Sociology 709: Linear and Nonlinear Regression
Models for Social Scientists
Spring 2011
9:00-10:45 am Tuesdays and Thursdays in Hamilton 151

Instructor: Ted Mouw
Hamilton 268 Hamilton Hall
tedmouw@email.unc.edu
Office hours: click on the office hours link on www.tedmouw.org

Course overview: This is the second semester of the two-semester course in statistics for sociology graduate students. Starting with linear regression models and ending with models for categorical variables, we will attempt to cover many of the key statistical models used in modern empirical research. Considerable attention will be paid to the assumptions behind these models, and the consequences of violations of these assumptions. In addition, a major component of the course will be hands-on training in data analysis using Stata.

Requirements:
1) Attendance
Because the topics covered in the class build on each other, class attendance is required. Unexcused absences will result in a grade reduction. More than 3 unexcused absences during the semester will result in automatic failure of the course.
2) Class participation [10% of final grade]
It is expected that students will do the assigned reading before coming to class and be willing to discuss and ask questions on the key issues in the reading. Class participation—i.e., active engagement in class via questions, participation in discussion, and evidence of preparation—will count towards the final grade.
3) Homework assignments [25%]
Homework assignments will primarily involve data analysis questions based on lab work. These (weekly or bi-weekly) assignments will be due 1 week after they are assigned. See the online class schedule for details.
4) Midterm and final exam [20% each]
See the online class schedule page for the dates of the midterm and final exams. Any student receiving below a 70% on either exam will be able to take a make-up exam.
5) Final paper [25%]
A final paper of 15 pages is due on the day after our final day of class. Please email your paper to me. We will work on these papers during the semester.

Required textbooks

[note: to follow these links, go to the online schedule page]

Wooldridge, Jeffrey M. 2009. Introductory Econometrics: A Modern Approach. 4e. Mason, OH: South-Western. [Link to purchase pdf version] [Link for used copies on Amazon]
NOTE: This is the required textbook for Sociology 708. We will continue to use it in this class.

Cameron, A. Colin and Pravin K. Trivedi. ("C&T") 2009. Microeconometrics Using Stata. [Link to Stata bookstore] [Either the original (12/2008) or revised edition (3/2010) is fine] [Link for book on Amazon]

Class Schedule

[Note: The official class schedule is the online version] This version is for course overview only and may not apply for specific class sessions. Use the online version.

Class #	Month	Date	Class link	Topic	Text Reading
1	1	11		Class canceled due to inclement weather	
2	1	13	A	**Part 1: Linear Models for Continuous Variables** [Note: this part continues directly from Sociology 708. Students new to the class should review this material, particularly Wooldridge Chapters 1-2] Review: Multiple Regression Interpreting coefficients, omitted variable bias, sampling variance Matrices and the matrix version of multiple regression	Wooldridge Ch.3 Appendix D
3	1	18	B	Review: Inference from multiple regression Lab : Review of Stata [C&T Ch. 1, 2], making tables from regression output [C&T 3.4.4-3.4.5], also Wooldridge 4.6 Matrix calculation of OLS	Wooldridge Ch.4
4	1	20	C	Making sense of regression Reading: [On reserve] Allison, Multiple Regression: A Primer p. 2-14 ("What is Multiple Regression") and p.97-108 (How does Bivariate Regression Work), p.15-48 ("How do I interpret Multiple Regression Results")	
5	1	25	D	Dummy variables and interaction terms Lab: using and interpreting dummy variables & interaction terms	Wooldridge Ch. 7

6	1	27	E	Statistical theory and asymptotics Lab: simulation (testing the central limit theorem)	Wooldridge Ch. 5, appendix E C&T Ch.4 (for lab)
7	2	1	F	Problems with the error term, Heteroskedasticity & Multicollinearity	Wooldridge Ch. 8 C&T Ch. 3.5, Ch. 5 GLS regression
8	2	3	G	Influential cases and model specification Lab: influential cases & heteroskedasticity	Wooldridge Ch.9 C&T 3.3.3, 3.4
9	2	8	H	Quantile Regression	C&T ch.7
10	2	10	I	Specification error and omitted variable bias, Instrumental variables and causality	Wooldridge Ch. 15
11	2	15	J	Instrumental variables (continued) Lab: IV estimation Causality discussion (papers on low birth weight and infant mortality)	C&T Ch.6
12	2	17	K	Panel data methods Fixed and random effects Lab: Longitudinal data w/ the NLSY	Wooldridge Ch.13-14 C&T Ch.8
13	2	22		Panel data methods, continued	
14	2	24	L	Panel data: Mixed linear models and clustered data	C&T Ch.9
15	3	1		Midterm Exam	
16	3	3	M	**Part 2: Nonlinear Models for Categorical Variables** Nonlinear regression methods Poisson regression	C&T Ch.10
17	3	15	N	Loglikelihood and Nonlinear optimization methods	C&T Ch.11
18	3	17	O	Models for Binary outcomes: Logit & Probit models	Wooldridge Ch. 17 C&T Ch. 14
19	3	22		Binary outcomes (continued)	C&T Ch.14

				Lab 6: Logit and Probit models	
20	3	24	P	Models for ordinal data: ordered probit	C&T 15.9
21	3	29	Q	Multinomial models Multinomial Logit	C&T Ch15.1-15.4
22	3	31	R	Conditional logit Mixed logit	C&T 15.5
23	4	5	S	Tobit and selection models	C&T ch.16
24	4	7		Selection models (continued)	
25	4	12	T	Count data models	C&T ch.17
26	4	14	U	Nonlinear panel models	C&T ch.18
27	4	19	V	Bootstrap methods for standard errors	C&T Ch. 13
28	4	21	W	Review for final exam	