Use of deviance statistics for comparing models

A likelihood-ratio test can be used under full ML. The use of such a test is a quite general principle for statistical testing. In hierarchical linear models, the deviance test is mostly used for multiparameter tests and for tests about the random part of the model.

This approach is based on estimating two models, M_0 and M_1. It is assumed that M_0 excludes the effects hypothesized to be null, while these effects are included in M_1. For each model, a deviance statistic, equal to -2 ln *L* for that model, is computed. The deviance can be regarded as a measure of lack of fit between model and data. In general, the larger the deviance, the poorer the fit to the data. Note that the value of a deviance could be negative. For the interpretation of a negative deviance, please see Log likelihood values and negative deviances. The deviance in usually not interpreted directly, but rather compared to deviance(s) from other models fitted to the same data.

The difference between the deviances D_0 and D_1 has a large-sample chi-square distribution with degrees of freedom equal to the difference in the number of parameters estimated. Large values of the chi-square statistic are taken as evidence that the null hypothesis is implausible.

An approximate rule for evaluating the chi-square is given by Snijders & Bosker, p 49:
" ...a difference in deviance between models should be at least twice as large as the difference in the number of parameters estimated".

These authors also give two examples of using deviance statistics (p 89 of the text).

Note, however, the implications of using other estimation procedures: REML in the linear two-level models and PQL in the HGLM models.

Example 1: Multivariate tests of variance-covariance components specification

Below we compare the variance-covariance components of two Intercept-and-Slope-as-Outcome models. One treats β_1 as random and the other does not. The two models are shown below, with the deviance statistics of the model.

Model 1:

```
Level-1 Model
       Y = B0 + B1*(SES) + R

Level-2 Model
       B0 = G00 + G01*(SECTOR) + G02*(MEANSES) + U0
       B1 = G10 + G11*(SECTOR) + G12*(MEANSES)

 Statistics for current covariance components model
 --------------------------------------------------
 Deviance                       = 46502.952743
 Number of estimated parameters = 2
```

Model 2:

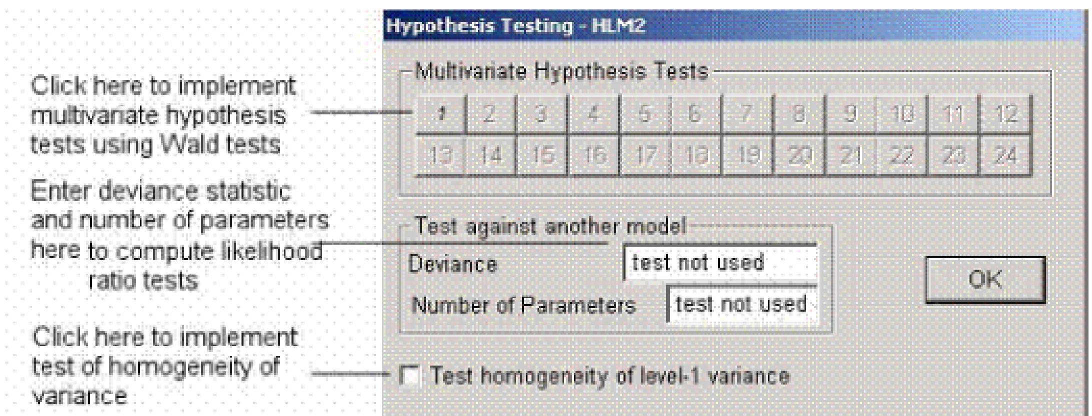
```
Level-1 Model
```

```
       Y = B0 + B1*(SES) + R

Level-2 Model
       B0 = G00 + G01*(SECTOR) + G02*(MEANSES) + U0
       B1 = G10 + G11*(SECTOR) + G12*(MEANSES) + U1

 Statistics for current covariance components model
 --------------------------------------------------
 Deviance                       = 46501.875643
 Number of estimated parameters = 4
```

The multi-parameter test for the variance-covariance components is implemented by a likelihood-ratio test, which compares the deviance statistic of a restricted model with a more general alternative. The user must input the value of the deviance statistic and related degrees of freedom for the alternative specification.

To specify a multivariate test of variance-covariance components

Enter the deviance and the number of parameters in the **Deviance Statistics** box and in the **Number of Parameters** box respectively (the two numbers for our example are 46501.875643 and 4).



The HLM2 output associated with this test appears below.

```
  Variance-Covariance components test
 -----------------------------------
 Chi-square statistic         =      1.07714
 Number of degrees of freedom =    2
 P-value                      = >.500
```

The test result indicates that a model that constrains the residual variance for the SES slopes, B1, to zero appears appropriate.

Example 2: Testing of variance-covariance structures

Consider the HMLM model from the Sustaining Effects (EG) data, we can assess the variance-covariance structures of the model using the deviance statistics of the following three models:

Model 1. Unrestricted model.

