FI=HLML02BB

NOTATION IN HLM FOR LEVEL II AND COMBINED MODEL

by
Ralph B. Taylor
breck@rbtaylor.net

In this section we review the notation for the Level-2 and the combined model. Further, we explore different types of Level-2 models.

I will continue with examples with fear of crime as our outcome variable and neighborhoods as our Level-2 units.

In this example we have age (x1) as a Level-1 predictor, and changes in neighborhood crime rate, and neighborhood stability, as Level-2 predictors

PREDICTORS IN LEVEL-1 ARE OUTCOMES IN LEVEL-2

This is an important idea:

> Everything that you have attempted to predict in your Level-1 model -- your intercepts, and your slopes -- becomes an outcome at Level-2.

You can either see how much these vary across your grouping units, or you can try and predict these intercepts and slopes using group level variables.

Whereas predictors in Level-1 are Xs, in the book the Level-2 predictors are Ws.

LEVEL-2 MODEL EQUATIONS

There are two equations in the Level-2 model. The first seeks to predict the intercepts of the different groups. In the case of one Level-2 predictor:

LEVEL-2 EQUATION FOR INTERCEPTS

B0J = G00 + G01(WJ) + U0J Eq. 2.1[1]

[1] B&R Eq. 2.4a, p. 14

where

B0J intercept (outcome scores), one value for each of the J groups.

These intercepts, one for each group, have been generated by the Level-1 model.

We want to be clear that these are **empirical Bayes estimates** of the **true** means for the entire neighborhood from which each sample came.

The program will also tell you about the **reliability** of B0J.

Reliability = how well does B0J correspond to the actual sample means on the outcome (Y.j)?
= average reliability of sample means (Y.j) as estimates of the "true" school mean outcome scores (B0J)
= λ

The biggest determinant of λ is sample size, within each neighborhood. As average sample size increases, reliability increases

How you interpret B0J <u>conceptually</u> depends on what you have done with your predictors in your Level-1 model; have you group centered them? Or grand mean centered them? Or left them as is?

G00 estimated mean score on the outcome for groups **in the population** that score 0 on WJ

It is the intercept (estimated mean outcome score) for all Level-2 units scoring zero on each Level-2 predictor **in the population** from which the sample was drawn.

For example, suppose your Level-2 predictor was a dummy variable for neighborhoods with an increasing vs. steady or decreasing violent crime rate.

W1 0 = steady or decreasing violent crime rate in neighborhood, for last 5 years

1 = increasing violent crime rate in neighborhood, for last 5 years

If our outcome is a fear variable, G00 represents an estimate of the mean fear score in the population of neighborhoods with steady or decreasing crime (W1=0).

Suppose we have two Level-2 predictor variables and they are both dummy variables? The same caution applies to interpreting G00 as appplied to interpreting B0J in the Level 1 model.

Suppose, in addition to W1, we have a second Level-2 predictor reflecting neighborhood instability

W2 0 = neighborhood relatively stable, with over 40% owner occupied households

1 = neighborhood relatively UNstable, with fewer than 40% owner occupied households

Our complete equation is

B0J = G00 + G01(W1) + G02(W2) + U0J Eq. 2.2

G00 now reflects the mean fear score **for relatively stable neighborhoods with a steady or declining violent crime rate**. These are the neighborhoods that will score zero on both Level-2 predictors.

Now let's move on to talk about the effects of neighborhood level characteristics on the neighborhood outcome.

G01 **estimate** of the slope of group mean fear on the first Level-2 predictor variable **in the population**. You have one of these for each Level-2 predictor.

In the case of our first equation (2.1) with only 1 Level-2 predictor variable, if that predictor is a dummy variable, G01 represents **a population estimate** of the mean fear difference between steady or decreasing violent crime neighborhoods, and increasing crime neighborhoods.

If we have **two** predictors as in Eq. 2.2, we have to remember that we are now controlling for the other predictors, as in multiple regression. So, in that equation

G01 reflects differences in mean neighborhood fear between increasing vs. steady or decreasing violent crime neighborhoods, **while controlling for neighborhood differences in stability**. This is an **estimate of what is happening in the population**.

G02 reflects differences in mean neighborhood fear between stable vs. unstable neighborhoods, **while controlling for differences in changing violent crime rates**. This is an **estimate** of what is happening the the **population**

If your Level-2 predictors are <u>not</u> dummy variables, you interpret these as you would an unstandardized regression coefficient (b weight) in OLS regression: how much of a unit Y (B0J) changes for each unit change on X (W1), while controlling for differences in other predictors (W2, W3, and so on).

So it is important to remember the metric of each of your W variables.

And the last term in our equation is:

U0J the unique effect of each J on the mean outcome score, after controlling for whatever Level-2 predictors (W1, W2, etc.) have entered into the model for B0J.

Stated differently: these are neighborhood level effects on the outcome that you have not captured with your Level-2 predictors.

