2. The General Two-Level Model

*See Chapter 2 in Bryk & Raudenbush (2002)*

Formally, there are ***i*** = 1, …, *nj* level-1 units (e.g., students) nested within ***j*** = 1, …, ***J*** level-2 units (e.g., schools).

 Y*i* = β0 + β1X*i* + *ri* 🡺 Achievement*i* = β0 + β1 SES*i* + *ri* where *ri* ~ N(0, σ2)

#  Association between Achievement & SES is linear: Figure 2.1

 We could also center SES to make the intercept meaningful: *Figure 2.2*

 We could run the regression in a second school: *Figure 2.3*

 , where *rij* ~ N(0, σ2)

E(β0j) = γ0

Var(β0j) = τ00

E(β1j) = γ1

Var(β1j) = τ11

Cov(β0j, β1j) = τ01



Generalization of the Level-1 Model

  , where

 (*q* = 0,1, ..., *Q*)

*Xqij*

*rij*

*rij* ~ N(0, σ2) Var(*rij*) = σ2

 , and  , where

 *u*0*j*

*u*1*j*

*u*0*j* ~ N(0, τ00) and *u*1*j* ~ N(0, τ11) with covariance τ01.

****

Generalization of the Level-2 Model

 , where

γ*qs* (*q* = 0,1,…,*Sq*)

*Wsj*

*uqj*

*Var*(*uqj*) = τ*qq*

*Cov*(*uqj, uq′j*) = τ*qq′*

Note that each level-1 coefficient can be modeled at level-2 as one of three general forms:

1. a fixed level-1 coefficient; e.g., β*qj* = γ*q0*,

2. a non-randomly varying level-1 coefficient, e.g., 

3. a randomly varying level-1 coefficient, e.g., β*qj* = γ*q*0 + *uqj* or a level-1 coefficient with both non-random and random sources of variation, .

Simpler Submodels from ANOVA and Regression

The simplest possible HLM is equivalent to a one-way ANOVA with random effects.

 *Yij* = β0*j* + *rij* , where *rij* ~ N(0, σ2).

β0*j* = γ00 + *u*0*j* , where *u*0*j* ~ N(0, τ00).

*Yij* = γ00 + *u*0*j* + *rij*

Var(*Yij*) = Var(*u*0*j* + *rij*) = τ00 + σ2

The intraclass correlation coefficient is  .

This model is an **unconditional model** and is an important tool for comparison with conditioned models with explanatory variables at either or both levels.

Other submodels are described in the text, including

* 1. Means-as-outcomes regression
	2. One-way ANCOVA with random effects
	3. Random-coefficients regression model
	4. Intercepts- and slopes-as-outcomes
	5. A model with nonrandomly varying slopes

Centering – Location of *X* and *W*

With HLM, the outcomes at Level-2 are “derived” from the data and as such, may not be meaningful in a precise and practical way. Selecting the location of our variables secures clear understanding of all estimated effects.

# Level-1 intercept

In the usual regression, β0 is the intercept, the point on the Y axis where X is zero – or the value of the outcome variable when there are no explanatory variables or when all explanatory variables are zero.

In our earlier example of achievement and SES, Y*ij* = β0*j* + β1X*ij* + *rij* , the intercept is the expected achievement for students in school *j* who have a value of zero on SES.

If we are to build a model to account for variation in β0*j* , we must be precise about what a zero on SES means. Perhaps our SES scale has a meaningful zero; many variables do not. In many cases, there are no observations of students who have zero SES. So the mean achievement for students with zero SES is not meaningful. This is common with many variables, including GRE scores (which range between 200 and 800), GPA (is a GPA=0 informative?), etc. In HLM, centering is often needed to secure stability in estimating the models.

# Level-2 intercepts

Similarly, interpretation of the intercepts in the Level-2 models depends on the location of the W*j* variables. Stability of estimation is, however, unaffected by their location. Centering will certainly improve interpretation of results.

# Grand-Mean Centering

When we center a variable around the grand-mean, we subtract the grand-mean from each observed value . With grand-mean centering, the intercept is the expected outcome for a subject whose value on the explanatory variable X is equal to the grand-mean; the value  is zero when the observed value *Xij* is equal to the grand-mean. Now, Var(β0*j*) = τ00 is the variance among the Level-2 units in the adjusted means (adjusted to accommodate the grand mean).

# Group-Mean Centering

The other option is to center the explanatory variable values around their corresponding Level-2 unit means . In this case, the intercept is the unadjusted mean for group *j*; . This removes the adjustment made to the group mean by the explanatory variable scale on X. Now, Var(β0*j*) is simply the variance among the Level-2 unit means, .

# Centering Dummy Variables

Consider a variable coded such that *Xij* takes on a value of 1 if subject *i* in school *j* is participating in free and reduced priced lunch (FRL) and 0 if not FRL. In this case, the intercept (β0*j*) is the expected outcome for a non-FRL student in school *j*. In addition, Var(β0*j*) = τ00 is the variance in non-FRL outcome means across schools.

If we grand-mean center this dummy variable, this can take on two values.

If the subject is FRL (*Xij* = 1), then  will equal the proportion of non-FRL students in the sample (one minus the proportion of ones or FRL students).

If the subject is non-FRL (*Xij* = 0), then  will equal minus the proportion of FRL students in the sample (zero minus the proportion of FRL).

As earlier, the intercept (β0*j*) is the adjusted mean outcome in school *j*; adjusted for differences among schools in the proportion of FRL students.

If we group-mean center this dummy variable, this can also take on two values.

If the subject is FRL (*Xij* = 1), then  will equal the proportion of non-FRL students in school *j* (one minus the proportion of ones or FRL in school *j*).

If the subject is non-FRL (*Xij* = 0), then  will equal minus the proportion of FRL students in school *j* (zero minus the proportion of FRL in school *j*).

Now as earlier, the intercept (β0*j*) is the average outcome for school *j*, .

**Centering Predictors in HLM**

# School 1

 **Math**

 **Achievement**

 

 

  **SES**

 

# School 2

 **Math**

 **Achievement**

 

 

  **SES**

  

1.  Uncentered
2.  Grand-mean centered
3.  Group-mean centered