

MULTIVARIATE ANALYSIS I
SOCIOLOGY 311
VANDERBILT UNIVERSITY, FALL 2005

Professor Michael E. Ezell

Time: Thursday 1:10 – 3:25
Email: mike.ezell@vanderbilt.edu
Office Hours: Tuesdays 11:30-1:00 and by appointment
Office: 312 Garland Hall

Emily E. Tanner-Smith (Teaching Assistant)

Office Hours: Tuesdays 1-2 and by appointment
Email: emily.e.tanner-smith@vanderbilt.edu

COURSE OVERVIEW

This is a graduate course on the statistical methods commonly employed in the social sciences. This class, coupled with (its second-semester sibling class) Sociology 312, focuses on the application of multivariate quantitative methods, with eventual particular emphasis on alternative specifications of the regression model.

Sociology 311 has two main sections. The first section will cover the basics of statistical analysis, including coverage of methods of graphically describing data, measures of central tendency and variability, probability and probability distributions, sampling distributions, confidence intervals and hypothesis testing, t-tests, and analyzing tabular data.

The second part of Sociology 311 begins a systematic exposition of the linear regression model involving a single, continuous dependent variable (i.e., the ordinary least squares regression model). The method of ordinary least squares (OLS) is derived on the basis of several assumptions, and we will devote attention to understanding these assumptions. Sociology 312 will pick up where this course ends, and it will finish covering the *general* linear model (OLS) and then move on to discussions of the *generalized* linear model (e.g., logistic regression, multinomial logit regression, Poisson regression, negative binomial regression).

Sociology 311 will involve both a practical, applied focus and a theoretical focus, as it is important to know not only how to obtain statistical estimates and their statistical properties, but it is also important (perhaps most important) to know what the estimates substantively communicate to you about the social process under study. Thus, there will be emphasis on both theory and substance.

PREREQUISITES

A previous calculus course is not required for this class but it is assumed that all students in this class have a sound understanding of algebra.

If you need a quick review of basic math rules, please see me for a book or visit this web site:

<http://publish.uwo.ca/~fernando/soc231/skills.htm>

TEXTS & READINGS

Required Textbooks (in bookstore)

David S. Moore and George P. McCabe. (2006). *Introduction to the Practice of Statistics, 5th Edition*. W. H. Freeman and Company: New York. [Denoted as “Moore” on the schedule below]

Michael A. Flinger. (2006) *Study Guide with Selected Solutions*. (in the *Upgrade Study Pack*) W. H. Freeman and Company: New York.

Recommended Textbooks

Kleinbaum *et al.* (1998). *Applied Regression Analysis and Other Multivariable Models, 3rd Edition*. Duxbury: Pacific Grove, CA. [Denoted as “Kleinbaum *et al.*” on the schedule below]

Hamilton (2004). *Statistics with Stata 8*. Duxbury: Pacific Grove, CA.

“Required Readings”

On the class schedule below, “Required Readings” denote the readings that you are expected to have read *prior* to the beginning of class for that day.

“Recommended Readings”

I also highly recommend that you also read the “Recommended Readings” denoted on the syllabus either before or after class to further enhance your understanding.

STATISTICAL SOFTWARE

The statistical software package “Stata” (Version 8) is the official software package of this course. This software package will be used frequently in class throughout this course. There will be two general introductions to this software package during the first two weeks (and you must attend one of them), and there will be several labs throughout the duration of this course.

The professor and the teaching assistant in this course will support the use of this package only. Although we have experience with other packages (SPSS, R, S-PLUS, SAS, LIMDEP), we will not be able to effectively support the use of all other packages. If you have a deep commitment to one of these other programs (and are very knowledgeable regarding the use of another package) and wish to use it in this course, please discuss this issue with the professor immediately.

Unlike previous versions of Stata. Version 8 allows for users to engage in the practice of statistics via a menu-driven command language (i.e., point and click menu system). Previous versions of Stata only allowed the user to run Stata via knowledge of the command syntax language. You will need to submit a Stata “do file” for your assignment sets, which is simply a list of commands used to generate the statistical output.

work or programs. A group solution to the assignment sets is not acceptable and will result in an F grade and referral to the Vanderbilt Honor Council; similarly, submission of identical software code is not acceptable. In other words, you must do your own work, but you may collaborate and help each other (not free load off one person) with the various problem sets in this class. Please let me know if we need to discuss this further.