U0J therefore will change as you add in more Level-2 predictors. When you have no Level-2 predictors U0J will be the largest that it can be, because you are letting all between-group differences in the outcome go into this "unique" parameter. You are not trying to "model" between group differences in the outcome, but rather you are just seeing how big they are.

If you start to add Level-2 predictors to predict B0J, **and** those predictors are effective, the size of U0J will decrease. If those predictors are **not** effective, U0J will stay about the same size.

NUMERICAL EXAMPLE WITH FIRST LEVEL-2 EQUATION, PREDICTING INTERCEPTS

Say we are using NITFEAR as our outcome variable. Scores range from 1 to 4, with a higher score indicating higher fear at night.

B0J (mean outcome score for each Level - 2 unit)	G00	GO 1	W1 0=stable or decreasing violent crime 1=increasing	G02	W2 0 = 40% or more owner occupied; 1 = less than 40% owner occupied	U0J
3	1.5	.5	1	.25	1	.75
2	1.5	.5	0	.25	0	.5
1.15	1.5	.5	0	.25	1	-.60
2	1.5	.5	1	.25	0	0

As the above table shows, each neighborhood's intercept is the sum of its scores on

G00 +
slope of each predictor (G01...) * score on each predictor (W1...) +
U0J

NOTE:

B0J	differs for each neighborhood	-- estimate of neighborhood "true score"
		-- better estimates have higher reliability
G00	same for all neighborhoods	-- estimate of population "true score"
G01	same for all neighborhoods	-- estimate of population "true score"
G02	same for all neighborhoods	-- estimate of population true score
U0J	differs for each neighborhood	

LEVEL-2 EQUATION FOR SLOPES

Here is our second Level-2 equation:

B1J = G10 + G11(W1) + U1J   Eq. 2.3[2]

There will be one of these equations for **each** Level-1 predictor.

The outcome being predicted here are the different slopes, on X1, across the different neighborhoods.

This equation only makes sense if you **do** in fact have variation on your Level-1 slopes. This is tested by seeing if the variance of T11 is significantly larger than 0. If the slopes (B1J) for X1 are all the same across all the different neighborhoods, it does not make much sense to try and predict things here.

G10 Average slope of your X1 variable on your outcome, e.g., average age-fear slope, **for all neighborhoods that score 0 on all your Level-2 predictors**.

---

[2] This is B&R Eq. 2.4b, p. 14

This is an **estimate** for the **population of neighborhoods** from which the sample was drawn.

So if we have only one Level-2 predictor (increasing vs. stable or declining violent crime rate), and only one Level-1 predictor, age, then

G10 = average slope of fear on age in neighborhoods with a stable or declining violent crime rate.

If we have two Level-2 dummy predictors (1 = increasing violent crime rate, 0 = stable/declining rate; 1 = less than 40% owner occupied, 0 = more than 40% owner occupied), and the same Level-1 predictor

G10 = average slope of fear on age in neighborhoods with a stable or declining violent crime rate **and** at least 40% owner occupancy.

Both Level-2 predictors score at 0.

If your Level-2 variables are not dummy variables, the interpretation is still the same: the average slope of fear on age when all Ws = 0.

G11 Estimate of effect of W1 (Level-2 predictor) on the population slope (B11) of fear (Y) on age (X1) in the population of Level-2 units (neighborhoods).

You will have a G1n for each of your Level-2 predictors.

The interpretation depends on whether your Level-2 (W) predictors are dummy variables or not.

If we have just 1 W (increasing vs. stable or declining crime rate) and it is a dummy you would interpret as follows

G11 = mean difference in age-fear slope in stable or declining violent crime neighborhoods vs. increasing violent crime neighborhoods.

So the slope of fear on age in all the increasing violent crime neighborhoods would be:

G10 + G11*(1)

FI=HLML02.WP5

If we have two dummy variables - the crime change and owner occupancy variables decribed above

B1J = G10 +G11(W1) + G12(W2) + U1J

G11 = mean difference in age-fear slope in increasing vs. stable or declining crime neighborhoods **while controlling for** differences across neighborhoods in owner occupancy.

G12 = mean difference in age-fear slope in high vs. low owner occupancy neighborhoods **while controlling for** differences in changing violent crime

U1J unique effect of neighborhod J on the age-fear slope holding constant other Level-2 predictors (W1, W2, etc.)

The different U1Js across neighborhoods reflect between neighborhood differences in slope after taking account of between-neighborhood slope differences due to the Level-2 predictors entered into the equation.

The comment made earlier about U0J applies here as well: U1J will be largest, if between neighborhood slope differences exist, when there are **no** Level-2 predictors entered into the equation.

You assume the mean of U1J is 0.

VARIANCES AND COVARIANCES

PARAMETER	VARIANCE
U0J	T00
U1J	T11

COVARIANCE of U0J and U1J = T01

If you enter Level-2 predictors these become **conditional** or **residual** variances and covariances of B0J and B1J

G10 Same for each neighborhood -- estimate for population of Level-2 units
G11 Same for each neighborhood -- estimate for population of Level-2 units
U1J Different for each neighborhood

COMMENT ON PRECISION

In constructing Level-1 and Level-2 results, HLM uses a variety of estimation procedures (Bayesian, Maximum Likelihood, Generalized Least Squares, and so on) to generate the results. The results are different than straight OLS. See B&R p. 78.

