Econometric Models for Discrete Dependent Variables

CIDE Summer School in Econometrics
June 9 - June 14, 2008

Rainer Winkelmann

Description: This course provides an introduction to the analysis of the two most common types of discrete dependent variables, counts and ordered responses (of which binary responses are a special case). In discrete data models, we seek to predict the *probability distribution* of the dependent variable, conditional on covariates. Parametric and semi-parametric models are considered. Starting from the benchmark Poisson and (ordered) probit regression models, we consider generalizations arising from various departures, such as unobserved heterogeneity, measurement error, endogeneity and selectivity. We review theoretical results as well as applications in fields such as health economics, labor economics, population economics and empirical industrial organization.

Short Course Outline

Part 1: Count data

- Probability distributions for count data
- Specification and estimation of count regression models
- Unobserved heterogeneity
- Endogeneity
- Two-part models and zero-inflation

Part 2: Ordered Responses

- Basic concepts
- Generalized ordered response models
- Multivariate ordered response models
- Endogeneity & self-selection with copula functions

General Literature:

I have written a book on "Econometric Analysis of Count Data" (Springer, 5th ed, 2008).

Topic 1

Case study – German health care reform of 1997

> Winkelmann, R. 2004, Co-payments for prescription drugs and the demand for doctor visits - Evidence from a natural experiment, Health Economics, 13, 1081-1089.

A differences in differences model for count data

- Poisson regression
- Maximum likelihood, QML
- Unobserved heterogeneity in cross-sections and in panels

> Cameron, A.C. and P.K. Trivedi 1986, Econometric models based on count data: comparisons and applications of some estimators and tests, Journal of Applied Econometrics 1: 29-53.
>
> Gourieroux, C., A. Monfort and A. Trognon 1984, Pseudo maximum likelihood methods: Theory, Econometrica 52: 681-700.
>
> Hausman, J.A. , B.H. Hall and Z. Griliches 1984, Econometric models for count data with an application to the Patents-R&D relationship, Econometrica 52: 909-938.

Topic 2

Endogeneity in Count Data Models

- Simultanous equation models for counts
- Non-linear IV and two-stage methods
- Parametric models
- Panel data methods

> Mullahy, J. 1997, Instrumental Variable Estimation of Count Data Models: Applications to Models of Cigarette Smoking Behavior, Review of Economics and Statistics 79(4): 586-593.
>
> Rose, N.L. 1990, Profitability and product quality: Economic determinants of airline safety performance, Journal of Political Economy, 98, 944-964.

Terza, J.V. 1998, Estimating Count Data Models with Endogenous Switching: Sample Selection and Endogenous Treatment Effects, Journal of Econometrics 84(1): 129-154.

Windmeijer, F.A.G. and J.M.C. Santos Silva 1997, Endogeneity in Count Data Models: An Application to Demand for Health Care, Journal of Applied Econometrics 12(3): 281-294.

Topic 3

Multi-Index Models for Extra Zeros

- Marginal probability effects in Poisson and Negbin models
- Hurdle versus finite mixture
- Multi episode model
- Zero inflated Poisson

Winkelmann, R. 2004, Health Care Reform and the Number of Doctor Visits - An Econometric Analysis, Journal of Applied Econometrics, 19, 455-472.

Mullahy, J. 1986, Specification and testing in some modified count data models, Journal of Econometrics 33: 341-365.

Pohlmeier, W. and V. Ulrich 1995, An econometric model of the two-part decision making process in the demand for health care, Journal of Human Resources 30: 339-361.

Santos Silva, J.M.C. and F. Windmeijer 2001, Two-Part Multiple Spell Models for Health Care Demand, Journal of Econometrics 104: 67-89.

Topic 4:

Generalized Ordered Response Models

Boes, S. and R. Winkelmann (2006) Ordered Response Models, *Advances in Statistical Analysis*, 90, 165-179, 2006

Income and Happiness: New Results from Generalized Threshold and Sequential Models (with S. Boes), University of Zurich, Socioeconomic Institute Working Paper 0407.

Topic 5:

Multivariate Ordered Response Models;
Modelling endogeneity & self-selection using copula functions

Luechinger, S., A. Stutzer and R. Winkelmann (2007) Self-selection and subjective-well being: Copula models with an application to public and private sector work, mimeo.

Trivedi, P.K and Zimmer, D.M. (2007), Copula Modeling: An Introduction for Practitioners", Foundations and Trends in Econometrics, Volume 1, Issue 1.

Computer Lab

Sessions will include hands-on exercises using Stata and data used originally in the following papers:

- Cameron and Trivedi, 1986
- Rose, 1990
- Winkelmann, 2004