H1 model.

General assumptions of chi-square difference testing:

- The H0 model is nested in the H1 model.
- The H1 is correctly specified / fits the data.
- The additional restrictions in the H0 model must not include fixing parameter to its boundary value, e.g. fixing a variance to zero, a correlation parameter to 1/-1, or a probability parameter to 0 or 1. See Stoel, Galindo Garre, Dolan & van den Wittenboer (2006).

Theoretical Foundations of Information Criteria

Information Criteria (see for example Schermelleh-Engel et al., 2003) are descriptive statistics for model comparisons. In contrast to the chi-square difference test, they can be used to compare models that are not hierarchically nested. Information Criteria take into account goodness of fit as well as model parsimony. A more restrictive model with fewer parameters is more parsimonious. Given equally well fitting models, the more parsimonious one is preferred (Occam's razor).

The better model in terms of fit and parsimony will have lower information criteria values. To date there is no test of significance for the difference in these values. Mplus reports *Akaike's Information Criterion* (AIC), *Bayesian Information Criterion* (BIC) and *sample size adjusted BIC*.

Mplus

The statistical software Mplus can be used to estimate SEM and LCA models (among others). Up to now (version 5.2), chi-square difference testing for maximum likelihood (ML) estimation has to be executed manually. We are lazy and developed the CDC.

The Chi-Square Difference Calculator

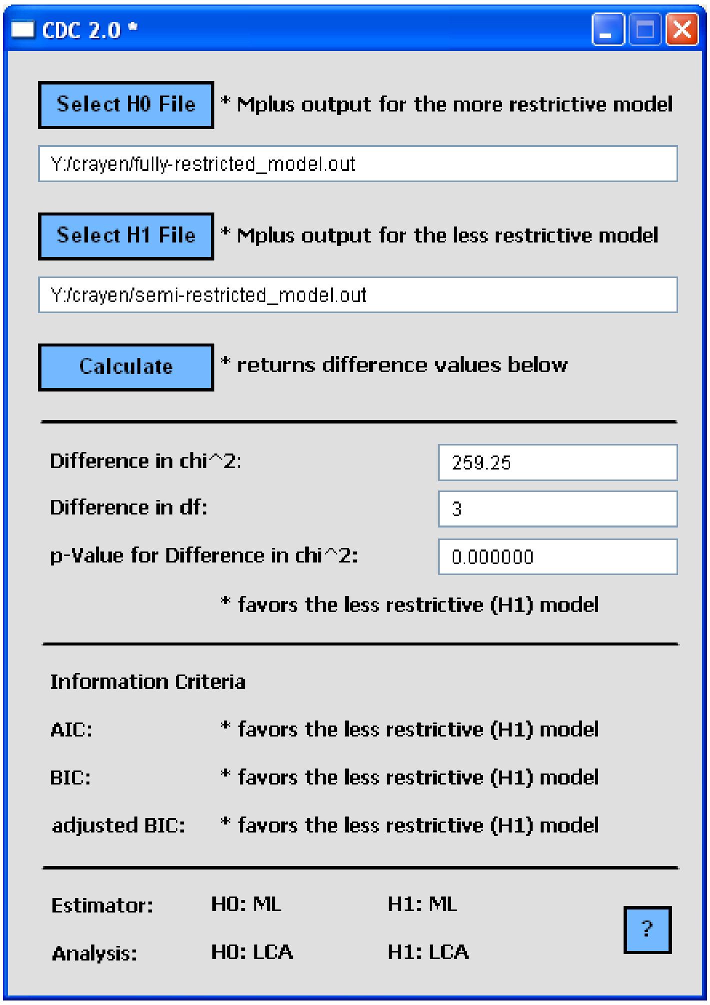
The CDC saves the user manually subtracting the values and entering the difference scores into a chi-square calculator. All you need are the Mplus outputs of the H0 and H1 model. Start the CDC by double-clicking on the **CDC2.exe** file. Since this is an executable file, a Windows warning message may appear. You have to ignore this in order to start the CDC.

The graphical user interface is quite simple. First, choose the output files of the models you'd like to compare. You can either just type in the filepath or search through the folders by clicking on the Select-Buttons. Your H0 file is the one for your more restrictive model with fewer parameters and more degrees of freedom.
When you've chosen your files, you can hit Calculate. The CDC will look for the necessary information in the output files and return the values in the lower part of the GUI.

The **Difference in chi^2** is simply the difference of the likelihood ratio values (H0 minus H1). The **Difference in df** is, as you may have guessed, the difference in degrees of freedom, (also H0 minus H1). These are the parameters used for calculating the **p-Value for Difference in chi^2**. When the p-value ≤ .05, the program will state that the H1 model is to be preferred, when the p-value > .05, the statement reads that the H0 model is to be preferred. The cut-off value of .05 is somewhat arbitrary, the user should decide which alpha error to base his or her decision on. The CDC also compares the Information Criteria and states for which model that particular criterium turned out more favorable (i.e., lower).
The program will also draw the information on analysis type (SEM vs LCA) and estimator (ML or MLR) from the ouput and return it in the lowermost section. Make sure that these are as intended. Note that Mplus does not report scaling factors for the MLR chi-square when latent class models are estimated. In this case, the CDC will inform you that it cannot execute the difference testing. Refer to www.statmodel.com/chidiff.shtml for ways to obtain chi-square difference values. When the H0 model and H1 model do not coincide on estimator and analysis type, you will get an error message because the selection of the outputs is most likely incorrect. You will also get an error message when the H0 and H1 model are mixed up (the difference in df will be negative and the CDC will notice).

There is NO check, however, whether the models are in fact nested within each other or fulfill the other assumptions (see above). In this case, the information criteria may still be interpretable, but the chi-square-difference test may be incorrect.

In the following example, we tested two LCA models that were estimated with regular ML. The difference in the likelihood-ratio values is 250.25. Based on a chi-square distribution with 3 df, this value is highly unlikely given the null hypothesis (p < .001). We therefore reject the null hypothesis (that the H0 model does not fit the data worse) and decide in favor of the less restrictive H1 model. All Information Criteria point to the same decision.



Caution

The CDC will only work with Mplus outputs. So far, we only used it for structural equation models (including path analysis and confirmatory factor analysis) and simple latent class analysis. We provide no warranty for the results. The CDC is freeware and may be distributed.

References

Bollen, K. A. & Long, S. (Eds.) (1993). *Testing structural equation models*. Newbury Park: Sage Publications.

Geiser, C. (2010). *Datenanalyse mit Mplus: Eine anwendungsorientierte Einführung*. Wiesbaden: VS Verlag.

Muthén, L. K. & Muthén, B. O. (1998-2007). *Mplus User's Guide. Fifth Edition*. Los Angeles, CA: Muthén & Muthén. http://statmodel.com/ugexcerpts.shtml

Schermelleh-Engel, K., Moosbrugger, H. & Müller, H. (2003). Evaluating the fit of structural equation models: Test of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research - Online, 8*, 23–74. http://www.dgps.de/fachgruppen/methoden/mpr-online

Stoel, R. D., Galindo Garre, F., Dolan, C. & van den Wittenboer, G. (2006). On the Likelihood Ratio test in structural equation modeling when parameters are subject to boundary constraints. *Psychological Methods, 4*, 439–455.