Center for Response to Intervention in Early Childhood
University of Minnesota
Michael C. Rodriguez
2013

Estimating Growth in HLM

Hierarchical Linear Modeling (HLM) can be used to estimate the intercept and slope of growth curves for individual children. These estimates also can be tested for differences, for example, given ELL and IEP status. The HLM model described here is a two-level growth model, with time points (or waves of administration) within child at level one and children at level two. This could be expanded to three levels, including children nested within childcare center at level three. Rasch scores in the logit metric are well suited for HLM analyses, given their linear characteristics.

In models estimate for CRtIEC, time-point is based on the age of the child at time of administration in months. To facilitate interpretation, age is centered at 60 months. Thus the intercept and slope can be interpreted as follows:

INTERCEPT: The interpretation of the intercept is the IGDI score at the age of 60 months; age was computed by taking the difference between date of administration and date of birth, centered around 60 months (age at administration – 60).

SLOPE: The interpretation of the slope is the change in IGDI score per month.

Unconditional Model

First, an unconditional model (without child-level characteristics or predictors) is estimated to provide a baseline for the intercept and growth rate. This includes the following specification, for child *i* of a given age at time *t*. The Level-1 model estimates the intercept (π_{0i}) and slope (π_{1i}) for each child *i*. The Level-2 model estimates the grand-mean for both the intercept (β_{00}) and slope (β_{10}) across children.

Level-1 Model: $\text{Score}_{ti} = \pi_{0i} + \pi_{1i}(\text{Age}–60)_{ti} + e_{ti}$

Level-2 Model:

$$\pi_{0i} = \beta_{00} + r_{0i}$$
$$\pi_{1i} = \beta_{10} + r_{1i}$$

This model is sufficient to estimate performance at age 60 months (the intercept) and change in performance per month (the slope). From the HLM analyses, we can also estimate variability in these estimates across children and obtain information about the reliability or certainty of the estimates – more on this below.

Full Model

An optional full model can be estimated to provide a test of the role of child characteristics on the intercept and slope – for example, we may wonder if performance differs given child ELL status and IEP status. The Level-2 model can estimate the effects of child characteristics, ELL and IEP status, on the intercept (β_{01}, β_{02}) and slope (β_{11}, β_{12}) respectively. Here the ELL and IEP indicators are entered uncentered, so the grand mean effects, intercept (β_{00}) and slope (β_{10}), are the means for students who are not ELL nor have IEPs. The remaining coefficients of level-2 are then deviations for ELL status (β_{01}, β_{11}) and IEP status (β_{02}, β_{12}).

Level-1 Model: $\text{Score}_{ti} = \pi_{0i} + \pi_{1i}(\text{Age}–60)_{ti} + e_{ti}$

Level-2 Model: $\pi_{0i} = \beta_{00} + \beta_{01}(\text{ELL})_i + \beta_{02}(\text{IEP})_i + r_{0i}$
$\pi_{1i} = \beta_{10} + \beta_{11}(\text{ELL})_i + \beta_{12}(\text{IEP})_i + r_{1i}$

Effect Sizes

To support the interpretation of the effects of ELL and IEP on the intercept, the coefficients can be standardized based on the SD of the intercept from the unconditional model. Similarly, the effects on the slope can be based on the SD of the slope from the unconditional model. These are stated in terms of SDs – a standardized effect size.

A note on the use of reliability to estimate fixed-effects in HLM.

Level-1 parameters (intercepts and slopes of participants) are based on “Bayesian shrunken estimates”. They are empirical Bayes estimates of the true intercept and slope – which means that they are based on information from the individual’s data and some information from the data across all individuals (the means). If the information from an individual is not reliable, we supplement that information with more reliable information from the whole sample – thus, the Bayesian shrinkage idea – we shrink the estimates of the individual toward the mean if the information from the individual is not reliable.

To facilitate this process, each estimate (each intercept and slope) has an associated reliability. This reliability is a function of the ratio of true variance to observed variance (the proportion of variance that is true, the classical test theory notion). The estimate for person *i* can be computed:

$$\lambda_i = \frac{\tau_{00}}{(\tau_{00} + V_{00i})} = \frac{\textit{Parameter Variance}}{\textit{Parameter Variance} + \textit{Error Variance}} = \frac{\textit{True Variance}}{\textit{Observed Variance}}.$$

The reliability λ_i for the intercept, as an example, will be high when the intercepts for individuals, π_{0i}, vary a great deal across individuals or the sample size n_i is large (the number of observations or time points within individual). The same is true for slopes – reliability is high when individuals vary in their slopes and the number of observations within individual is large.

