CORE QUANTIATIVE ANALYSIS SAMPLE SYLLABI

SIPA U6500 – SAMPLE 1:

Course Description:

This course introduces students to the fundamentals of statistical analysis. We will examine the principles and basic methods for analyzing quantitative data, with a focus on applications to problems in public policy, management, and the social sciences. We will begin with simple statistical techniques for describing and summarizing data and build toward the use of more sophisticated techniques for drawing inferences from data and making predictions about the social world.

The course will assume that students have little mathematical background beyond high school algebra. Students will be trained on STATA, which is supported in the SIPA Computer lab. This powerful statistical package is frequently used to manage and analyze quantitative data in many organizational/institutional contexts. A practical mastery of a major statistical package will be an important proficiency for many of you down the road. All students are strongly encouraged to enroll in the SIPA Computer Skills Course. You can obtain more information about the course and your lab sticker at the SIPA lab, which is located on the 5th floor of IAB.

The formal mathematical foundation of statistics is downplayed; students who expect to make extensive and customized use of advanced statistical methods may be better served by a different course. This course also offers less practice in writing research papers using quantitative analysis than some courses (e.g., Political Science 4910). Most MPA students, however, should benefit from our emphasis on generating and interpreting statistical results in many different practical contexts.

Requirements:

Class attendance is required. Lectures will sometimes cover matters related directly to the homework assignments that are not covered fully in the assigned readings.

Students are required to review and obtain any relevant material (e.g., weekly handouts) in advance of each class by going to Courseworks at https://courseworks.columbia.edu. This site will include all course materials including: the syllabus, weekly class handouts, homework assignments, solution sets, midterm and final exam review sheets, and information on data as well as downloadable datasets.

Students are required to come to class having already completed the assigned readings for that class. The purpose of this requirement is to ensure that our twelve classes (not counting an in-class midterm) can focus on learning how to bring to bring statistical concepts and methods to life in an applied context. Class will be conducted in a manner that assumes this advance preparation has been done.

Students are required to attend a weekly lab session in addition to the regular lecture. These labs will be important supplements to each lecture, where concepts and methods will be reviewed and students will receive direction and support as they learn STATA.

Problem sets will be assigned at least a week in advance of their due dates. Late problem sets will not be accepted for credit. You are encouraged to be actively engaged in the completion of every problem set since hands-on work (computer-based or otherwise) is essential to fully absorbing the material presented in this course. Problem sets will be done in groups of two.

Grading:

The three components to the final course grade will include weekly problem sets (30%), a midterm exam (35%), and a final exam (35%). In "borderline" cases, the quality of your class attendance and participation will weigh heavily in determination of a final grade.

Academic Integrity Statement:
The School of International & Public Affairs does not tolerate cheating and/or plagiarism in any form. Those students who violate the Code of Academic & Professional Conduct will be subject to the Dean's Disciplinary Procedures. Click here to view the Code of Academic & Professional Conduct online. http://sipa.columbia.edu/resources_services/student_affairs/academic_policies/deans_discipline_policy.html Please familiarize yourself with the proper methods of citation and attribution. The School provides some useful resources online; we strongly encourage you to familiarize yourself with these various styles before conducting your research: http://sipa.columbia.edu/resources_services/student_affairs/academic_policies/code_of_conduct.html Violations of the Code of Academic & Professional Conduct should be reported to the Associate Dean for Student Affairs.
Readings:

The required and recommended textbooks may be purchased at Bookculture (536 West 112th Street). All readings have been put on reserve in Lehman Library (3rd Floor IAB).

Required Texts:
D. Moore and G. McCabe. Introduction to the Practice of Statistics. 5th Ed (2005)
W.H. Freeman and Company
M. Lewis-Beck, Applied Regression. (1980)
SAGE

Recommended:
Lawrence C. Hamilton. Statistics with STATA: Updated for Version 9. (2007)
Thomson-Brooks/Cole

Supplemental:
Wonnacott & Wonnacott, Knoke & Borhnstedt and Kachigan are statistics textbooks that also cover much of the same material as Moore & McCabe. Students may want to supplement the weekly readings for some units. .
T. Wonnacott and R.Wonnacott, Introductory Statistics. 5th Ed (1990)
(Chapters 1-4, 6, 8, 9, 11-14, 17)
D. Knoke and G. Bohrnstedt, Statistics for Social Data Analysis. 3rd Ed (1994)
(Chapters 1-8)
S.K. Kachigan, Statistical Analysis: An Interdisciplinary Introduction to Univariate and Multivariate Methods (1986)
(Chapter 1-13)

Units on research methods/design and sampling will refer to the readings in Hoyle, et al. Multiple copies of this textbook are on reserve in Lehman.
R. Hoyle, M. Harris, C. Judd, Research Methods in Social Relations. 7th Ed (2002)

C. Achen, Interpreting and Using Regression. (1982)
(For Multiple Regression Units)

COURSE OUTLINE

Session 1: Introduction to Statistics & Research Design
September 8 (Sect. 3 & 4)/September 9 (Sect. 5)

• How can we ask and answer meaningful empirical questions? Testing Theories (Concepts and Variables, Unit of Analysis, Independent and Dependent Variables, Path Diagrams)
• Internal/External Validity, Observational Studies, Randomized Experiments
• Surveys, Sampling Design (Population, Sample, SRS, Probability Sampling)

