**Ordered, Multinomial and Multivariate Logits and Probits:**

**A Cheat Sheet**

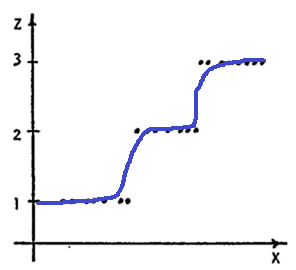
**The Basics**

Discrete dependent variable: A logit or probit, only with three of more outcomes.

Examples:

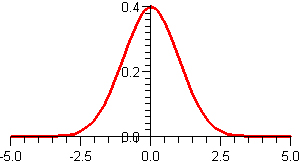
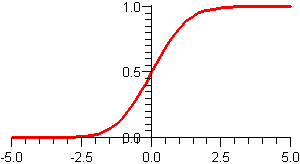
* A parliamentary election (or the U.S. 1992 presidential election with Ross Perot)
* Letter grades (A,B,C,D,F)
* Choice of transportation (car, bus, walk, bike)

Visual of three-outcome logit or probit



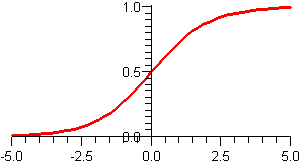
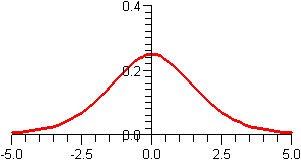
**Differences between logit and probit**

* Basic differences
  + Logit errors follow standard logistic distribution, while probit errors are assumed to follow the standard normal distribution (PDF = probability density function, CDF=Cumulative Density Function)



CDF of St. Normal Distribution (probit)

PDF of the St. Normal Distribution (probit)



PDF of the St. Logistic Distribution (logit)

CDF of the St. Logistic Distribution (logit)

* So, what
  + In other words, the PDF of a logit has a lower peak and fatter tail than the PDF of the probit. Additionally, the CDF of the logit is less steep than that of the probit, taking longer to shift from the values of zero to one.
    - So what the heck does these mean? Essentially, they PDF and CDF differences leak to different estimates, as the logit model has a larger estimate due to the fatter distribution. However, even with this difference, the coefficients almost end up the same.
* Differences in logits and probit for multinomial and multivariate models
  + For logics with logistics distribution, random components are independently and identically distributed.
    - This means that individuals with identical observable characteristics have identical tastes, and that any effects of unobservable characteristics of individuals of alternatives are uncorrelated. A very rigid approach.
  + For probits with normal distribution, tastes are allowed to vary among individuals with identical observable characteristics, and effects of unobservable characteristics are allowed to be correlated across alternatives.
    - Computationally this is more difficult, but with Stata, not as hard as it used to be.

**Ordered Logit/Probit**

* Ordinal variables are one type of discrete dependent variable, with hierarchy or ranking.
  + Examples: life satisfaction (awesome, meh, not good), employment status (employed, underemployed, unemployed, skill level (high, medium, low).
* This model assumes effects of independent variables are constant between levels of dependent variable
  + Example: with grades, difference between A and B is the same as B and C

**Example Output (Probit)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **School Grade** | **Teacher Grade** | **Facilities Grade** | **Principal Grade** |
| Charter School | .59\*\*\* | .52\*\*\* | .54\*\*\* | .45\*\*\* |
|  | (.09) | (.10) | (-.40\*) | (.31) |
| Cutpoint 1 | -1.60 | -1.59 | -1.15 | -1.65 |
| Cutpoint 2 | -1.04 | -1.24 | -.53 | -1.30 |
| Cutpoint 3 | -.26 | -.50 | .38 | -.60 |
| Cutpoint 4 | .84 | .61 | 1.23 | .32 |

This output, from the charter school paper, shows that parents with children in charter schools rate their schools, teachers, principals, and their school’s facilities higher than public schools (st. errors in parentheses). However, as an ordered probit, there is no natural odds-ratio interpretation for the effects of the ‘charter school’ variable.

Parents of kids in charter schools are 59% more likely to rate their school higher than public schools, 52% more likely to rate their teacher better, etc.

