

Scenario: Fitting a Linear Mixed Model with Nested Classification

Suppose you are an analyst working with a large school district. The district wants to study four sets of instructional materials and their relationship to test scores on a standardized test. The data is organized first by material. The four materials are identified as A, B, C, and D. The district randomly selects 20 teachers to use the materials in their classes. However, each teacher does not use all of the materials. Instead, the 20 teachers are divided into four groups of five teachers each. The teachers in each group are assigned to one of the four sets of materials. Each teacher uses the assigned material in a class consisting of six randomly selected students. The study data is stored in the scores data set.

A few sample observations are shown here. Notice that the observations represent students. The variables are **Material**, **Teacher**, **Student**, and **Score**. Let's take a closer look at how the data is classified and the role of each variable in the study.

The response variable is **Score**, which stores the students' scores on the standardized test. The two predictor variables (that is, the two factors) are **Material** and **Teacher**. In the **Material** variable, the instructional materials are coded A, B, C, or D. **Teacher** stores a number for each teacher selected for the study: 1 to 5 for each type of material. The **Student** variable stores a number for each student selected for the study: 1 to 6 for each teacher. The **Student** variable is not included in the model. Instead, **Student** represents the observations that are selected from the data set. Note that the values for **Teacher** (1 to 5) and **Student** (1 to 6) are arbitrary. They could be coded with any other values, if the teachers within a material, and the students within a teacher, have distinct values. This data is an example of nested classification because teachers are nested within materials.

Material is considered a fixed effect because only four materials are used in the study and you are interested only in making inferences about these four materials. Is **Teacher** a fixed effect or a random effect? **Teacher** is considered a random effect because the twenty teachers in the study are randomly selected from a population of teachers. Teachers are nested within materials. The purpose of the study is to accomplish the following: estimate and compare the treatment means (that is, the **Material** means) over the entire population of teachers, and account for the variability in the response variable (**Score**) due to the **Teacher** variance.

To see details about the **scores** data set, click the Information button.