Cornell Statistical Consulting Unit

Variance Estimates of Zero in Multilevel Models

Francoise Vermeylen

1 Introduction

Multilevel models, also called mixed or hierarchical models, allow us to carry out regression analyses when we have data collected at two or more levels. These models contain both fixed and random effects. Common examples include data collected on children within schools and randomized block designs in agriculture. In the former, there is random variation both among schools and among children within a school. In the latter, there is random variation from block-to-block as well as variation within blocks.

In multilevel models, the variance at each level is estimated. For example, if we are studying students' test scores, then a multilevel model may involve the student-to-student variance of scores at a specific school as well as the school-to-school variation in scores. When fitting a multilevel model, some software packages may output estimated variances equal to zero. This newsletter addresses what it means to have variance estimates equal to zero.

2 An example in education research

Let us consider an example looking at the standardized test scores obtained by 4th graders across several schools. These test scores depend on the individual child characteristics, but they can also be influenced by individual school characteristics. Stated differently, test scores of children in the same school are not independent of each other. The total variance of the test results of the children from all of the schools taken together can be decomposed in two parts: the variance among schools and the variance among the children within a school. So it is important when running a regression on these test results to account for the school-to-school variability. This is done by including school as a random effect. In doing so, the output will provide two variance estimates: the variance among schools and the variance within schools or "residual" variance. These are the estimates of variance not explained by fixed effects at the school and child levels.

3 Variance estimates equal to zero

It is possible to end up with a school variance estimate of zero. This fact often puzzles the researcher since each school will most certainly not have the same mean test result. An estimated among-school variance being zero, however, does not mean that each school has the same mean, but rather that the clustering of the students within schools does not help explain any of the overall variability present in test results. Stated another way, the average within-school variance can be larger than the total variance of the scores across all schools, leading to an estimate of residual variance that is equal to the total variance in the data. In this case, all of the variance in the data is attributed to within-school (residual) variance, and the between-school variance is estimated as being equal to zero. This suggests that school-to-school variation may be negligible and the test results of students can be treated as independent of each other whether or not they are from the same school.

The "technical details" section at the end of this document presents a numerical example of this phenomenon.

4 Technical details

Let us look at the following hypothetical two-level dataset on students' test results. There are three schools, with four children in each school. In this example, the mean test results vary greatly across schools, although the scores vary little within schools . Table 4.1 displays the within-school mean and variance of the test scores, as well as the mean and variance of the within-school means.

Table 4.1: Within-school and overall means and variances for hypothetical test scores in three schools with 4 observations per school.

	Mean Score	Variance of Scores
school 1	4.0	10.0
school 2	12.5	9.67
school 3	33.25	9.58
overall	mean of means = 16.58	mean of variances =9.75
	variance of means = 226.4	

The overall mean of the 12 observations, across the schools, is 16.58, and the variance of all 12 individual scores is 172.63. Running this as a two-level model with school as a random effect will yield a residual variance parameter estimate of 9.75 and an among-school variance estimate of 223.96. By definition, the residual variance is the same value as the mean of the within-school variances. The among- school variance estimate (223.96) is almost identical to the variance of the means (226.4). This difference is due to the fact that the within-school variance also slightly contributes to the among-school variance. The among-school variance is actually estimated by the variance of the means minus the within school variance (9.75) divided by the number of children within a school (4).

Consider a second hypothetical example with four students in each of three schools. The overall and within-school means and variances for this example are given in Table 4.2

Table 4.2: Within-school and overall means and variances for hypothetical test scores in three schools with 4 observations per school. In this example, the school-to-school variation is much lower.

	Mean Score	Variance of Scores
school 1	12.75	202.91
school 2	15.75	192.92
school 3	21.25	187.58
overall	mean of means = 16.58	mean of variances =194.47
	variance of means = 18.58	

In this example, the overall variance of the 12 scores, across all three schools, was 172.63. The overall mean and variance, as well as the mean of means, are the same as in the previous example. In contrast to the previous example, however, the mean test results vary much less across schools, and the within-school variances are much larger. The variances for each school are now so large that the average within-school variance (194.47) is greater than the total variance in the dataset (172.63).

Running the same two-level model on this data set now yields a residual variance parameter estimate of 172.63 and an among-school estimate of 0. Remember that by definition, the residual variance should equal the mean of the variances, but in this case, the mean of the variances (194.47) is larger than the total amount of variation in the data set (172.63). Since the residual variance cannot be more than the total variance, the residual variance is set to equal the total variance of 172.63.

Similarly, the among-school variance is estimated by the variance of the means (18.58) minus the within-school variance (194.47) divided by the number of children within a school (4). Since this estimate is negative and variances cannot be negative, the among-school variance is reported as being zero.

As this second example illustrates, an estimated among-school variance being zero does not mean that each school has the same mean, but rather that the clustering of the students within schools does not help explain any of the overall variability present in the data. In this second dataset, one might thus think of the students as independent and not clustered within schools and run the model accordingly.

Created April 2006. Last updated April 2022.