SOC385K/ SSC385: Discrete Multivariate Models

Fall 2010 Unique # 45690/57635 MW 2:00-3:30 BUR 220

Instructor: Dan Powers

Office Hours: MW 10:00-11:00AM, F 1:00-3:00PM and by appointment 502 Burdine (232-6335)

Course Description

This course deals with regression models for discrete and categorical dependent variables. Categorical data analysis is appropriate when the dependent variable is measured as a count, or consists of events and trials, or when responses are in the form of binary, ordered, or unordered categories. Regression-like models for discrete and categorical outcomes are widely used in applied research. The methods for estimating and interpreting results from these model are different from those of classical linear regression, however there are many similarities. Students in this course should have some prior exposure to linear regression models. This class serves a wide range of students; the material presented here aims to be useful for students at all levels. Most of the material in the exercises and handouts is based on applied, as opposed to theoretical (i.e., mathematical statistical) problems. In keeping with the applied nature of this course, we will provide examples drawn mainly from sociological and demographic research.

Course Requirements

Depending on the needs and interests of students, we will spend 2–3 weeks on each of the subjects below. Grades will be based on scores from approximately 5 assignments or problem sets (50%) and a 15–20 page (double spaced) research paper (50%). Students who are enrolled in the MS in Mathematical Statistics program must complete the extra credit problems in the assignments. Paper proposals must be submitted for approval no later than Nov. 15. In lieu of a substantive paper, students may undertake an in-depth exploration of any methodological issues pertaining to the analysis of categorical data (i.e., a methods or applied statistics paper).

Topics

Topics covered in this course will include:

- an overview of the classical linear regression model
- models for binary data
- models for count data and contingency tables
- models for ordered and unordered categorical data

Extensions to the models above will also be examined, such as hierarchical/multilevel models for categorical responses, as well as treatment-effect and selection models.

Recommended Texts

Powers, Daniel A., and Yu Xie (2008) Statistical Methods for Categorical Data Analysis, 2nd Edition, London: Emerald.

Long, J. Scott, and Jeremy Freese (2001/2005) Regression Models for Categorical Dependent Variables Using Stata, College Station: Stata Press. (Edition is not important)

Course Materials

Handouts and data sets detailing specific topics will be placed on Blackboard. A web site with programming examples reproducing all the models in the Powers and Xie (2000, 2008) books can be found at:

https://webspace.utexas.edu/dpowers/www/

Useful Background Material

Particularly useful books on mathematical and statistical concepts are:

Fox, John (2009) Mathematical Primer for Social Statistics, Beverly Hills: Sage Publications.

Gill, Jeff (2006) Essential Mathematics for Political and Social Research, New York: Cambridge.

Syllabus

All readings are from Powers and Xie [PX] (2008) and Long and Freese [LF] (2001/2005).

Topic	Reading	Schedule
Introduction to Categorical Data Analysis — definition of categorical data — categorical response variables — transformational and latent variable approach	[PX] Ch. 1 [LF] Ch. 1	Week 1
Review of Linear Regression — regression models — model structure and assumptions — estimation, and inference in linear Regression	[PX] Ch. 2 [LF] Ch. 2, Ch. 8	Weeks 2–3
Models for Binary Data — the transformational approach — logit and probit models — the latent variable approach — model selection and model fit — estimation, inference, and interpretation — extending binomial response models	[PX] Ch. 3, 5 [LF] Ch. 3–4	Weeks 4–7
Models for Counts — Poisson regression models for counts — models for two-way tables — models for ordinal data — extending count models — models for rates	[PX] Ch. 4 [LF] Ch. 7	Weeks 8–9
Models for Ordinal and Nominal Dependent Variables — ordered logit and probit models — extending ordinal response models — the multinomial logit model — the conditional logit model — sequential logit models	[PX] Ch. 7–8 [LF] Ch. 5–6	Weeks 10-12
Selection Models, Switching Regressions, and Treatment Effects Models — partial observation, missing data and selection — parametric sample selection models selection — propensity score matching and related methods	handouts	Weeks 13-15