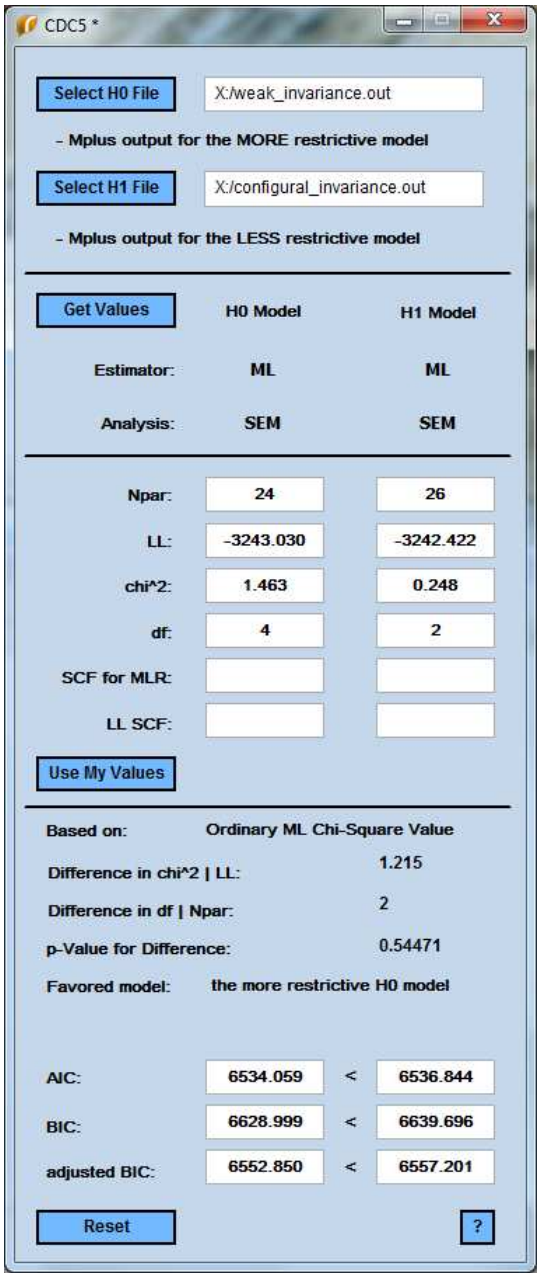
# Chi-Square Difference Calculator




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, 1998-2010). 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 6.1), chi-square difference testing for maximum likelihood (ML) and robust maximum likelihood (MLR) 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 **CDC5.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 divided into four sections.

## File section

Here, you can choose the output files of the models you would like to compare. You can either just type in the file path or search through the folders by clicking on the "**Select File**" 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 the "**Get Values**" button. The CDC will extract the necessary information from the output files and return the values in the Analysis and Value sections of the GUI. The left column of properties belongs to the H0 model, the right column to the H1 model.

## Analysis section

The program will draw information on analysis type (SEM or LCA) and estimator (ML or MLR) from the ouput file. Make sure that these are as intended. There are two basic checks implemented. When the H0 model and H1 model do not coincide on estimator and analysis type, you will get a warning message because the selection of the outputs is most likely incorrect. You will also get a warning message when the difference in degrees of freedom is negative (most likely, H0 and H1 model were swapped by mistake) or zero. This warning message can also occur if the specified file cannot be found.
There is NO check, however, whether the models are in fact nested within each other or fulfil the other assumptions (see above). In this case, the chi-square-difference test may be incorrect.

## Values section

In the **Npar** you'll find the number of parameters for each model. In the **LL** row, the Log-Likelihood values are stated and in the **df** row, the degrees of freedom. You can find the chi-square values in the **chi^2** line. For latent class models, these are the Likelihood Ratio chi-square values (not the Pearson chi-Squares). The **SCF for MLR** row gives the scaling factors for the chi-square value reported in the Mplus output if robust maximum likelihood estimation was used. The **LL SCF** is the H0 scaling correction factor for the Loglikelihood (if available in the output). If you are having trouble importing the values from the output or if

you are using other software packages, you can enter the according values into the fields by hand. If you click on the button "**Use My Values**", these values will be used for difference testing instead of the ones imported from the Mplus outputs. If you provide LL values, a test based on the Loglikelihood will be performed. If you provide scaling factors for the chi-square values, a scaled test will be performed. If you provide neither, the ordinary chi-square test will be performed.

### Results section

In the **Based on** line, you will find details on the method used to calculate the difference. For structural equation models estimated with ML, this is simply the difference between the ordinary chi-square values. For structural equation models estimated with MLR, the chi-square difference using the scaling correction factor (SCF) is calculated according to the Mplus website (www.statmodel.com/chidiff.shtml). For latent class models, the difference is calculated using the Loglikelihood (with according scaling factor if MLR is used).
The **Difference in chi^2 | LL** is simply the difference of the chis-square or likelihood ratio values (H0 minus H1). This difference is scaled for the MLR. The **Difference in df** is the difference in degrees of freedom (scaled if necessary) or number of parameters, (also H0 minus H1). These are the parameters used for calculating the **p-Value for Difference**. When the p-value ≤ .05, the program will state that the H1 model is to be preferred, when the p-value > .05, the statement is that the H0 model is to be preferred. The cut-off value for this automated decision is at the common .05, but the user should decide which Type I error to base his or her decision on. The CDC also states the Information Criteria and indicates for which model that particular criterion is lower (i.e., better).

## Caution

This CDC version was only tested with Mplus 6 outputs. So far, we only used it for structural equation models (including path analysis and confirmatory factor analysis) and latent class analysis. We provide no warranty for the results. You should check your final results by hand.

## 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. (1998-2010). *Mplus User's Guide. Sixth 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.