

Sociology 2010
Multivariate Statistical Methods I
Syllabus, Fall 2009: Brown University

Where and When: 108 Maxcy Hall; Thursdays, 2-4:50pm

Instructor: Margot Jackson, 303 Maxcy Hall, 401.863.3192, margotj@brown.edu

Office Hours: Wednesdays, 1:30-2:30pm; Thursdays, 5-6pm; or by appointment

Course website: <http://mycourses.brown.edu> I'll use this site to post lecture slides in advance of class, to post assignments, and to make announcements.

Goals:

This course is designed as an applied introduction to statistical concepts and multiple regression for sociologists. We will discuss the methods used to collect quantitative evidence and begin to apply those methods to analyze quantitative data. I will also provide an introduction to Stata, software for statistical analysis, in the class sessions. Our focus will be applied—I will provide an introduction to underlying theory, but our focus will be on application and interpretation rather than derivation. We will briefly review measures of central tendency and variability, but I will assume a basic knowledge of descriptive statistics. Students who seek a refresher might do a bit of reading from the references listed below before the first class.

A Note on Stata:

Although I will provide some introduction to Stata in class sessions (and will be happy to answer questions to help students complete assignments), this is not a course about Stata, and you will be expected to spend time independently becoming comfortable with programming. The department also sponsors a Stata proctor, Inku Subedi, who will be offering a Stata introduction throughout the semester. You should feel free to solicit Inku's help as a complement to what we cover in class.; where possible, we have tried to coordinate our use of particular data sources. Although Inku is available for assistance with general programming-related questions, you are responsible for completing homework assignments and exam questions on your own. More generally, although you are welcome to work on the homework assignments in small groups, each student should produce his or her own write-up and interpretation.

Grading:

- 1) Five exercises (65% of grade). Each exercise will involve a short report of your analyses, to give you experience in writing about data analyses. The first 4 exercises will each count for 10% of your grade. The last report will be more like a mini-article (10-12 pages) and will be worth 25%.
- 2) Final exam, take-home format (35% of grade). The exam will involve a combination of problems to solve and results to interpret and discuss.

I will not take attendance but I do expect you to attend and to be prepared to contribute.

Books/Materials:

- 1) There are two required texts for the course (and one strongly recommended optional text). These books should be available in the Brown bookstore. You can also find them new or used at many sites online:

David S. Moore. 2010. *The Basic Practice of Statistics* (5th edition). W.H. Freeman and Co.
<http://bcs.whfreeman.com/bps5e/>

Jeffrey M. Wooldridge. 2009. *Introductory Econometrics* (4th edition). South-Western Cengage Learning.

[Optional but recommended: Lawrence C. Hamilton. 2009. *Statistics with Stata: Version 10*. Cengage. <http://www.stata.com/bookstore/sws.html#contents>]

- 2) The syllabus includes some articles and book chapters to download for free via our website. These papers are listed at the bottom of the syllabus, below the course outline. Note: articles require that you be connected to the Brown network for access.
- 3) You will need access to Stata. There are a number of ways to obtain it:
 - a. Purchase it. This is the best long-term option, if you plan to continue in quantitative analysis. You can purchase Stata at a discounted price through their “Grad Plan:” <http://www.stata.com/order/new/edu/gradplan.html> Recommended: Stata/SE 10 or 11
 - b. You can access Stata for free through Brown’s software distribution site: <http://software.brown.edu/dist/> You can download Stata on your computer, and access will be permitted on campus, or off-campus if you’re connected via VPN.
 - c. You can use Stata on any of the machines in the university labs or servers. Stata has extensive online documentation accessible through “help” commands. In addition, Academic Technology Services at UCLA has a very extensive set of online tutorials for Stata: <http://www.ats.ucla.edu/stat/stata/default.htm> These resources, along with my and Inku's help, should be sufficient for this course. If you plan to use Stata a lot in the future, consider purchasing a set of manuals.

Recommended Supplementary Materials:

A number of other books cover the material in this course, and will be useful references for the future: Alan Agresti and Barbara Finlay. 2009. *Statistical Methods for the Social Sciences* (4th edition). Prentice Hall.

John Fox. 1997. *Applied Regression Analysis, Linear Models, and Related Methods*. Sage Publications, Inc.

Peter Kennedy. 2003. *A Guide to Econometrics*. MIT Press.

J. Scott Long. 2009. *The Workflow of Data Analysis Using Stata*. Stata Press.

Jane E. Miller. 2005. *The Chicago Guide to Writing about Multivariate Analysis*. University of Chicago Press.

Stephen Morgan and Christopher Winship. 2007. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Cambridge University Press.