The cutpoints (aka threshold parameters) are estimated to help match the probabilities associated with each discrete outcome. Thus, the odds to a school receiving higher grades is much higher than receiving lower grades.

**Multinomial Logit/Probit**

* This model differs from the ordered model because of a lack of inherent ranking or order in the categorical variable
  + Example: Red cars, green cars, and yellow cars are all still cars, but you would get different levels of utility from the different colors. Here, “red” will be chosen if “red”>”green” and “red”> “yellow”.
* More flexible than the ordered logit/probit, with different each category choice, except the reference category
* Setup: Right Side
  + Instead of one equation, you have J-1 equations, where J= the number of possible equations).
  + The explanatory and control variables remain the same
  + Because of multiple equations, the coefficients of different equations mirror the differing effects of the various choices.
* Setup: left side
  + One of the possible outcomes becomes the ‘base outcome’ from which all the other outcomes are compared
  + The dependent variable becomes a log equation that gives the probability
* Reading the output:
  + The base outcome is the natural comparison group, so it’s usually not a problem comparing dependent variables A to B, B to C, and A to C. Just look at the relative sign and size of the coefficients of the two comparisons you’re looking at.

**Example Output (Logit)** (Note: this is not the raw output, which would be in log-odds, but instead with the marginal effects command in use (code listed below) focused on a specific outcome within the multiple possible outcomes).

|  |  |
| --- | --- |
| **Variable** | **Coefficient (dy/dx)** |
| IQ | .182 |
| Sibling | -.02 |
| Mother’s Education | .043 |
| Father’s Education | .033 |

Here, the dependent variable is the probability of a person having some sort of college education, compared to their chances to dropping out of school. The relative probability of going to college increases by 1.82% for every IQ point above the average, and 4.3% and 3.3% for every year of education your mother and father have. Having a sibling reduces your probability of going to college by 2%.

**Multivariate Logits/Probits**

* Unlike ordered and multinomial, multiple binary outcomes are related but not mutually exclusive
  + Thus, there is a covariance matrix of variables that correlates outcomes
  + Potential problem: estimation of the model is problematic if the outcomes are being explained by characteristics which vary for each individual but not across the different outcomes.

**Example Output**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Men’s Accounts | | Women’s Accounts | | Joint Accounts | |
| Variable | Coefficient | Marginal Effect | Coefficient | Marginal Effect | Coefficient | Marginal Effect |
| Percent Earnings from Female Partner | -.158 | -.041 | .571 | .141 | .057 | .014 |

Here, because families could have a joint account as well as sole accounts, the outcomes are not mutually exclusive. Here, women’s income is shown to be associated with an increased likelihood that they had their account, negatively associated with a sole male account, and positively associated with a joint account.

Marginal effects are the weighted average of the effects calculated in each observation and combined across the multiple imputations.

**In Stata**

Ordered logit: *ologit*

Ordered probit: *oprobit*

Marginal Effects: *mfx: (example: mfx, predict (p outcome(1)) varlist (variable)*

Multinomial Logit: *mlogit*

Multinomial Probit: *mprobit*

Multivariate Logit:

**Resources**

Multinomial Logit and Probit: <http://espin086.wordpress.com/2010/08/20/multinomial-probit-regression-probability-of-going-to-college-given-intelligence-siblings-parents-education-and-iq/>

Also for Multinomial Logit and Probit , <http://www.indiana.edu/~statmath/stat/all/cdvm/cdvm6.html>

Multivariate Probit: <http://www.actuaries.org/AFIR/Colloquia/Stockholm/Young.pdf>

Limited dependent variables in general (*a great source for the basics):* <http://faculty.arts.ubc.ca/tlemieux/econ594/lecture6.pdf>

**Lastly, a quick list of different methods for different limited independent variables**

Binary response variable (y=0,1) = *probit, logit*

Truncated variable or censored variable (y>0) = *tobit*

Count variable (y=0,1,2) = *poission model*

Discrete choice variable (y=bus, train, car) = *multivariate/multinomial logit or probit*

Ordered (y=a,b,c) = *ordered logit or probit*