

POLSCI-702: Advanced Techniques of Political Science Research **4:00-6:40PM, Thursdays (1/23/11-5/10/11), BOL 293**

Instructor: Dave Armstrong
Office: NWQ Building B #5531
Hours: Tuesday 2-3PM or by appointment
Phone: (414) 229-4239
E-mail: armstrod@uwm.edu
web: <http://www.quantoid.net>

This course builds explicitly on the foundation that was laid in POLSCI 701. In that class, you learned the nuts and bolts behind statistical inference. In this course, we extend those tools to cover the linear model. The linear model is the workhorse of political science research. Nearly all of the techniques you see in published Political Science journals are either direct extensions or close relatives of the linear model. It is a very powerful tool for understanding relationships among a wide variety of types of variables. The course will teach you the ins and outs of estimation, interpretation, diagnostics and presentation of the linear model. If time permits, we will move on to models for binary dependent variables, but I suspect there will be little time for that.

This course will also push you in the direction of becoming producers of research. Your work in this course will be, at least in part, more like “real” data analysis. The idea is that by the end of this course, you will be able to effectively both evaluate quantitative work and present your own analyses.

As you are all graduate students, I expect that you will take this class seriously regardless of your inherent interest in the subject matter. I expect that you will attend class regularly, do the readings and ask questions when something is confusing. Pay particular attention to this last instruction - **ASK QUESTIONS IF YOU ARE CONFUSED**. In a similar fashion to last term, the things we learn in later weeks depend fundamentally on previous material. If you do not understand something it will cause problems later on. Please don't hesitate to ask either in class or outside of class if you are confused. If you miss class, you are responsible for learning the material you missed in a manner that proves least distracting for the other participants in the course. Also, the late work policy is that make-up exams are not given to graduate students and late papers are not accepted (rare exceptions may be allowed on a case-by-case basis).

As with 701, the course will use **R** as its main piece of statistical software. The main immediate benefit to **R** is that it is free. It can be downloaded from <http://www.cran.r-project.org>. This will allow you to do homework and get more familiar with the software anywhere you have a computer (other options we might have chosen cost >\$100). You will also need a calculator for the course. This does not have to be a fancy calculator, it would be helpful if it had the square function (usually something like x^2) and the square-root function (usually either $\sqrt{\quad}$ or \sqrt{x}) is essential.

You final grade in the course will depend on the following:

Quizzes	15
Homework	30%
Short Data Papers	25%
Final Exam	30%

Quizzes

Given their wild popularity last semester, the weekly quizzes are back. We'll work on the same system - you can drop your three lowest quiz grades and there will be one every week covering the readings due that week.

Homework

You will get weekly assignments of various lengths and types some of the questions can/should be answered with pencil and paper while others will utilize the computer. You should consider your colleagues a resource and I encourage you to discuss the problem sets with your them. That said, each person must turn in their own, original answers to the homework problems. Further, you'll notice that only 30% of the grade comes from the homework, thus it makes sense for you to do your own work so you can pass the other 70% of the course.

Short Data Papers

One of the problems I have noticed in past semesters is that people get done with the course and are still lost about how to *do* empirical research on their own. The short data papers are meant to remedy this problem. I want you to come up with a question (presumably at least tangentially related to Political Science) and find some data to answer it. You don't need to do a literature review of any sort, just pose a question and find some data to answer it and answer it. There are five of these due throughout the semester. The first one will be a bit more guided than the others. We will go over an example of one of these in class. The due dates for the five papers are:

Exam

The final exam will test you on material from the entire course. This will be entirely a pencil and paper affair; you will not need to use the computer at all. To put your mind at ease now, you will not need to memorize formulae. For each exam, you will be allowed to bring with you one (1) standard sheet of letter-sized paper (8.5" × 11") with whatever types of information you want written on either side. You will obviously need to know which formulae apply to which problems, but that is not a task that requires memorization so much as it is a task that requires understanding of which procedures are appropriate for which types of problems.