Sophia Rabe-Hesketh and Brian Everitt. 2007. *Handbook of Statistical Analyses Using Stata* (4th edition). Chapman and Hall: <http://www.stata.com/bookstore/statanalyses.html>

Donald Treiman. 2009. *Quantitative Data Analysis: Doing Social Research to Test Ideas*. John Wiley & Sons, Inc.

Tentative Course Outline:

Date	Topic	Reading*	Assignment
Sept. 10	Introductions and Course basics Introducing quantitative analysis/data & sampling Visualizing data with graphs	M: 8, 9, 1, 4 Morgan/Winship (web) Tufte (web)	
Sept. 17	Describing data with numbers and tables	M: 2, 6 Hout	
Sept. 24	Probability and Sampling Distributions	M: 3, 10, 11, 12	Ex. # 1 Due
Oct. 1	Inference I: Estimators and Tests	M: 14-15, 17-21	
Oct. 8	Inference II: Estimators and Tests Linear Model: Bivariate Regression	M: 22, 24 W: 2 Freedman (web)	
Oct. 15	Linear Model: Multivariate Regression	W: 3	Ex. # 2 Due
Oct. 22	Interpretation, Hypothesis Testing, Regression Inference	W: 4, 5	
Oct. 29	Categorical Independent Variables Interactions among Independent Variables	W: 7 Brambor et al.	Ex. # 3 Due
Nov. 5	Interactions, cont. Nonlinearities and Variable Transformations	W: 6	
Nov. 12	Regression Diagnostics; Errors & Residuals; Influential & Missing Data	W: 8, 9 Jasso, Kahn/Udry reply, Jasso reply	Ex. # 4 Due
Nov. 19	Reading, critiquing and producing papers	Miller: 3, 6, 7 (web) Articles TBA	
Nov. 26	No Class: Happy Thanksgiving!		
Dec. 3	Logical extensions of linear regression: pooled cross sections, first differencing	W: 13, 14 England et al.	Ex. # 5 Due (mini-paper)
Dec. 10	Last class: Catch-up, review, discussion	Duncan Fischer et al. (web) Winship/Morgan (SKIM)	
Dec. 14	Final exam due		

*M= Moore; W= Wooldridge. Note about Moore readings: Each chapter contains a large number of examples and problems. Although I'd encourage you to read some of the examples, we'll also cover examples with real data in class. So focus most heavily on the main text of each chapter.

Articles and Book Chapters:

- Thomas Brambor, Williams Roberts Clark and Matt Golder. 2006. "[Understanding Interaction Models: Improving Empirical Analyses.](#)" *Political Analysis* 14: 63-82.
- Greg J. Duncan. 2008. "[When to Promote, and When to Avoid, A Population Perspective.](#)" *Demography* 45(4): 763-784.
- Paula England, George Farkas, Barbara Stanek Kilbourne and Thomas Dou. 1988. "[Explaining Occupational Sex Segregation and Wages: Findings from a Model with Fixed Effects.](#)" *American Sociological Review* 53(4): 544-558.
- Claude S. Fischer, Michael Hout, Martin Sanchez Jankowski et al. 1996. Chapter 4 in *Inequality by Design: Cracking the Bell Curve Myth*. Princeton University Press: Princeton.
- David A. Freedman. 1991. "[Statistical Models and Shoe Leather.](#)" *Sociological Methodology* 21: 291-313.
- Michael Hout. 1999. "[Abortion Politics in the United States, 1972-1994: From Single Issue to Ideology.](#)" *Gender Issues* 17(2): 3-34.
- Guillermina Jasso. 1985. "[Marital Coital Frequency and the Passage of Time: Estimating the Separate Effects of Spouses' Ages and Marital Duration, Birth and Marriage Cohorts, and Period Influences.](#)" *American Sociological Review* 50(2): 224-241.
- Guillermina Jasso. 1986. "[Is It Outlier Deletion or Is It Sample Truncation? Notes on Science and Sexuality.](#)" *American Sociological Review* 51(5): 738-742.
- Joan R. Kahn and J. Richard Udry. 1986. "[Marital Coital Frequency: Unnoticed Outliers and Unspecified Interactions Lead to Erroneous Conclusions.](#)" *American Sociological Review* 51(5): 734-737.
- Stephen L. Morgan and Christopher Winship. Chapter 1 in *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Cambridge University Press: Cambridge.
- Edward Tufte. 1997. Chapter 2 in *Visual Explanations: Images and Quantities, Evidence and Narrative*. Graphics Press: Cheshire, CT.
- Christopher Winship and Stephen L. Morgan. 1999. "[The Estimation of Causal Effects from Observational Data.](#)" *Annual Review of Sociology* 25: 659-706.