Quantitative Research in Political Science I
Syllabus : Sept 4, 2006

Prof. Jonathan Nagler — Phone: 992-9676
Office: 726 Broadway, room 752 — jonathan.nagler@nyu.edu
Office hours: 3:00-5:00 Thursdays, or by email anytime

TA: Alejandro Quiroz Flores — aqf200@nyu.edu
Stata Assstance: Melanie Goodrich — melaniegoodrich@nyu.edu

Course Summary: This is two courses. First, we will cover probability and statistics. We will go over basics of probability, and then consider different distributions of discrete and random variables. We will also cover hypothesis testing, everything from comparing means across two samples to draw inferences about the relative population means to considering whether or not two variables are independent of each other. After that, we will move on to regression analysis. We begins with basic hypothesis testing using Ordinary Least Squares regression. In this course we will examine how to build more sophisticated models allowing us to test more complex hypotheses, and also learn more sophisticated statistical tests enabling us to proceed with analysis even when the Gauss-Markov assumptions are violated. The primary emphasis is on identifying statistical techniques appropriate to the question being examined, and correctly applying those techniques. The course will first present most material in scaler form, then move on to the matrix representation of the fundamental estimators we will examine. Students will be expected to be able to prove the properties of the OSL estimator and other more general linear estimators. We will also devote considerable time to computing quantities of interest based on estimates of model parameters, and determining the distribution of those quantities. [**There is no way we will cover the whole syllabus.**]

Everything about this syllabus is subject to change as we proceed thru the course!!!!

Course Requirements: There will be approximately 8 to 10 problem sets. These will be a mix of analytical problems and computer-based problems requiring students to analyze data with the techniques covered in class. You are free to work together on these problem sets, and **are encouraged to do so!** However, you must write up your problem set on your own, and run the code yourself. In addition there will be a midterm, and a final exam. Class attendance and participation are mandatory.

Your grade is determined as follows:

Problem Sets*:	40%
Midterm Exam:	25%
Final Exam:	35%

*Note: Your problem set grade is computed as the percentage of points you received from the total points available on all problem sets. Each problem set is worth anywhere from 30 to 80 points.

Required Books

Suggestions for purchasing books: *www.campusi.com*, *amazon.com*, *barnesandnoble.com*.

- Wackerly, Mendenhall, and Scheaffer, *Mathematical Statistics with Applications*, 6th ed, ISBN 0534377416, May 2001.

    1. Note: I do not know the difference between the 6th and 5th editions of *WMS*.

- Wooldridge, Jeffrey. *Introductory Econometrics*, 3^{rd} edition, 2005.

Recommended Books

- Caffo, Biran and Galin Jones, *Student Solutions Manual for W/M/S's Mathematical Statistics with Applications, 6th edition*, ISBN: 0534463258. [**Strongly recommended. So strongly, that it is almost required.**]

- Hamilton, Lawrence C. *Statistics with Stata Version 9*, 2005, Duxbury. ISBN: 049510972X. [**You might find the version 7 or 8 book very useful as well, and perhaps substantially cheaper.**] This is a very useful book. There should be copies available in the lab when you are working there.

- Greene, Bill. *Econometric Analysis*, 5^{th} edition, 2005. **Ask Neal if you'll use it in Q2, if so - you might want to get hold of it early.**

- Goldberg, Samuel. *Probability: An Introduction*, 1960, Dover. ISBN: 0486652521.

- Klepner, Daniel. *Quick Calculus*, 2^{nd} edition, 1985, John Wiley and Sons. ISBN: 0471827223. ($18 paperback).

- Thompson, Silvanus Phillips. *Calculus Made Easy*, 1998, 3^{rd} edition, St. Martin's Press. ISBN: 0312114109. ($10 paperback).

- Wooldridge, Jeffrey M. *Econometric Analysis of Cross Section and Panel Data*, 2002, MIT Press. ISBN: 0-262-23219-7.

Articles

- Achen, Christopher H., "Measuring Representation: Perils of the Correlation Coefficient", *American Journal of Political Science*, 1977, Vol 21:805-815.

- Alvarez, R. Michael and Jonathan Nagler, "When Politics and Models Collide: Estimating Models of Multi-Party Elections," *American Journal of Political Science*, 1998, 42:55–96.

- Berk, Richard A., "An Introduction to Sample Selection Bias in Sociological Data," *American Sociological Review*, June, 1983, Vol 48:386-398.
- Green, Donald Philip , and Jonathan S. Krasno, "Salvation for the Spendthrift Incumbent: Reestimating the Effects of Campaign Spending in House Elections," *American Journal of Political Science*, Nov, 1988, Vol 32:884-907.
- Heron, Michael. "Post-Estimation Uncertainty in Limited Dependent Variable Models", *Political Analysis'*, 1999, Vol 8(1): 83-98.
- Jacobson, Gary, "The Effects of Campaign Spending in Congressional Elections," *American Political Science Review*, 1978, Vol 72:769-783.
- King, Gary, "How Not To Lie With Statistics: Avoiding Common Mistakes in Quantitative Political Science, " *American Journal of Political Science*, 1986, Vol 30:666-687.
- King, Gary, "Stochastic Variation: A Comment on Lewis-Beck and Skalaban's "The R-Squared"', *Political Analysis*, 1990, Vol 2:185-200.
- King, Gary; Michael Tomz; and Jason Wittenberg. "Making the Most of Statistical Analyses: Improving Interpretation and Presentation," American Journal of Political Science, Vol. 44, No. 2 (April, 2000): 341-355.
- Kramer, Gerald H., "The Ecological Fallacy Revisisted: Aggregate versus Individual-level Findings on Economics and Elections, and Sociotropic Voting", *American Political Science Review*, 1983, Vol 77:92-111.
- Nagler, J., "Coding Style and Good Computing Practices", *PS: Political Science and Politics* September, 1995 - Volume 28, No 3. Pages 488-492.
- Robinson, W.S., "Ecological Correlations and the Behavior of Individuals," *American Sociological Review*, 1950, Vol 15:351-357.
- Wright, Gerald C., "Linear Models for Evaluating Conditional Relationships,", *American Journal of Political Science*, May, 1976, Vol 20:349-373.