COURSE WEBSITE

There will be a web site devoted to this class. The web site will contain links to assignments, readings, data sets, and other course materials. The web site location is:

<http://people.vanderbilt.edu/~mike.ezell/courses/soc311/soc311.htm>

Also note that since Stata is “web aware,” you can directly load the data sets into Stata over the web using the following command:

use <http://people.vanderbilt.edu/~mike.ezell/courses/data/insertfilenamehere>

COURSE OUTLINE

Below is a tentative outline of the topics to be discussed in class each day and the due dates for each of the problem sets. This schedule is more or less firm, but if I have inaccurately projected the time required for each topic of discussion, I reserve the right to make modifications to the topics to be discussed and/or the dates of assignments/exam.

Date	Topic
<u>August</u>	
Week 1	
Thursday 25 th	Course Introduction; Types of Data; Summation; Graphically Describing Data Required Reading: Moore Chapter 1 (Pp. 1-40) Recommended Reading: Kleinbaum <i>et al.</i> Chapters 1 & 2 (Pp. 1-13)
<u>September</u>	
Week 2	
Thursday 1 st	Measures of Central Tendency and Dispersion Required Reading: Moore Chapter 1 (Pp. 40-63)
Week 3	
Thursday 8 th	Density Curves and Normal Distributions Required Reading: Moore Chapter 1 (Pp. 64-100) Assignment 1 Due

Week 4

Thursday 15th Scatterplots, Correlation, Regression
Required Reading: Moore Chapter 2
Recommended Reading: Kleinbaum *et al.* Chapters 4 & 6 (Pp. 88-96)
Assignment 2 Due

Week 5

Thursday 22nd Two-Way Tables; Introduction to Probability; Producing Data
Required Reading: Moore Chapter 9 (Pp. 581-590), Moore Chapter 4 (Pp. 251-310);
Moore Chapter 3 (Pp. 191-197; 218-230)

Week 6

Thursday 29th Sampling Distributions; Central Limit Theorem; Estimators; Properties of Estimators
Required Reading: Moore Chapter 5; Moore Chapter 3 (Pp. 231-250)
Assignment 3 Due

October

Week 7

Thursday 6th Confidence Intervals
Required Reading: Moore Chapter 6 (Pp. 381-399); Moore Chapter 7 (Pp. 449-453);
Moore Chapter 8 (Pp. 536-540)
Recommended Reading: Kleinbaum *et al.* Chapter 3 (Pp. 21-24)

Week 8

Thursday 13th Hypothesis Testing
Required Reading: Moore Chapter 6 (Pp. 400-448); Moore Chapter 7 (Pp. 454-459);
Moore Chapter 8 (Pp. 540-548)
Recommended Reading: Kleinbaum *et al.* Chapter 3 (Pp. 24-29);
Assignment 4 Due

Week 9

Thursday 20th Comparing Groups
Required Reading: Moore Chapter 7 (Pp. 459-515)
Assignment 5 Due

Week 10

Thursday 27th Tabular analysis (chi-square)
Required Reading: Moore Chapter 9 (Pp. 591-632)
In-class portion of Mid term at beginning of class

November

Week 11

Thursday 3rd Simple Regression
Required Reading: Moore Chapter 10
Recommended Reading: Kleinbaum *et al.* Chapter 5

Week 12

Thursday 10th Simple Regression
Recommended Reading: Kleinbaum *et al.* Chapter 6 (96-100), Chapter 7
Assignment 6 Due

Week 13

Thursday 17th Multiple Regression I
Required Reading: Moore Chapter 11
Recommended Reading: Kleinbaum *et al.* Chapter 8
Assignment 7 Due

Week 14

Thursday 24th **Thanksgiving Break Week**

December**Week 15**

Thursday 1st Multiple Regression II
Required Reading: Moore Chapter 11
Recommended Reading: Kleinbaum *et al.* Chapter 9
Assignment 8 Due

Week 16

Thursday 8th Multiple Regression III