

SOC 2020 Multivariate Statistical Methods II
Michael White
Spring 2009

DRAFT SYLLABUS

Course Objective

This course is a graduate-level introduction to multivariate regression models for categorical and limited dependent variables. The course is designed to provide students with the basic knowledge and skills needed to analyze social science data with categorical and limited continuous outcomes. The main objectives of the course are to: (1) introduce students to the family of logistic regression models for modeling binary, nominal, and discrete ordered outcomes; (2) introduce students to more advanced models for truncated distributions and for addressing issues of selectivity; (3) acquaint students with basic concepts and strategies in sample design and the use of sample weights in descriptive and multivariate analyses; (4) instruct students in current methods for dealing with missing data; and (5) teach students how to use the statistical software package STATA to estimate regression models for categorical and limited dependent variables. The course will place special emphasis on model selection and interpretation.

Prerequisites: Sociology 2010.

Course Structure

The course will meet once a week for a 3 hour lecture. The material is cumulative so it is essential that students do the required reading and exercises before each class. This course will be computer intensive with a significant portion of the weekly exercises involving the use of STATA. There is a required computer lab section for the course that meets once a week for two hours during the first half of the semester. Student in the course will often be writing an MA thesis or paper intended for publication. Hence session(s) for student presentation of this work-in-progress is reserved.

Student Evaluation

- (1) Weekly exercises (25% toward the final grade)
- (2) Mid-term Exam (25%) and Final Exam (25%)
- (3) Analysis paper due at the end of the semester (20%)
- (4) Class presentations (5%)

Week	Topic
1	Logistic Regression <ul style="list-style-type: none">• Why logistic regression• Multiple logistic regression model
2	Logistic Regression <ul style="list-style-type: none">• Odds ratios

- Estimating probabilities
 - Comparing the fit of models
- 3 Probit Regression, Multinomial Logistic Regression
- 4 Ordered Logistic Regression, Tobit Regression, Conditional Logistic Regression
- 5 Predicted Values
- Graphing
 - Intro to simulation
- 6 EXAM
- 7 Selection Models
- What is Selection and Why is it important
 - Heckman Selection Model
- 8 Hierarchical Regression Models: Introduction
- Context: concept and measurement
 - Assumptions and Sources of Bias
- 9 Hierarchical Regression Models: alternative formulations
- Adding a contextual variable
 - Context from within data
 - Alternative software and assumptions
- 10 Missing Data
- Assumptions and Sources of Bias
 - Strategies re Missing Independent Variable information
 - Strategies re Missing Dependent Variable information
- 11 Time Series Data
- Data Structure
 - Autocorrelation
- 12 Panel Data and repeated measures
- Origins of panel data
 - Link to Event history data
 - Error Structures and Estimation Techniques
- 13 Student Presentations
- 14 FINAL EXAM