

# POL 681: Applied Data Analysis II

## Spring 2006

Brad Jones  
Office: SBS 325  
Office Hours: T-TH, 10:00–11:15  
Phone: 621-7605  
E-mail: [bsjones@email.arizona.edu](mailto:bsjones@email.arizona.edu)

## Course Overview

This course is the second in a three-course methodology sequence for the graduate program in Political Science. As such, we will begin where POL 582 ends and end where POL 682 will begin. The principle focus of this class will be on the traditional linear model, and variants of this model as well as alternatives to the classic linear model, for example, models for categorical data. In particular, we will cover in detail issues pertaining to estimation and interpretation of the classic linear model and its alternatives. To that end, we will examine in detail: 1) the least squares principle; 2) its relationship to correlational analysis; 3) the assumptions underlying the principle; 4) alternative functional forms for covariates; 5) the conditions under which violations of classic assumptions lead to the consideration of alternative models; 6) diagnostic methods for linear models; 7) substantive interpretation of results produced by the models presented in this class; 8) maximum likelihood estimation; 9) models based on MLE; and 10) interpretation of results from these kinds of models.

Students will be expected to complete in a timely fashion, a number of problem sets and applied problems using political data, complete two in-class exams, and complete a substantive research paper using some of the statistical methods considered in this class. By way of prerequisites, students are expected to have mastered the material covered in POL 582. This implies that the student understands elementary probability theory and classical hypothesis testing. Moreover, it is assumed that the student has a strong working knowledge of basic inferential statistics.

## Objectives

There are several objectives I hope to accomplish in this class. The primary objective of this course is for the student to exit with a thorough understanding of how and when the traditional linear model is appropriately and inappropriately applied to substantive problems. Also, I expect the student to exit the class knowing how to sensibly apply alternative modeling strategies that are suitable to given kinds of data. Relatedly, students will be expected to exit this course possessing the ability to clearly and precisely interpret *both* the substantive results generated from one's analysis, *as well as* convey to the intended audience, information regarding the uncertainty around one's substantive conclusions. Moreover, students will be expected to have an understanding of the mathematical principles underlying the statistical models discussed in the class, and further, students will be expected to implement these models using statistical software. Students will be expected to take results from statistical software output and present them in a sensible way, for example, through the use of graphical displays, such that the results are understandable to non-technically trained political scientists. Finally, students are expected to be trained sufficiently well such that they can confidently enter more advance methods classes.

## Requirements

In addition to attending class and completing the reading assignments, students will complete 10 problem sets or applied analysis using data I provide or data obtained by the student. In general, assignments will be given out on a weekly basis and will be expected to be completed within 1 week. Of your total grade, problem sets and applied problems will account for **25 percent**. In addition to problem sets, there will also be two exams given during the semester. The exams will constitute **50 percent** of your total grade (that

is, 25 percent each). Finally, each of you will complete a technically and substantively detailed research paper. The paper project will entail the following. By March 23rd, each of you will need to turn in a paper prospectus outlining the research question and theoretical motivation behind the question. Additionally, a description of the data you will be using in this project will need to be included. The final paper will be due on Friday, May 5. The research paper is worth **25 percent** of your total grade. Grades will be determined in the following way: A (90–100 percent); B (80–89 percent); C (70–79 percent); D (60–69 percent); E (59 percent or below).

## Statistical Software

There will be several problem sets assigned requiring the use of statistical software. While I do not care which software you use, I have a strong preference for either R or Stata. I am beginning to make the switch to the R program and will make every effort to provide examples using both R and Stata code. The R package is freely available on the web. Stata costs money.

## Readings

The primary text will be John Fox's *Applied Regression Analysis, Linear Models, and Related Methods*. In addition to the Fox text, we will also use *Statistical Methods for Categorical Data Analysis* written by Daniel Powers and Yu Xie (2000, Academic Press) and Janet M. Box-Steffensmeier and Bradford S. Jones' *Event History Modeling: A Guide for Social Scientists (2004, Cambridge University Press*. We will have a couple of weeks devoted to matrix algebra. I've adopted the Sage Monograph, *Matrix Algebra: An Introduction* written by Krishnan Namboodiri to cover this topic.

## Topics and Assigned Readings

The following gives an outline of the expected path we will take over the course of the semester. Note that Spring Break is March 14-18 and the Midwest Political Science Association Meetings are on April 20. There will be no class on these dates.

1. Regression and Research Design (Jan. 12)

Read: Fox, Chapters 1 and 2 (omit section 2.3).

Problem Set 1: Regression and Correlation Review

2. The Regression Model (Jan. 17)

Read: Fox, Chapter 5.

