



## StatNews #02

### Multiple-Level Data June 1996

Recently, increasing attention has been paid to models for describing data that are collected at several levels. These models are called multilevel, hierarchical, or mixed models. For example, data may be collected on characteristics of children (1st level), classrooms (2nd level), and schools (3rd level). The researcher may be interested in how the outcome variable measured for each child depends on characteristics of the child, the classroom, and the school. Factors at the child level may include sex, race, and socioeconomic status. Factors at the classroom level may include teaching method and teacher's level of experience. Factors at the school level may include location and type of school.

An interesting example of multilevel data collected by Sue Rhee in Tom Brenna's lab comes from a study in which subjects are randomly assigned to one of three diets. Fatty acids are measured at the beginning of the study and at four subsequent weeks. In order to deal with potential measurement error, two measurements of fatty acids are taken for each subject at each time period. The study evaluates how two measurements of fatty acids taken on the same day for the same subject vary (1st level), how the fatty acids for each subject vary over time (2nd level), and how the fatty acids vary across subjects depending on the diet (3rd level).

In multilevel data, a different source of variability is present at each level. In this second example there are three sources of random variation in the outcome: 1) the variation between two measurements taken on the same subject at the same time, 2) the variation across time for each subject, and 3) the variation among subjects, after accounting for the effect of diet. The researcher may be interested in estimating how much of the total variation comes from each source.

Split-plot and repeated measures experiments and longitudinal studies are familiar designs that produce multilevel data. In the past, analysis of data from such designs was problematic if there were continuous factors or missing data. Specialized software for estimating multilevel models has been made available over the past few years. Multilevel models estimated using this software can account for different factors and different sources of variability at each level. A future newsletter will discuss and compare this software.

Statistical Consulting staff can help you develop and implement a model for analyzing multilevel data. We have also researched a number of publications that address this topic in more detail. A particularly clear and non-technical article, "An Introduction to Hierarchical Linear Models," by Carolyn L. Arnold (Measurement and Evaluation in Counseling and Development, July 1992, pp. 58-90), is available from anyone in the Statistical Consulting office.

Author: Cara Olsen