Cornell University
Cornell Statistical Consulting Unit

StatNews #69

Variance Estimates of Zero in Multilevel Models

April 2006
Updated 2012

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. Researchers are often puzzled when variance estimates are reported as zero. This is often accompanied by an intriguing message from the software informing that some matrix is not positive definite and by the researcher's exclamation but I know that there is some variance at that level. This newsletter addresses what it means to have variance estimates equal to zero.

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, labeled school in the output, and one, labeled residual, for the variance within school. These are the estimates of variance not explained by fixed effects at the school and child levels.

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. In this case, test results of students can be considered as all independent of each other regardless if they are from the same school or not.

Detail on how the mixed procedure estimates both among- and within-school variability goes beyond the scope of this newsletter and is explained in a separate document that can be found in the below appendix.

If you need additional help with this, feel free to contact the CSCU.

Author: Francoise Vermeylen < fmv1@cornell.edu>

(This newsletter was distributed by the Cornell Statistical Consulting Unit. Please forward it to any interested colleagues, students, and research staff. Anyone not receiving this newsletter who would like to be added to the mailing list for future newsletters should contact us at cscu@cornell.edu. Information about the Cornell Statistical Consulting Unit and copies of previous newsletters can be obtained at http://www.cscu.cornell.edu).

Appendix: StatNews #69 Examples

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 .

School	Response	Mean:	Variance:
1	0		
1	3	4.0	10.0
1	6		
1	7		
2	9		
2	11	12.5	9.67
2	14		
2	16		
3	29		
3	33	33.25	9.58
3	35		
3	36		
		Mean of Means: 16.58	Mean of Variances: 9.75
Overall Mean	16.58	Variance of means:	
Variance	172.63	226.40	

We can calculate an overall mean (16.58) and variance (172.63) across these 12 students. We can also calculate a mean and variance for each school and then take again the mean of them. Since the data are completely balanced, the mean of means is the same as the overall mean, 16.58, but the mean within-school variance is 9.75. Additionally, we can also calculate the variance of the three school means (226.4).

Running this as a two-level model with school as a random effect will thus yield a residual variance parameter estimate of 9.75 and an among-school 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 now the following dataset with again four students in each of three schools. This data set contains the same responses as in the previous example but the responses have been distributed differently across the schools such that scores within a school are no longer similar:

School_2	Response
1	0
1	7
1	11
1	33
2	3
2	9
2	16
2	35
3	6
3	14
3	29
3	36
Overall Mean	16.58
Variance	172.63

Mean:	Variance:
12.75	202.91
15.75	192.92
21.25	187.58
Mean of means: 16.58	Mean of Variances: 194.47
Variance of Means: 18.58	

First, notice that 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 though, the mean test results vary much less across schools, and the scores vary much more within schools.
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.

Likewise, 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. A likelihood ratio test would probably show that school_2 is not significantly different than zero. In this second dataset, one might thus think of the students as independent and not clustered within schools and run the model accordingly.