Problem Set 2: From Output to Interpretation.

3. Inference for the Linear Model (Jan. 19-Jan. 24)

Read: Fox, Chapter 6.

Problem Set 3: Hypothesis Testing and Uncertainty

4. When  $x$  is Categorical (Jan. 26)

Read: Fox, Chapter 7

Problem Set 4: Regression with Categorical Covariates

5. Multiplicative Terms (Jan. 31-Feb. 2)

Read: Fox, Chapter 7, Section 7.3

Read: Robert Friedrich, "In Defense of Multiplicative Terms in Regression Equations." *AJPS*, 26: 797-833. (available on J-Stor).

Problem Set 5: Working with Interaction Terms

6. Transformations of  $x$  (Feb. 7)  
Read: Fox, Chapter 4.

7. Heteroskedastic Error and Multicollinearity (Feb. 9-Feb. 16)  
Read: Gujarati, Chapters 10-11.<sup>1</sup>

**Exam: Feb. 21 Covering Parts 1-7**

8. Matrix Algebra and Regression in Matrix Form (Feb. 23-Mar 2)  
Read: Fox, Chapter 9.  
Read: Namboodiri, Chapters 1-2.  
Problem Set 6: Working with Matrices

9. Regression Diagnostics (Mar. 7)  
Read: Fox, Chapters 11-12.  
Problem Set 7: Diagnostic Methods for Regression

10. Introduction to MLE and Models for Binary Response Variables (Mar. 9-Mar. 28)  
Read: Powers and Xie, Chapters 1-3.  
Read: Fox, Chapter 15.  
Read: Powers and Xie, Appendix B.  
Problem Set 8 Assigned: Logit, Probit and MLE

11. Models for Ordinal Outcomes (Mar. 30-April 4)  
Read: Powers and Xie, Chapter 6.

12. Models for Unordered Outcomes (April 6-April 11)  
Read: Powers and Xie, Chapter 7.  
Problem Set 9 Assigned: Ordered and Unordered Categorical Data Analysis

13. Models for Event Count Data (April 13)  
Read: TBA

14. Introduction to Survival Analysis (April 18)  
Read: Box-Steffensmeier and Jones, Chapters 1-2.  
Problem Set 10 Assigned: Duration Models

15. Duration Modeling Strategies (April 25-May 2)  
Read: Box-Steffensmeier and Jones, Chapters 3-5.  
Problem Set 8 Assigned: Duration Modeling

**Final Exam: May. 9 (Covering 8-15 above).**

## Class Policies

### Policy on Incompletes

Students are expected to complete the course on time. Incompletes will be permitted only in extraordinary circumstances where the student is unable to complete the work for reasons beyond his/her control. If a grade of "I" is assigned, the student must complete all work prior to the completion of the Fall 2006 semester.

---

<sup>1</sup>This is in reference to Gujarati's *Basic Econometrics*, published by McGraw Hill (2003) My copy can be borrowed for photocopying, but this must be done well in advance of Feb. 9.

Any work not completed by then will result in the grade of “E” and be averaged with the student’s other work in computing an overall course grade.

### **Policy on Academic Integrity**

Students are expected to maintain the highest standards of academic integrity. Cheating, plagiarism, and other forms of academic dishonesty will be subject to disciplinary action consistent with University rules and regulations. Students are expected to familiarize themselves with University regulations regarding plagiarism and academic dishonesty. The University of Arizona student code of conduct can be found at the following website: <http://w3.arizona.edu/studpubs/policies/studcofc.htm>. Any questions about these policies should be asked of me within the first two weeks of class. Additionally, students should be aware of the newly instituted policy on “threatening behavior.” Information on this can be found at: <http://info-center.ccit.arizona.edu/policy/threaten.shtml>.

### **Policy on Physical or Mental Disabilities**

If a student has any condition such as a physical or mental disability which will make it difficult for her or him to carry out the work in this course as outlined above, or which will require extra time on examinations, please notify me in the first two weeks of the course so that we may make appropriate arrangements.

### **Policy on Due Dates**

Dates for papers and other assignments have been noted above. Any changes to these dates will be announced in class. Assignments are to be turned in on time. *No late assignments will be accepted and missed assignments/quizzes/exams cannot be made up.*

### **Policy on Classroom Behavior**

Students are expected to behave in a civil and professional manner and in a way consistent with the University of Arizona’s Student Code of Conduct. Additionally, use of cell phones, use of digital cameras, reading outside materials (newspapers, books, etc.), or accessing the internet via wireless connections is strictly forbidden. Failure to comply with these rules may result in your being asked to leave the classroom, or in the extreme case, may result in your expulsion from this class.