Textbooks

The required text for the course is:

Fox, John. 2008. “Applied Regression Analysis and Generalized Linear Models” Thousand Oaks: Sage Publications, Inc. (ISBN: 9780761930426 New: \$101, Used: \$69.95 [if available])

Fox, John. 2011. “An R Companion to Applied Regression, 2nd edition” Thousand Oaks: Sage Publications, Inc. (ISBN: 9781412975148 New: \$65.95, Used: \$45.50 [unlikely to be available used])

I may provide some additional reading materials that will elucidate points covered in the lecture, but I will make those available electronically as they are needed.

Miscellaneous

You may obtain information on UWM policies concerning academic issues and course conduct here: <http://www.uwm.edu/Dept/SecU/SyllabusLinks.pdf>

Outline

The outline below offers a chronological list of topics covered, but each topic does not necessarily represent a complete lecture’s worth of material. Rather, some topics will require more than one class period to complete, while others can be covered in less than a single class period.

1. Loose Ends/Intro (1/26)
 - (a) Finish up with Correlation
 - (b) Go over 701 Final
 - (c) Talk about short data papers
2. Intro to Regression (2/2)
 - (a) What is Regression?
 - (b) Non-parametric Regression.
 - (c) Why parametric regression is “better”.
 - (d) Examining Data

Read: Fox (2008) Chapters 1-3
3. Essential Mathematics (2/9)
 - (a) Basic math, matrix algebra and scalar calculus, functions

Read: Readings from Jeff Gill’s “Essential Mathematics for Political and Social Research”, available electronically on D2L.

* First Short Data Paper Due

4. Least Squares Regression (2/16)

- (a) Simple Linear Regression in Scalar Form
- (b) Multiple Regression
- (c) Standardized Coefficients **Read:** Fox (2008) Chapter 5, Fox and Weisberg (2011) 149-157

5. Least Squares Regression II (2/23)

- (a) Statistical Properties of OLS Estimators
- (b) Empirical versus Structural Relationships
- (c) Regression in Matrix Form
Read: Fox (2008) Chapter 6, Fox and Weisberg (2011) 190-197

6. Presentation and Categorical Independent Variables (3/1)

- (a) Regression with Dummy Regressors
- (b) Dealing with Categorical Regressors
- (c) General advice on model presentation
Read: Fox (2008) 120-131, Armstrong (2011) on D2L, Fox and Weisberg (2011) 157-164 and 171-175

* Second Short Data Paper Due

7. Testing and Presenting Conditional Hypotheses (3/8)

- (a) Specifying and interpreting conditional hypotheses
Read: Fox (2008) 131-140, Fox and Weisberg (2011) 164-166 and 175-177, Brambor, Clark and Golder (2006) and Berry, Golder and Milton (2012) (available electronically on D2L)

8. Assessing the Linearity Assumption I (3/15)

- (a) Diagnostic plots and methods for Non-linearity including transformations.
Read: Fox (2008) Chapter 4, 277-290 and 291-296, Fox and Weisberg (2011) 125-140, 285-293, 302-314

9. Spring Break on 3/22

10. Assessing the Linearity Assumption II (3/29)

- (a) Non-parametric regression
Read: Fox (2008) 451-462 and 476-508, Fox and Weisberg (2011) 177-181

* Third Short Data Paper Due

11. Assessing and Homoskedasticity and Normality (4/5)

- (a) Diagnostic plots and methods for non-constant error variance and non-normality of the errors

Read: Fox (2008) 267-277, Fox and Weisberg (2011) 184-186, 314-317

12. Outliers and Influential Observations (4/12)

- (a) How to diagnose and treat influential observations.

Read: Fox (2008) Chapter 11, Fox and Weisberg (2011) 294-302

13. Logit and Probit I (4/19)

- (a) Understanding the basics of models for binary dependent variable.

Read: Fox (2008) 335-355, Fox and Weisberg (2011) 229-240

* Fourth Short Data Paper Due

14. Logit and Probit II (4/26)

- (a) Presentation and interpretation of logit and probit models

Read: Fox “Effect Displays in R for Generalized Linear Models” (Available on D2L)

15. Missing Data and Multiple Imputation (5/3)

- (a) Why are missing data bad?

- (b) What are the options for dealing with missing data?

Read: McKnight et al. (2007) 1-64 and 196-212

16. Tying up loose ends (5/10)

* Fifth Short Data Paper Due

Final Exam - Thursday, May 17, 4:00 PM