Multinomial logistic regression is used to model nominal outcome variables, in which the log odds of the outcomes are modeled as a linear combination of the predictor variables.

This page uses the following packages. Make sure that you can
load them before trying to run the examples on this page. If you
do not have a package installed, run:
install.packages("packagename"), or if you see the version is out of
date, run: update.packages().

```
require(foreign)
require(nnet)
require(ggplot2)
require(reshape2)
```

Multinomial Regression with R

Examples of multinomial logistic regression

Example 1. People's occupational choices might be influenced by their parents' occupations and their own education level. We can study the relationship of one's occupation choice with education level and father's occupation. The occupational choices will be the outcome variable which consists of categories of occupations.

Example 2. A biologist may be interested in food choices that alligators make. Adult alligators might have different preferences from young ones. The outcome variable here will be the types of food, and the predictor variables might be size of the alligators

Description of the data

For our data analysis example, we will expand the third example using the hsbdemo data set. Let's first read in the data.

```
ml <- read.dta("http://www.ats.ucla.edu/stat/data/hsbdemo.
```

The data set contains variables on 200 students. The outcome variable is prog, program type. The predictor variables are social economic status, ses, a three-level categorical variable and writing score, write, a continuous variable. Let's start with getting some descriptive statistics of the variables of interest.

Multinomial Regression with R

```
with(ml, table(ses, prog))
prog ses general academic vocation low 16 19 12
middle 20 44 31 high 9 42 7
with(ml, do.call(rbind, tapply(write, prog,
function(x) c(M = mean(x), SD = sd(x)))))
M SD general 51.33 9.398 academic 56.26 7.943
vocation 46.76 9.319
```

Multinomial Regression with R

Analysis methods you might consider

- ▶ Multinomial logistic regression, the focus of this page.
- ▶ Multinomial probit regression, similar to multinomial logistic regression with independent normal error terms.
- ▶ Multiple-group discriminant function analysis. A multivariate method for multinomial outcome variables
- ▶ Multiple logistic regression analyses, one for each pair of outcomes: One problem with this approach is that each analysis is potentially run on a different sample. The other problem is that

Multinomial Regression with `R`

- ▶ Collapsing number of categories to two and then doing a logistic regression: This approach suffers from loss of information and changes the original research questions to very different ones.
- ▶ Ordinal logistic regression: If the outcome variable is truly ordered and if it also satisfies the assumption of proportional odds, then switching to ordinal logistic regression will make the model more parsimonious.
- ▶ Alternative-specific multinomial probit regression, which allows different error structures therefore allows to relax the IIA assumption. This requires that the data