Computing the Explained Variance in Multilevel Models with Mplus

Tihomir Asparouhov and Bengt Muthén

July 8, 2021

In this note we illustrate how to compute the explained variance in multilevel models in Mplus. A general formulation of the explained variance is given in Rights and Sterba (2019). This formulation focuses on the overall explained variance and is different from the cluster specific explained variance that can be obtained in Mplus with the command **output:stand(cluster)**. Nevertheless, it is possible to compute the overall explained variance in Mplus using Rights and Sterba (2019) method and the **Model Constraint** command in Mplus.

We illustrate this with the following simple example. Let Y_{ij} be the dependent variable for individual i in cluster j and X_{1ij} and X_{2ij} be two predictors that are assumed to be cluster mean centered. The model is given by the following equations

$$Y_{ij} = \alpha_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \varepsilon_{ij}$$

where β_{1j} and β_{2j} are the random slopes for the two predictors, α_{0j} is the random intercept, and ε_{ij} is the residual. The three random effects are assumed to have a normal distribution

$$\begin{pmatrix} \alpha_{0j} \\ \beta_{1j} \\ \beta_{2j} \end{pmatrix} \sim N\left(\begin{pmatrix} m_0 \\ m_1 \\ m_2 \end{pmatrix}, \begin{pmatrix} v_{00}, v_{01}, v_{02} \\ v_{01}, v_{11}, v_{12} \\ v_{02}, v_{12}, v_{22} \end{pmatrix} \right).$$

In addition,

$$\varepsilon_{ij} \sim N(0, \sigma).$$

The random effects can also be written as follows

$$\alpha_{0j} = m_0 + \xi_{0j}$$

$$\beta_{1j} = m_1 + \xi_{1j}$$
$$\beta_{2j} = m_2 + \xi_{2j}$$

where

$$\begin{pmatrix} \xi_{0j} \\ \xi_{1j} \\ \xi_{2j} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} v_{00}, v_{01}, v_{02} \\ v_{01}, v_{11}, v_{12} \\ v_{02}, v_{12}, v_{22} \end{pmatrix} \right).$$

Therefore

$$Y_{ij} = m_0 + \xi_{0j} + m_1 X_{1ij} + m_2 X_{2ij} + \xi_{1j} X_{1ij} + \xi_{2j} X_{2ij} + \varepsilon_{ij}.$$

If the variance covariance matrix for the covariates X_{1ij} and X_{2ij} is

$$\begin{pmatrix} \sigma_{11}, \sigma_{12} \\ \sigma_{12}, \sigma_{22} \end{pmatrix}$$

then we can compute the total model estimated variance for Y_{ij} as follows

$$Var(Y_{ij}) = Var(\xi_{0j}) + Var(m_1 X_{1ij} + m_2 X_{2ij}) + Var(\xi_{1j} X_{1ij} + \xi_{2j} X_{2ij}) + Var(\varepsilon_{ij}) =$$

$$v_{00} + (m_1^2 \sigma_{11} + 2 m_1 m_2 \sigma_{12} + m_2^2 \sigma_{22}) + (v_{11}\sigma_{11} + 2 v_{12} \sigma_{12} + v_{22} \sigma_{22}) + \sigma$$

The within level variance of Y_{ij} is

$$V_w = (m_1^2 \sigma_{11} + 2 m_1 m_2 \sigma_{12} + m_2^2 \sigma_{22}) + (v_{11}\sigma_{11} + 2 v_{12} \sigma_{12} + v_{22} \sigma_{22}) + \sigma.$$

The within level proportion of the variance explained by the fixed effects is

$$R2A = (m_1^2 \sigma_{11} + 2 m_1 m_2 \sigma_{12} + m_2^2 \sigma_{22}) / V_w.$$

The within level proportion of the variance explained by the random part of the effects is

$$R2B = (v_{11}\sigma_{11} + 2 v_{12} \sigma_{12} + v_{22} \sigma_{22}) / V_w.$$

The Mplus input file for computing these quantities is as in Figure 1. The values $\sigma_{11} = 1.169$, $\sigma_{12} = 0.404$ and $\sigma_{22} = 0.798$ are the variances and the covariance for the covariates. Those can be obtained in Mplus with a preliminary run and the option **output:samp**. Alternatively, if the variances for the covariates are mentioned in the model these quantities can be estimated as a part of the multilevel model.

Figure 1: Computing explained variance in Mplus for multilevel models

```
variable:
       names are y x1 x2 c;
       within = x1 x2;
       cluster=c;
data:  file = ex.dat;
define:    center x1-x2 (groupmean);
ANALYSIS:  TYPE = TWOLEVEL RANDOM;

model:
       %within%
       b1 | y on x1;
       b2 | y on x2;
       y (sigma);

       %between%
       [b1-b2] (m1-m2);
       y (v00); b1 (v11); b2 (v22);
       y with b1-b2;
       b1 with b2 (v12);

output:samp;

model constraint:
 new(t1); t1=m1*m1*1.169+2*m1*m2*0.404+m2*m2*0.798;
 new(t2); t2=v11*1.169+v22*0.798+2*v12*0.404;
 new(Vw); Vw=t1+t2+sigma;
 new(V); V=Vw+v00;
 new(R2a); R2a=t1/Vw;
 new(R2b); R2b=t2/Vw;
```

References

[1] Rights, J. D., & Sterba, S. K. (2019). Quantifying explained variance in multilevel models: An integrative framework for defining R-squared measures. Psychological Methods. 24, 309–338.