

ECONOMICS 6002 CLASS 21
DISCRETE CHOICE MODELS

1. Categories of limited dependent variable (LDV) models
 - a. Discrete Choice Models
 - i. Binary (Dichotomous) variables
 - ii. Polychotomous variables
 - b. Dependent variables subject to upper or lower limits
 - c. Most of these models cannot be consistently estimated using fitting criteria (e.g., ordinary or non-linear least squares), but can be consistently estimated using the maximum likelihood principle.

2. Binary dependent variable models
 - a. Estimating equation specifies the **probability** that the dependent variable = 1
 - b. Therefore functional form must be interpretable as a Cumulative Distribution Function (CDF), and be bounded between 0 and 1,
 - c. Linear probability model is deficient in that it does not satisfy this characteristic.
 - d. Interpretation as a CDF can be reinforced by interpreting the limited dependent variable as arising from a continuous (but unobserved) **latent variable**, as in the index function model or the random utility model.
 - e. The **probit** model is based on a normal CDF, and the **logit** model is based on a logistic CDF.
 - i. These are estimated by maximum likelihood.
 - ii. Because the model is non-linear, its coefficients are not easily interpretable. They must be scaled appropriately in order to obtain the **marginal effects** of the independent variables on the dependent variable.
 - iii. Estimates are not robust with respect to misspecification of
 - (1) error distribution
 - (2) omitted variables
 - (3) heteroskedasticityThese should be tested (LM tests are available) if their presence is suspected.

3. Multiple (unordered) choice models
 - a. These can be motivated by a random utility model, in which the choice that provides the highest utility is chosen.
 - b. The probit model involves evaluation of multiple integrals, and becomes difficult when there are more than three choices (but recent advances in the simulation of these integrals is promising).
 - c. The logit model scales to multiple choices more easily.
 - i. Major limitation is the Independence of Irrelevant Alternatives (IIA) characteristic: the odds of choosing one alternative over another depends on the attributes or parameters of the two choices only, and not those of the other choices.

4. Ordered choices
 - a. If the dependent variable is scaled ordinally, the order provides additional information which can be used for greater efficiency in estimation.
 - b. A latent regression model can motivate an ordered probit or logit model, in which the ordered dependent variable is an ordinal representation of the latent variable.
 - c. Maximum likelihood estimation will estimate the latent regression, along with the mapping of the latent variable into the ordered dependent variable.