Intercept-Only Model and Random Effects ANOVA

Group Means and Individual Deviations

Consider a special regression equation for a particular group in which there are no predictors. For this regression equation, we have some number of individual scores (e.g., 10 students in a particular classroom), which we could say were measured at the micro level. We could call this special regression equation an "intercept-only" model.[1]

$$Y_{ij} = \beta_{0j} + R_{ij}$$

where Y_{ij} is the dependent variable, β_{0j} is the (unstandardized) intercept, and R_{ij} is the error or residual.[2] The subscript i is the index for an individual within a group, and j is an index for a particular group. β_{0j} can be thought of as the mean of each group. So, the equation shows that an observed value, Y_{ij}, is a function of the group mean, β_{0j}, and some difference in value or deviation from that group mean, R_{ij}. Notice that the R_{ij} values are simply $Y_{ij} - \beta_{0j}$ and are parallel to the first step of computing the variance, $Y_i - \bar{Y}$, or more relevantly here, are parallel to the first step used to compute the within-group sum of squares $Y_{ij} - \bar{Y}_j$ in analysis of variance (ANOVA).[3] In the text, the variance of R_{ij}, is referred to as σ^2 (i.e., $\mathrm{var}(R_{ij}) = \sigma^2$).

Grand Means and Group Mean Deviations

Because there is one of these equations for every group, as emphasized by the j subscript, we could theoretically use the intercepts from each group in another regression equation. In actuality, this is not really done in the estimation of MLR models, but we could save out these values if we wanted. This equation is for the macro level, because we now have an aggregated mean of individual scores for each particular classroom as the outcome. The micro and macro levels are commonly referred to as level 1 and level 2, respectively.

$$\beta_{0j} = \gamma_{00} + U_{0j}$$

In this level-2 equation, β_{0j} is the dependent variable, γ_{00} (lower case gamma) is the level-2 intercept, and U_{0j} is the level-2 residual. This level-2 equation is exactly parallel to the level-1 equation but uses different symbols, with γ_{00} representing the grand mean or the mean of the intercepts (i.e., the mean of the group means or *grand mean*). Observe that the subscripts for the terms on the right-hand side of the equation all start with 0 because β_{0j} is on the left. γ_{00} has two zero subscripts, where the first refers to the fact that β_{0j} is on the left and the second is for its role as the intercept in the equation.

Because each level-2 residual, U_{0j}, is the deviation of each group mean from the grand mean, the variance of U_{0j} is just referring to the variance of all of the group means around the grand mean. The variance of the intercepts, $\mathrm{var}(U_{0j})$, is referred to as τ_0^2 (tau-squared) by Snijders and Bosker (2012).[4] If $U_{0j} = \beta_{0j} - \gamma_{00}$, then U_{0j} parallels the deviation of group means from the grand mean, $\bar{Y}_j - \bar{Y}_{..}$, used in computing the sum of squares for the treatment effect (or A effect) in one-way ANOVA. When the deviation of means from the grand mean (U_{0j}) is large, there are large group differences, and, if we test the intercept variance for significance, it can tell us whether groups differ significantly. Thus, it can be shown that our model is really the same as the ANOVA model (Raudenbush, 1993). One difference from traditional ANOVA is that, in multilevel models, we assume that the groups are drawn from a larger

---

[1] In other contexts, such as logistic regression, this might be called an empty or null model.

[2] I follow the Snijders and Bosker (2012) notation. β_{0j} (and later β_{1j}) is unstandardized, and, in a non-nested case, refers to unstandardized coefficients as they are used and interpreted in ordinary least squares regression. Also, be sure to remember that R_{ij} in this context is not at all referring to multiple *R*-squared. Raudenbush and Bryk (2002) use lower case r_{ij} for the residual, which also would potentially cause confusion.

[3] Remember that the sample variance formula is $\sum\left(Y_i - \bar{Y}\right)^2 / \left(n-1\right)$ and the within group sum of squares is $\sum\sum\left(Y_{ij} - \bar{Y}_{.j}\right)^2$.

[4] Be aware that in their notation, Raudenbush & Bryk (2002), the symbol for intercept variance is τ_{00} with no superscripted 2.

population of groups, so this simple multilevel model is really equivalent to a *random effects* ANOVA. In my courses until this point, we have only considered *fixed effects* ANOVA.

We can rewrite our two MLR equations as a single equation, if we plug in the level-2 terms into level-1 equation.

$$Y_{ij} = \gamma_{00} + U_{0j} + R_{ij}$$

In this equation, γ_{00} represents the grand mean, the U_{0j} term represents the deviations from the mean (or mean differences or treatment effect, if you would like), and R_{ij} represents the within-group variation, or error variation. In more mathematically oriented ANOVA textbooks, you will see an ANOVA model written something like:

$$Y_{ij} = \mu + \alpha_j + e_{ij}$$

In the ANOVA model μ is the grand mean, α_j represents deviation of group means from the grand mean (between-group variation), and e_{ij} represents the within group deviation. If the model is a random effects ANOVA, the equation is the same, but α_j is assumed to have a random distribution instead of representing fixed deviations from the mean. I will have more to say about the random and fixed distinction in the subsequent handout "Distinguishing Between Random and Fixed: Variables, Effects, and Coefficients."

When we add predictors to the level-1 equation, they are covariates and the model becomes a random effects analysis of covariance (ANCOVA) in which the means are adjusted for the covariate.[5] We will talk about and illustrate the multilevel regression equivalent of ANCOVA in more detail later.

References

Raudenbush, S.W. (1993). Hierarchical linear models and experimental design. In Lynne K. Edwards (Ed.), *Applied analysis of variance in behavioral science* (pp. 459-496). New York: Marcel Dekker.
Raudenbush, S.W., & Bryk, A.S., (2002) *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage.
Snijders, T.A.B., & Bosker, R.J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling (2nd Edition)*. London: Sage.

[5] *If* the slopes are nonvarying, because ANCOVA usually assumes slopes are equal for all groups (i.e., the homogeneous or parallel slopes assumption). In the MLR context, the slopes would be said to be nonvarying in this case, but we will see later that this can be relaxed.