Topics Covered

1 Intro : Week 1

- **Read:** Nagler, *PS: Political Science and Politics* 1995.

- Overview of Models, Estimation, Causality
- Probability:

2 Probability : Week 1 - N

- Probability: Definitions
- Probability: Discrete Random Variables
- Probability: Continuous Random Variables

3 Descriptive Statistics

- Mean, Variance, Standard Deviation
- Other things

4 Hypothesis Testing - 1 Sample

- Drawing inferences from 1 sample.
- Hypothesis tests
- confidence intervals
- type 1 and type 2 errors
- 1-tail v 2-tail testss
- t-test
- χ^2 statistic

5 Hypothesis Testing - 2 Samples

- Comparing the means of 2 populations

6 Cross-Classification of Data, Measures of Association

- χ^2 test for independence

- Proportional Reduction in Error (PRE)
- Correlation

7 Matrix Algebra

- What you need to know.

8 The Basic Linear Model : Weeks ? - ?

- Theories, Models, and Hypotheses. {**Read:** Wooldridge: Chapter 1}
- Terminology, Notation, Data.
- **The General Linear Model - 2 Variable OLS Case**
  - **Read:** Wooldridge: Chapter 2.
  - parameters vs. estimates
  - Deriving OLS Estimates in 2-var case.
  - Assumption-free properties of OLS estimates.
  - Brief digression on units of measurement, functional form.
  - Unbiasedness of OLS, variance of OLS estimators, estimating, the variance of the disturbance.
- **The General Linear Model - K Variable OLS Case**
- {**Read:** Wooldridge: Chapter 3.}
  - Deriving OLS Estimates.
  - Interpreting OLS estimates in multivariate regression.
  - Fitted Values and Residuals.
  - Goodness of Fit.
  - Omitted Variable Bias.
  - Variance of the OLS estimates.
  - The Gauss-Markov Theorem → BLUE.
- **More OLS Estimation**.
- {**Read:** Wooldridge: Chapter 4.}
  - Sampling distributions of the OLS estimators.
  - Simple Hypothesis tests: one-sided, two-sided.
  - Confidence Intervals.
  - Testing multiple linear restrictions (F-tests).
  - Reporting regression results.
  - Using Statistical Software.
- R^2 vs. (std. error of $\hat{Y}$)
  - **Read:** Achen (*AJPS*, 1977)

  - **Read:** King (*Political Analysis*, 1990)

- OLS Asymptotics
- {**SKIP:** Wooldridge: Chapter 5.}

9 Multiple Regression and Specification

- **Read:** Wooldridge: Chapter 6.
- Functional Form
  - Quadratics
  - Interaction Terms
- More Goodness of Fit and Specification
- Race of the Variables
- {**Read:** King (*AJPS*, 1986), p 669-678.}
- Confidence Interval for Predictions
- **Read:** Wooldridge: Chapter 7. (Omit 7.5)
- **Read:** Wright, *AJPS*, 1976.
  - Dummy RHS Variables.
  - Interactions Involving Dummy RHS Variables.

10 Heteroscedasticity

- **Read:** Wooldridge: Chapter 8.
  - Consequences of Heteroskedasticity.
  - Testing for Heteroscedasticity.
  - Heteroskedastic-Robust Inference.
  - Weighted Least Squares.

11 More on Specification

- **Read:** Wooldridge: Chapter 9.
  - Brief things on specification.

12 Limited Dependent Variables

- **Read:** Wooldridge: Chapter 17.
- **Read:** Alvarez and Nagler, *AJPS*, 1998.
- **Read:** Herron, *Political Analysis*, 1999.
- Probit and Logit
- LHS variables bounded by 0-100

13 2SLS, IV Estimators, Identification and Simultaneous Equation Models

- **Read:** Wooldridge, Chapter 15.
- **Read:** Jacobson (*APSR*, 1978)
- **Read:** Greene and Krasno (*AJPS*, 1988)

14 Individual Level data vs. Aggregate Data

- **Read:** Robinson (*ASR*, 1950)
- **Read:** Kramer (*APSR*, 1983)
- **Read:** King (*PUP*, 1997), selected pages.

15 Simulation

- Statistical Theory Behind Simulation
- Sampling Error Versus Estimation Uncertainty
- Probability Distributions about quantities of interest.
- Computing a Confidence Interval about quantities of interest.
- Software: Using Clarify **Read:** King (*AJPS*, 2000)

16 What's a Bayesian?