



## StatNews #32

### Generalizations of Logistic Regression November 1998

Researchers are often interested in performing regression when the response variable is categorical. Logistic regression may be used when the response variable has only two categories. When the response variable has more than two categories, methods that extend the technique of logistic regression are available. The choice of method depends on whether the response variable is measured on an ordinal or nominal scale. This newsletter gives examples of nominal and ordinal response variables and describes techniques that can be used to analyze these kinds of data.

A nominal scale has categories that are not ordered. For example, an education researcher may be interested in a nominal response variable such as what each student decides to do after high school: attend college, attend vocational school, find a job, join the military, or something else. Logistic regression for a nominal response variable is called nominal, multinomial, or polytomous logistic regression.

An ordinal scale has categories that are ordered. Letter grades for a particular class (A, B, C, D, or F) and academic tracks (remedial, regular, or advanced placement) are examples of ordinal variables. While nominal logistic regression can be used to analyze ordinal response variables, it is often not the best choice because it ignores the ordering of the categories. An alternative is to fit a proportional odds model, which assumes that the effect of the explanatory variables on movement between categories of the response variable is the same across the entire scale. This approach is often referred to as ordinal logistic regression.

Nominal and ordinal logistic regression can be performed in several different statistical software packages including Minitab and SAS. SAS uses different procedures for the two types of logistic regression; PROC LOGISTIC for ordinal logistic regression and PROC CATMOD for nominal logistic regression. SAS users should be careful when interpreting results from PROC CATMOD, since it treats explanatory variables in a different way than most other packages do. SPSS does not currently perform logistic regression for response variables with more than two categories. However, the logit loglinear models procedure in SPSS can do nominal logistic regression if all of the explanatory variables are categorical.

If you need to implement logistic regression or interpret the results, statisticians in the CSCU can help you. Feel free to contact one of us.

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](mailto:cscu@cornell.edu). Information about the Cornell Statistical Consulting Unit and copies of previous newsletters can be obtained at <http://www.cscu.cornell.edu>).