HMLM – Hierarchical Multivariate Linear Models EPSY 8268

This provides a modeling option for repeated observations (level 1) nested within persons (level 2) with a fixed design, for example with *T* observations, different outcome variables, or experimental conditions per person. Given the fixed design, this can be conceived of as a repeated measures design. The HMLM provides the ability to complete estimation in the presence of missing data, such that *nj* observations are collected (where *nj* ≤ *T*).

In HMLM, a number of models can be compared.

1. **Unrestricted model**

There is a fixed design for data collection, but some observations are missing at random. This model assumes a constant *T* × *T* covariance matrix for each person’s complete data.

Level-1 Model:

*Yhi* is the *r*th outcome for person *i* at time *h*. In this model, is the value of *Y* that person *i* would have obtained if the person was observed at time *t*, and *mthi* is an indicator variable = 1 when that observation did occur at time *t*, 0 if not.

 is where *t* = 1, …, *T* for the complete data for person *i*.

 is where *h* = 1, …, *Ti* for the observed data with *mthi* indicators for the pattern of missingness.

Consider the example where *T* = 5 and person *i* has data from occasions 1, 2, and 4, but not the others.

In matrix notation, this is *Yi* = *MiYi*\*

The structural model for Level 1:

In matrix notation, *Yi*\* = *A*π*i* + ε*i*

 ~ *N* (0, Δ) where Δ captures all of the variation within and between persons. This estimates the variance at each time point and all covariances.

Level-2 Model:

In matrix notation, π*i* = *Xi*β

There is no random variation between persons in the regression coefficients .

The combined model, *Yi*\* = *AXi*β + ε*i*

And in the model of observed data, *Yi* = *Mi* *AXi*β + *Mi* ε*i*

The number of parameters is *f* + *T*(*T* + 1)/2, where *f* is the number of fixed effects and *T* is the number of intended observations per person.

This is the most complex of the models to be compared. The remaining models impose constraints, and thus can be compared using a likelihood ratio test.

1. **Homogenous Level-1 Variance**

Level 1 in matrix notation, *Yi*\* = *A*π*i* + ε*i* where ~ *N* (0, Σ) and Σ = σ2 I*T*. This indicates that the variance is the same for all *t* time points.

Level 2 includes covariates, and because of the level-1 homogenous variance, degrees of freedom are available to estimate randomly varying intercepts and slopes:

In matrix notation, π*i* = *Xi*β + *ri* , provided *T* is large enough, intercepts/slopes can be randomly varying.

The combined model is *Yi*\* = *A Xi*β + *Airi* + *ei* which is = *A Xi*β + ε*i*

Var(ε*i*) = Var (*Airi* + *ei*) = *A*τ*A* + σ2 I*T* = Δ

The number of parameters is *f* + *r*(*r* + 1)/2 + 1, where *r* is the dimension of τ. In this case, *r* must be less than *T*.

1. **Heterogeneous (varying) Level-1 Variance**

The level-1 model is the same as the homogenous level-1 variance model, where

Var(*ei*) = Σ = *diag*{}, where Σ is diagonal with elements , the variance associated with occasion *t*, where *t* = 1, …, *T*. This indicates that the variance varies over time.

The number of parameters is *f* + *r*(*r* + 1)/2 + *T*, where *r* can be no larger than *T* – 1. When *r* = *T* – 1, the results duplicate those from the unrestricted model.

1. **Log-Linear Model for Level-1 Variance**

The model above has varying level-1 variance, which assumes a unique level-1 variance for every occasion. A more parsimonious model would specify a functional association between occasion (time or age) and the variance. In this case, Σ = *diag*{}, and also

The natural log of the level-1 variance may be a linear or quadratic function of age. If the explanatory variables *cl* are *T* – 1 dummy variables, indicators of measurement occasions, this is similar to the varying level-1 variance.

The number of parameters is *f* + *r*(*r* + 1)/2 + *L* + 1, where *r* can be no larger than *T* – 1 and *L* can be no larger than *T* – 1.

1. **First-Order Auto-Regressive Model for Level-1 Residuals**

Level-1 residuals are correlated, but only through the immediately preceding residuals.

Cov(*eti*, *et'i*) = σ2 ρ| *t*-*t'* |

The variance at each time point is σ2 and each correlation diminishes between time points. As the distance between occasions increases from 1, 2, 3, …, the correlations become ρ, ρ2, ρ3, etc.

The number of parameters is *f* + *r*(*r* + 1)/2 + 2, where *r* can be no larger than *T* – 1.

Level-1 predictors are assumed to have the same values for all level-2 units. If the design for *apti* varies over individuals *i*, its coefficient cannot vary randomly at level 2. The standard 2-level model will be more flexible in these cases (as in the example presented in class).

**A Multilevel Multivariate Model**

The extends the above models to the case where individuals with multiple outcomes are nested within higher-level units or clusters. In this case, HMLM2 is the appropriate approach to create MDM models for HLM analysis.