

# Data III: Generalized Linear Models

PSCI7108/SOCY7111  
Fall 2007

## Course Description

Time: Tuesday 3:00-5:50pm  
Location: Ketchum 33  
Instructor: Ying Lu  
Office: Ketchum 3E  
Office hours: Wed 10:30am-1:30pm or by appointment  
Email: Ying.Lu@colorado.edu

## Prerequisites

This course is designed for graduate students who have taken SOCY 5021, or PSCI 7095 or other quantitative classes at equivalent level. I will assume that students have a good understanding of the basic statistical concepts and are competent with linear regression analysis techniques. Students should also be able to manage data and conduct regression analysis with STATA. If you need to review some of the basic materials, you can refer to books such as,

- Allison, P. D. (1999). *Multiple Regression: A Primer*. Thousand Oaks, CA: Pine Forge Press.
- Moore, D. and McCabe George. (2004). *Introduction to the Practice of Statistics*. 5th Edition, Freeman.
- McClelland, G. (1999). *Seeing Statistics*. Duxbury Press. The online version [www.seeingstatistics.com/](http://www.seeingstatistics.com/) is available through any CU IP address, or via VPN.
- Hamilton, L. C. (2006). *Statistics with STATA*. Belmont CA: Duxbury.
- STATA Corp. *Getting Started with STATA 9*. copy available at 3 Ketchum.

## Textbooks

The textbook we will use for this class is

- Hoffmann, John P. (2004). *Generalized Linear Models: An Applied Approach*. Pearson: Allyn and Bacon.

In addition, I will make the lecture notes available online. If you would like to read more about generalized linear models, I recommend the following books which offer slightly different perspectives:

- McCullagh, P. and Nelder, J.A. (1989) *Generalized Linear Models* 2nd Edition. Chapman & Hall/CRC. (The "Bible" of GLMs).
- Hardin, James W. and Hilbe, Joseph M. (2001) *Generalized Linear Models and Extensions*, 2nd Edition. STATA press. (Covers a wider range of interesting models)
- Agresti, Alan (2002) *Categorical Data Analysis*, Wiley. (A view from contingency tables, more traditional social science perspective)

I would also recommend the following very affordable publications by Sage where you can read more about applied regression methods.

- Menard, S. (2001) *Applied Logistic Regression Analysis*. 2nd Edition.
- Breen, R. (1996) *Regression Models: Censored, Sample Selected, or Truncated Data*.

## Evaluation

Final grades will be calculated based on the following system:

1. Three group projects 25% each. Each group consist of at most three people, and there is a lead person to each project. Each student needs to be the lead person once.
2. One individual term project, 25%, You will need to pick a research question, answer the question by analyzing a real dataset you choose and write a 10-page double spaced report, 20%. each student needs to do a 30 min presentation.

## Auditing Policies

Auditing is in general not encouraged. If you do choose to audit, you will still have to do all the readings and homework (but your homework will not be graded, I will post the solutions as soon as possible.) For students who audit this class in preparation for the comprehensive exam, you should pair up to grade each other's work and tests.

## Computation

In this course we will use statistical software called **STATA**. You should obtain a copy of

- Hamilton, L. C. (2006). *Statistics with STATA 9*. Belmont CA: Duxbury. (\$65.95, an earlier version is acceptable)

To learn more about STATA, there is a complete set of STATA manuals located in 3 Ketchum, but please do not remove them. You can also look for some online help such as by joining the STATA listserv <http://www.stata.com/statalist/>. At the social science data lab website, you will also find some useful information.

day	topics
08.28	review of linear regression Chp 1 of Hoffmann
09.04	theory of GLM Chp 2 Hoffmann
09.11	Logit and Probit models Chp 3 Hoffmann
09.18	Logit and Probit models, Lab session Chp 3 Hoffmann
09.25	Ordered Logit and Probit models Chp4 Hoffmann project 1
10.02	Ordered Logit and Probit models, Lab session Chp4 Hoffmann
10.09	Multinomial model Chp 5 Hoffmann, project 1 due
10.16	Poisson (log linear) model Chp 6 Hoffmann, project 2
10.23	Negative Binomial model, Lab session Chp 6 Hoffmann
10.30	Event history model Chp 7 Hoffmann, project 2 due, project 3
11.06	Event history & survival analysis, Lab session Chp 7 Hoffmann
11.13	Wrap up, discuss your term paper project project 3 due
11.20	Break!
11.27	Presentations
12.04	Presentations
12.11	Presentations

## Course Plan