In constructing Level-2 results, HLM uses precision weighting, which means it pays more attention to units where

-- the results are in more agreement; the group is more homogeneous; and

-- the results are based on larger samples

. See B&R Ch. 3 for more details than you probably want to contemplate.

PUTTING LEVEL-1 AND LEVEL-2 TOGETHER

In effect, the Level -1 and Level-2 equations **together** give us our prediction of YIJ. It is worth looking at this entire equation, because, among other things, it shows us how HLM decomposes error terms.[3] It also gives us a single prediction equation for the outcome for each individual.

In OLS we simply have residuals. We make certain assumptions about them, but do not "pull apart" the different factors contributing to them. HLM does otherwise.

Let's assume we have one Level-1 predictor, age, and that it has been centered by the neighborhood mean age. Also let's assume we have one Level-2 predictor, the dummy variable for stable or declining (0) or increasing (1) violent crime rates. Our outcome is one of our fear measures.

Well, here we go....

YIJ = score of individual I in group J on fear

G00 + mean fear in all neighborhoods with stable or declining violent crime rates

[3] This is Eq. 2.5, B&R, p. 14

G01*(W1) + mean fear difference in increasing vs. stable or declining violent crime neighborhoods. It is the slope difference * 1, or * 0

G10*(XIJ - X.J) + the average slope of fear on age in stable or declining violent crime neighborhoods * that individual's group centered score on age.

G11*W1*(XIJ - X.J) + the mean difference in the fear-age slope between increasing vs. stable or decreasing violent crime neighborhoods * that neighborhood's score on the crime change dummy (1 or 0) * that individual's group centered score on age.

U0J + the unique effect of neighborhood J on mean neighborhood fear, after controlling for between-neighborhood differences in changing violent crime. It is that neighborhood's difference from the overall fear mean that has not been modeled.

U1J*(XIJ-X.J) + the unique effect of neighborhood J on on the age-fear slope, after controlling for between neighborhood differences in the slope due to increasing vs. stable or declining violent crime rates * that individual's group centered score on age.

RIJ within-neighborhood, pooled between-person differences on the outcome score that have not been captured by parameters in the model.

COMMENT ON ERROR TERMS

In standard OLS we have a single error term: E. All of E ends up in residuals, that we then examine to be sure they meet our assumptions about the error terms.

Those assumptions are that the errors are **i.i.d.**[4] "Errors out to be normal i.i.d if they represent the sum of many small random influences."

i = errors have identical distributions at different levels of x, i.e., a mean of 0 and comparable variance

---

[4] Hamilton (1992) Regression with graphics. Monterey: Brooks-Cole, p. 32

i = errors are independent: "unrelated to X variables or to the errors of other cases" (p. 31)

d = errors are normally distributed

By contrast, in HLM, errors are treated differently (B&R p. 15). Here is how.

First, errors have been decomposed into the last three terms in the above equation:

U0J + U1J*(XIJ-X.J) + RIJ

Second, you explicitly anticipate **dependent** or **correlated** errors.

"Errors are dependent within each [neighborhood] because the components U0J and U1J are common to every [resident] within [neighborhood] J" (B&R: 15).

Third, you explicitly anticipate **unequal error variances** across your Level-2 units.

U0J + U1J*(XIJ-X.J) relies upon both U0J and U1J, and we already know these vary across neighborhoods, and will be bigger in some places than others. Therefore it is likely that the overall error terms will vary more in some neighborhoods than others.

EXPLORING DIFFERENT MODELS

THE ONE-WAY ANOVA[5]

Your Level-1 model is:

YIJ = B0J + RIJ

Your Level-2 model is:

B0J = G00 + U0J

Notice you have no predictors. You are merely decomposing variance in the outcome into between- and within-group levels. Nevertheless, HLM will give you a lot of information.

[5] B&R pp. 17-18; 61-64

FI=HLML02.WP5

G00 + or - 2 SE around G00	= maximum likelihood point estimate of grand mean of Y in the population
sigma (σ) squared	= estimate of person-level variance (W/I unit variance)
T00	= estimate of between unit variance = variance of U0J = variance of estimated true means (B0J) around the estimated true grand mean (G00)
Reliability (λ)	= averaged over all J Level-2 units, average reliability of the observed mean score on the outcome (Y.J) as an indicator of the true mean score (B0J) on the outcome in each J.

From the output you also can derive the

Intraclass correlation (ρ)	= percentage of outcome variance that is between Level-2 units (e.g., neighborhoods or schools) = agreement between all possible pairs of residents within one neighborhood, on the outcome score, averaged over all the neighborhoods = T00 / (T00 + σ squared)