

PADP 8120: Data Analysis and Statistics (Fall 2014)

Hye-Sung Kim

Office: Baldwin Hall 204D

Office Hours: Monday & Wednesday 1:00-2:30 pm

Email: hkim82@uga.edu

Course Website: <http://www.hyesungkim.org/teaching.html>

Course Information

Classroom: Baldwin 202

Class Time: Mon., 6:50 – 9:50 PM

Course Description

This course provides students an opportunity to develop quantitative analysis skills that can be applied to social science research questions. There are two main objectives of this course: the first is to provide students with the mathematical foundation essential for quantitative analyses; the second is to provide students with statistical foundations for data analysis. The course is composed of two parts: in the first part, topics will include matrix algebra, differential and integral calculus, and optimization. The mathematical tools taught in the first part of the course will serve as a foundation for the second part of the course as well as advanced quantitative analysis courses in following semesters, i.e., PADP 8130 and PADP 8140. The second part of the course will include sampling, descriptive statistics, probability distributions, confidence intervals, hypothesis testing, comparison of means, ordinary least squares, and factor analysis. In addition, I will provide three STATA tutorial sessions for statistical analysis, which you will need to use for data analysis assignments.

Required Materials

- *Fundamental Methods of Mathematical Economics* by Alpha C. Chiang and Kevin Wainwright, 4th ed.
- *Statistical Methods for the Social Sciences* by Alan Agresti and Barbara Finlay, 4th ed.

You will need access to Stata. There are 2 ways to obtain it:

1. Purchase it. You can purchase Stata at a discounted price through their "Grad Plan": <http://www.stata.com/order/new/edu/gradplans/>. Recommended: Stata/IC or Stata/SE 13.
2. Use it at the computer lab in Baldwin Hall.

Recommended Texts

For those who want to see more rigorous approaches to the topics, the following texts are recommended although not required for purchase:

- *Mathematics for Economists* by Carl P. Simon and Lawrence E. Blume.
- *Probability and Statistics* by Morris H. DeGroot and Mark J. Schervish

Grading

Class attendance is not required, but is strongly recommended.

Two Exams (60%)

There will be two in-class exams on October 13 and November 17. Each exam counts for 30% of the course grade. Any known scheduling problems should be brought to my attention as soon as possible. If you miss the midterm exam without a university-approved excuse, you will receive a zero. If you have a university-approved absence, a make-up exam will be given. For a second exam, students are allowed to use a one-page formula sheet.

Weekly Problem Sets (20%)

You will be assigned 12 problem sets, which will count for 20% of the course grade. You are encouraged to work in groups (composed of no more than 3 students). The problem sets handed in on time will be graded on a 3 point scale (check plus, check and check minus). Since a solution set will be available on the course website as soon as a problem set is handed in, late problem sets will not be accepted (which means you will receive a zero).

Final Data Analysis (20%)

This data analysis will count for 20% of the course grade. Conduct empirical research on the question of your choice using a dataset of your choice. A research proposal that includes the dataset, the research question and details about the key variables of interest is due on September 29. A data description is due November 17. The final data analysis is due on December 15. The final data analysis will include a brief introduction to the research question, a description of the data and measures, including descriptive statistics, an explanation of the empirical methodology, results, and the discussion/interpretation of the results. No literature review is necessary.

Academic Integrity

All students are responsible for maintaining the highest standards of honesty and integrity in every phase of their academic careers. The penalties for academic dishonesty are severe and ignorance is not an acceptable defense.

Course Topics

Note: The course syllabus is a general plan for the course; deviations announced to the class by the instructor may be necessary. We may not cover all of these topics. Conversely, time permitting, other topics might be covered in this course.

- August 18** **Introduction and Mathematical Background**
- Introduction to Functions; Exponential and Logarithmic Functions
 - CW Ch. 1, Ch. 2, Ch. 10.1-10.4.
 - STATA tutorial 1
 - Problem Set 1 handed out
- August 25** **Matrix Algebra 1**
- Vectors and Matrices, Dimensions, Transposes, Matrix Multiplication, System of Linear Equations, Solving System of Linear Equations
 - CW Ch. 4-5; Assigned Readings from Greene
 - Problem Set 2 handed out
- September 1** **Labor Day - NO CLASS**
- September 8** **Matrix Algebra 2**
- Rank, Existence of Solutions, Inverse of A Matrix, Linear Systems and Inverses; Determinants; The Determinant Formula for An Inverse, Cramer's Rule, Eigenvectors and Eigen Values, (Semi) Definiteness
 - CW Ch. 4-5; Assigned Readings from Greene
 - Problem Set 3 handed out
- September 15** **Differential Calculus 1**
- Derivatives and Rules of Differentiation of Function with One Variable:
 - CW Ch.6, Ch.7
 - Problem Set 4 handed out
- September 22** **Differential Calculus 2**
- Differentials, Total Differentials and Total Derivatives
 - CW Ch.8, Ch. 10.5
 - Problem Set 5 handed out
- September 29** **Optimization**
- Intro to optimization, Optimization with Constraints
 - CW Ch.9.1-9.4; 12.1-12.3
 - Problem Set 6 handed out
 - Research proposal (one-page) due
- October 6** **Integral Calculus**
- Indefinite Integrals and Definite Integrals.
 - CW Ch.14
 - Problem Set 7 handed out
 - Midterm 1 review materials handed-out

- October 13** **Midterm 1**
- October 20** Probability; Random Sampling; Describing Data
- Probability, Conditional Probability, Histograms and Scatterplots, The Center of the Data, Measuring Spread
 - AF Ch. 1-3
 - STATA Tutorial 2
 - Problem set 8 handed out
- October 27** **Probability Distribution**
- Distribution, Z-scores, Normal Tables, Samples, Population, Central Limit Theorem, Repeated Samples, Standard Errors
 - AF Ch. 4
 - Problem set 9 handed out
- November 3** **Estimation and Testing Hypotheses**
- Estimation, Hypothesis Testing, Confidence Interval, Significance Tests
 - AF Ch.5-6
 - Problem set 10 handed out
- November 10** **Comparing Groups**
- Population Difference in Means, Binary Variables, Causal Effects
 - AF Ch. 7
 - STATA Tutorial 3
 - Problem set 11 handed out
- November 17** **Midterm 2**
- Data description due
- November 24** **Thanksgiving Break - NO CLASS**
- December 1** **Ordinary Least Squares**
- Univariate and Multivariate Ordinary Least Squares
 - AF Ch. 9-11
 - Problem set 12 handed out
- December 8** **Measurement Models**
- Factor Analysis
 - Reading materials to be assigned on the course website (TBA)
- December 15** **FINAL DATA ANALYSIS DUE VIA EMAIL BY 11:59 PM**