

StatNews #05

Software for Multilevel Models

July 1996

Updated May 2012

In many studies, data are collected at several levels, and a different source of variability is present at each level. Multilevel models are useful for analyzing these data. Our newsletter of June 3, 1996 ([StatNews #02](#)) outlined when a multiple level analysis is appropriate.

Several software packages are available for fitting multilevel models. The choice depends on the type of response variable (i.e., continuous or categorical), the complexity of the model (i.e., number of levels, variance structure, and presence of covariates at one or more levels), and whether the model is linear or nonlinear.

For a continuous response variable and linear model (the multilevel version of linear regression or analysis of variance), we generally recommend SAS PROC MIXED. It is part of the SAS/STAT software, the syntax is familiar to users of SAS PROC GLM, and it accommodates a variety of variance structures. We have available two specialized multilevel programs, MLn and HLM, which may be able to solve certain problems that PROC MIXED cannot model. Other specialized packages, including BMDP and VARCL, can also perform this type of analysis. SAS can also estimate non-linear multilevel models using the NLINMIX macro in version 6.11.

When the response variable is categorical (i.e., binary, nominal, or ordinal), the multilevel version of logistic or probit regression may be appropriate. Charles McCulloch of the Biometrics Unit gave a presentation on this topic in June. Based on his review, the best software for this purpose is MIXOR. MIXOR may be downloaded free of charge from the World Wide Web at <http://www.uic.edu/~hedeker/mix.html>.

This package is available for Windows and requires the PKUNZIP utility for installation. MIXOR can fit logistic or probit regression models for binary or ordered categorical responses, but it can only estimate a two-level model. Both SAS (using the GLIMMIX macro, available with version 6.11) and MLn can fit more complex models with categorical responses. However, we hesitate to recommend these packages for logistic regression, since the estimation method they use can give biased estimates of regression parameters.

We can advise you about organizing data for use with these programs, and help with the proper syntax for fitting your multilevel model

Update 2012: most commonly used statistical software packages have now procedures for multilevel models with continuous and categorical response variable.

Author: Cara Olsen

(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>).