Growth (Slope) Precision and Reliability in HLM

In HLM, slopes are estimated based on multiple data points on a common scale over time, within person. So each person is providing information to estimate a slope or change in performance over time (time points are nested within person).

Each person will then have an intercept, and when the time points are centered in a meaningful way, the intercept can have a specific interpretation, such as baseline, or performance at a specific age, or whatever is desired. Each person will also have a slope, which estimates the change in performance per unit change over time. Both the intercept and slope are randomly varying over persons, such that there is a grand mean intercept and slope, and variances of intercepts and slopes.

The person intercept and slope are person parameters. In addition, there are estimates of precision for each parameter, based on the amount of information available for estimating the intercept and slope and variability or stability in those estimates – the reliabilities in HLM.

Reliability of the intercept for person *i* is denoted as λ_{0i}. This is the ratio of true variance (parameter variance) to total variance (parameter + error variance). The traditional definition of reliability is the ratio of true-score variance to observed-score variance – the proportion of observed variance that is true variance. These variances are estimated in HLM so that parameter variance in the intercept is τ_{00} (between person variance) and total variance is the sum of τ_{00} and error variance for person *i*.

The basic Level-1 model is:

$$Score_{ti} = \pi_{0i} + \pi_{1i}(Time_{ti}) + e_{ti}, \text{ where } e_{ti} \sim \text{N}(0, \sigma_e^2).$$

The basic Level-2 model is:

$$\pi_{0i} = \gamma_{00} + r_{0i}, \text{ where } r_{0i} \sim \text{N}(0, \tau_{00}), \text{ and}$$

$$\pi_{1i} = \gamma_{10} + r_{1i}, \text{ where } r_{1i} \sim \text{N}(0, \tau_{10}).$$

Averaging across the n_t observations within student *i* gives a model with the sample mean as the outcome: $\bar{Y}_{\bullet i} = \pi_{0i} + \bar{e}_{\bullet i}$. This is a model where the sample student mean (or baseline if time is centered correctly) is an estimate of the student mean (intercept) β_{0i}, and the error of estimation is $\bar{e}_{\bullet i} = \sum_{t=1}^{n_t} \frac{e_{ti}}{n_t}$, which has an error variance $Var(\bar{e}_{\bullet i}) = \frac{\sigma_e^2}{n_t} = V_{0i}$. V_{0i} is the error variance, essentially a sampling error variance or the variance of $\bar{Y}_{\bullet i}$ as an estimate of π_{0i}. This is analogous to the error variance of estimating the mean – the SE(M) = $\sqrt{\frac{\sigma^2}{n}}$ or SE(M) = $\frac{s}{\sqrt{n}}$.

Now we can estimate the reliability of π_{0i}, the intercept for student *i*:

$$\lambda_{0i} = \frac{Var(\pi_{0i})}{Var(\bar{Y}_{\bullet i})} = \frac{\tau_{00}}{(\tau_{00} + V_{00i})} = \frac{Parameter\ Variance}{Parameter\ Variance + Error\ Variance} = \frac{True\ Variance}{Observed\ Variance}.$$

In terms of classical test theory, $\bar{Y}_{\bullet i}$ is a measure of the true parameter β_{0i}. λ_{0i} is the reliability because it is the ratio of true-score variance to total or observed-score variance, the extent to which what we observe is true. The reliability λ_{0i} will be large when the student means, π_{0i}, vary a great deal across students or the n_t number of time points, is large.

Reliability λ_{1i} of the growth slope π_{1i} over time points is estimated similarly. For person *i*,

$$\lambda_{1i} = \frac{\tau_{11}}{(\tau_{11} + V_{11i})}$$

As with the reliability of the intercept, the reliability λ_{1i} of the slope is large when student slopes vary a great deal across students and the number of time points or observations within student is large.

The average across persons provides a summary index of the reliability of the slope estimate for this population. The HLM program computes an average reliability for the estimates of the slope across the set of *i* level-2 persons.

For more information, see:

Raudenbush, S.W. & Bryk, A.S. (2002). *Hierarchical Linear Models. Applications and Data Analysis Methods* (2nd ed., pp. 50-79). Sage Publications.