**Prerequisite:** Students should have completed an undergraduate statistics course. Knowledge of rudimentary matrix algebra and calculus is assumed (we will review as necessary).

**Course Summary:** This is a course in regression analysis. After briefly reviewing basic probability and statistics, we begin 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 scalar 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 OLS 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.

**Course Requirements**

There will be approximately 6 to 8 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. There will also be a short paper requiring you to test a hypothesis and write up the results. This can be a replication and improvement of published research. In addition there will be a short midterm, and a final exam. Class attendance and participation are mandatory.

Your grade is determined as follows:

- Problem Sets*: 35%
- Midterm Exam: 20%
- Paper: 20%
- Final Exam: 25%

*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


Recommended Books


Articles


Topics Covered

1 Probability and Statistics : Weeks 1-2

- **Read:** Gujarati, Appendix A.
- Sample Spaces and Probability
- Probability Distributions and Density Functions
- Inference
- Hypothesis Testing
- **Read:** Goldberg, selected chapters.

2 The Basic Linear Model : Weeks 3-6

- Matrix Algebra - review.  
  {**Read:** Namboodiri, Chapter 1 (Skip Markov Chains).}
- Theories, Models, and Hypotheses.  
  {**Read:** Guj: Introduction}
- Terminology, Notation, Data.  
  {**Read:** Guj: Chapter 1}
- Expected Value of Y Given X.  
  {**Read:** Guj: Chapter 2}
- The General Linear Model
  - parameters vs. estimates
  - a disturbance term
  - the interpretation of \( \beta \)
- **OLS Estimation.**  
  {**Read:** Guj: Chap 3.1-3.4}
  {**Read:** Nagler: OLS Notes (online)}
  - Sample Estimates versus Population Parameters
  - a disturbance term
  - the interpretation of \( \beta \)
  - criteria for choosing \( \hat{\beta} \)
  - Properties of OLS (*SANS assumptions!!!*)
  - Gauss-Markov Theorem \( \rightarrow \) BLUE
- **More OLS Estimation.**  
  {**Read:** Guj: Chap 3 (remaining), 4, 5}
  - Normalcy of the disturbances
  - Confidence Intervals and Hypothesis tests
  - Using Statistical Software.  
    {**Read:** Nagler, *PS: Political Science and Politics* 1995.}
- \( R^2 \) vs. (std. error of \( \hat{Y} \))  
  {**Read:** Guj: Chaps 3.5, 7.5, 7.8.}
  - **Read:** Achen (*AJPS*, 1977)
  - **Read:** King (*Political Analysis*, 1990)
• More OLS Estimation - Multiple Regression.  \{Read: Guj: Chap 7\}
  – 3 Variable Models
  – Estimation and Variance of OLS Estimators
  – Interpretation of ‘partial regression coefficients’
  – Multiple RHS Categorical Variables  \{Read: Guj: Chap 15.1-15.8\}

• OLS in Matrix Form  \{Read: Johnston: Chapter 3, omit 3.5-3.7.\}
  \{Read: Johnston: Appendix 3.1 and 3.2\}
• More OLS Estimation - F-Tests of Multiple Coefficients  \{Read: Guj: Chap 8\}
  – F-Tests  \{Read: Johnston: 3.4.5-3.4.7; Appendix 3.3, 3.4\}
  – Increment of an additional variable
• The Race of the Variables  \{Read: King (AJPS, 1986), p. 669-678.\}
• More Matrix Algebra  \{Read: Namboodiri, Chapter 2; Guj: Chap 9.1-9.3\}

• Common Violations of GM and Associated Symptoms
  – \Read: Guj: Chapter 11
  – \Read: Johnston: Chapter 6
  – heteroscedasticity
  – misspecification
• Common Non-Violations of GM and Associated Symptoms  \{Read: Guj: Chapter 10\}
• Common Violations of GM and Associated Symptoms  \{Read: Guj: Chapter 12\}

3 Limited Dependent Variables : Week 7

• \Read: Guj: Chapter 16
• \Read: Johnston: 13.1-13.7
• \Read: Alvarez and Nagler, AJPS, 1998.
• \Read: Herron, Political Analysis, 1999.
• Probit and Logit
• LHS variables bounded by 0-100

4 Model Specification: Week 8

• Some Tests of Specification Errors  \{Read: Guj: Chap 13\}
• Omitted Variables  \{Read: Guj: Chap 7.7\}
• Interactive Effects  \{Read: Wright (AJPS, 1976)\}
• Fixed Effects
• Leamer’s Method of Bounds, ETC
  – \Read: Guj: Chapter 14
5 2SLS, IV Estimators, Identification and Simultaneous Equation Models: Weeks 9-10

- Read: Wooldridge, Chapter 5.
- Read: Jacobson (APSR, 1978)
- Read: Greene and Krasno (AJPS, 1988)
- Read: Handout
- Read: Guj: Chap 18, 19.
- Read: Johnston 5.5, Appendix 5.3.

6 Time Series: Week 11

- Lags {Read: Guj: Chapter 17}

7 Individual Level data vs. Aggregate Data: Week 12

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

8 Simulation: Week 13

- 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)

9 Student Presentations: Week 14

- Students will present research papers in a poster session.