Level-1 model:

$$y = IND_1 \times y_1^* + IND_2 \times y_2^* + IND_3 \times y_3^* + IND_4 \times y_4^* + IND_5 \times y_5^* + IND_6 \times y_6^*$$
$$y^* = \pi_0 + \pi_1(YEAR) + e$$
$$Var(e) = \Delta$$

Level-2 model:

$$\pi_0 = \beta_{00}$$
$$\pi_1 = \beta_{10}$$

Level-3 model:

$$\beta_{00} = \gamma_{000} + u_0$$
$$\beta_{10} = \gamma_{100} + u_1$$

Model 2. Random effects model with homogeneous level-1 variance.

Level-1 model:

$$y = IND_1 \times y_1^* + IND_2 \times y_2^* + IND_3 \times y_3^* + IND_4 \times y_4^* + IND_5 \times y_5^* + IND_6 \times y_6^*$$
$$y^* = \pi_0 + \pi_1(YEAR) + e$$
$$Var(e) = Var(\mathbf{A} \times r + e) = \Delta = \mathbf{A}\boldsymbol{\tau}_{\pi}\mathbf{A}' + \mathbf{S}$$

Where $\mathbf{S} = \sigma^2\mathbf{I}$

Level-2 model:

$$\pi_0 = \beta_{00} + r_0$$
$$\pi_1 = \beta_{10} + r_1$$

Level-3 model:

$$\beta_{00} = \gamma_{000} + u_0$$
$$\beta_{10} = \gamma_{100} + u_1$$

Model 3. Random effects model with heterogeneous level-1 variance.

Level-1 model:

$$y = IND_1 \times y_1^* + IND_2 \times y_2^* + IND_3 \times y_3^* + IND_4 \times y_4^* + IND_5 \times y_5^* + IND_6 \times y_6^*$$
$$y^* = \pi_0 + \pi_1(YEAR) + e$$
$$Var(e) = Var(\mathbf{A} \times r + e) = \Delta = \mathbf{A}\boldsymbol{\tau}_{\pi}\mathbf{A}' + \mathbf{S}$$

Where $\mathbf{S} = Diag(\sigma^2(1), \sigma^2(2), \ldots, \sigma^2(6))$

Level-2 model:

$$\pi_0 = \beta_{00} + r_0$$
$$\pi_1 = \beta_{10} + r_1$$

Level-3 model:

$$\beta_{00} = \gamma_{000} + u_0$$
$$\beta_{10} = \gamma_{100} + u_1$$

The output for the fitness of the three models is as follows.

```
Summary of Model Fit

 -------------------------------------------------------------------
 Model                                        Number of         Deviance
                                              Parameters
 -------------------------------------------------------------------
 1. Unrestricted                                 26           15960.50733
 2. Homogeneous sigma-squared                     9           16326.23111
 3. Heterogeneous sigma-squared                  14           16140.15892

 -------------------------------------------------------------------
 Model Comparison                         Chi-square       df    P-value
 -------------------------------------------------------------------
 Model 1 vs Model 2                         365.72378      17     0.000
 Model 1 vs Model 3                         179.65159      12     0.000
 Model 2 vs Model 3                         186.07219       5     0.000
```

The first test of the model comparison tests, comparing "Model 1 vs Model 2", is testing for the hypothesis

$$H_0 : \Delta = \mathbf{A}\boldsymbol{\tau}_{\pi}\mathbf{A}' + \sigma^2\mathbf{I}$$

against the alternative

$$H_1 : \Delta \text{ is an unstructured } 6\times 6 \text{ matrix.}$$

The chi-square test statistic of 365.72 with 17 degree of freedom gives a *p*-value of 0.000, indicating that the null hypothesis is implausible, and we can conclude that model 1 is a better fitted model than model 2.

The second test of the model comparison tests, comparing "Model 1 vs Model 3", is testing for the hypothesis

$$H_0 : \Delta = \mathbf{A}\boldsymbol{\tau}_{\pi}\mathbf{A}' + \mathbf{S} \quad \text{where} \quad \mathbf{S} = Diag(\sigma^2(1), \sigma^2(2), \ldots, \sigma^2(6))$$

against the alternative

$$H_1 : \Delta \text{ is an unstructured } 6\times 6 \text{ matrix.}$$

The chi-square test statistic of 179.65 with 12 degree of freedom has a p-value of 0.000, indicating that the null hypothesis does not seem tenable, and we can conclude that model 1 rather than model 3 is a better model.

The third test of the model comparison tests, comparing "Model 2 vs Model 3", is testing for the hypothesis

$$H_0 : \Delta = \mathbf{A}\boldsymbol{\tau}_{\boldsymbol{\beta}}\mathbf{A}' + \mathbf{S} \text{ where } \mathbf{S} = \sigma^2\mathbf{I}$$

against the alternative

$$H_0 : \Delta = \mathbf{A}\boldsymbol{\tau}_{\boldsymbol{\beta}}\mathbf{A}' + \mathbf{S} \text{ where } \mathbf{S} = Diag(\sigma^2(1), \sigma^2(2), \ldots, \sigma^2(6))$$

The chi-square test statistic of 186.07 with 5 degree of freedom gives a p-value of 0.000, indicating again that the null hypothesis is implausible, and we can conclude that model 3 is significantly different from model 2.

Summarizing the three tests above, we can draw a conclusion with respect to the variance-covariance structure of the HMLM model: The model with unrestricted variance-covariance structure for the combined level-1 and level-2 model is the best fitted model for the given data.