STATA: Introduction to STATA

Required Readings:
M & M: Chapters 3.1-3.2 (Research Design)
Supplemental: Hoyle, et al.: Chapter 8, 9, 11

Courseworks: Session 1 Handout, Homework 1, Introduction to STATA Handout

Session 2: Exploratory Data Analysis—Single Variable
September 15 (Sect. 3 & 4)/September 16 (Sect. 5)

• Types of Data, Displaying Distributions (Bar Charts, Pie Charts, Histograms), Frequency Distributions, Measures of Central Tendency (Mean, Median, Mode), Measures of Dispersion (Quartiles, Boxplots, Variance, Standard Deviation)
• Density curves, Normal Distribution (Standardizing, Z-scores and table, Standard Normal, Normal Quantile plots)

STATA: Frequency tables, charts and graphs, univariate statistics

Required Readings:
M & M: 1.1-1.3

Courseworks: Session 2 Handout, Homework 2, General Social Survey Online Codebook Handout

Session 3: Exploratory Data Analysis—Relationships between variables
September 22 (Sect. 3 & 4)/September 23 (Sect. 5)

• Association between categorical explanatory variable and quantitative response variable
• Association between 2 categorical variables
• Association between 2 quantitative variables: Scatterplots, Correlation

STATA: Means tables, Crosstabs, scatterplots

Required Readings:
M & M: 2.1-2.2

Courseworks: Session 3 Handout, Homework 3

Session 4: Exploratory Data Analysis—Relationships between variables (cont.)
September 29 (Sect. 3 & 4)/September 39 (Sect. 5)
• Simple Linear Regression (Least Squares Criterion, Slope, Intercept, Predicted Value, Calculating and Interpreting regression coefficients), Residuals, Outliers, Lurking Variables, Causation

STATA: Simple Regression, residual analysis

Required Readings:
M & M: Chapter 2.3-2.4

Courseworks: Session 4 Handout, Homework 4

Session 5: Concepts and Applications in Probability
October 6 (Sect. 3 & 4)/October 7 (Sect. 5)

• Basic Definitions (Sample Space, Event, Properties of Probabilities), Compound Events (Addition rule), Joint Probabilities (Multiplication Rule), Conditional Probabilities, Statistical Independence, Contingency Probability Tables
• Random Variables (Discrete versus Continuous, Probability Distributions, Mean and Variance of Random Variables), Law of Large Numbers

STATA: Contingency probability tables (crosstabs)

Required Readings:
M & M: Chapters 4.1-4.5 & 2.5 (Contingency Probability Tables)

Courseworks: Session 5 Handout, Homework 5

Session 6: Sampling Distributions
October 13 (Sect. 3 & 4)/October 14 (Sect. 5)

• Sampling Distribution for counts (Binomial), Sampling Distribution for proportions, Normal approximation
• Sampling Distribution of the sample mean, Central Limit Theorem.

Required Readings:
M & M: Chapters 3.3 & 5.1-5.2

Courseworks: Session 6 Handout, Pseudo Homework 6, Midterm Review Practice Problems, Midterm Review Key Topics Handout

Session 7: MIDTERM EXAM
October 20 (Sect. 3 & 4)/October 21 (Sect. 5)

Session 8: Basics of Statistical Inference

October 27 (Section 3 & 4)/October 28 (Sect. 5)

• Application to a single mean, Student's t-distribution, Confidence Intervals around a Mean (Variance is known and unknown), Hypothesis testing (Null and Alternative Hypotheses, One-side versus two-sided tests, Alpha level, p-values, t-values)

STATA: One-Sample t-test

Required Readings:
M & M: Chapters 6.1-6.3 & 7.1

Courseworks: Session 8 Handout, Homework 7

Session 9: Inference for two sample means, a single proportion, and two proportions
November 4 (Sect. 5)/November 10 (Sect. 3 & 4)
Note: No class on Tuesday November 3 for Election Day Holiday

• Comparing two means, Difference in Means (Independent samples, Pooled Variance, Paired/Marched Samples), Confidence Interval around a Difference in Means, Hypothesis testing
• Application to a single proportion (Confidence Interval around a proportion, "Margin of error", Hypothesis testing)
• Comparing two proportions, Difference between two proportions, Robustness, Small sample issues

STATA: Paired Samples t-test, Independent Samples t-test, Two-sample test of proportion

Required Readings:
M & M: Chapters 7.2 (Difference in Means) & 8.1-8.2 (Inference for proportions)

Courseworks: Session 9 Handout, Homework 8

Session 10: Contingency Tables & Inference for Simple Linear Regression
November 11 (Sect. 5)/November 17 (Sect. 3 & 4)

• Contingency Tables Reviewed, One-sample Chi-Square test, Chi-Square test of statistical independence
• Simple Linear Regression (OLS, residuals, degrees of freedom), Goodness of fit, Gauss-Markov Assumptions, Hypothesis testing of b-coefficient (Standard error of b-coefficient, Confidence Intervals, t- and p-values), Standardized beta

STATA: Crosstabs, Chi-Square, Bivariate regression estimates, assess statistical significance, calculate predicted values and residuals, correlation matrix, goodness of fit and standardized beta

Required Readings:
M & M: Chapters 2.5 (review) 9.1-9.2 (Chi-Square) & 10.1-10.2 (Inference for Simple Regression)
Lewis-Beck: pp. 9-47

Courseworks: Session 10 Handout, Homework 9

Session 11: Multiple Linear Regression I

November 18 (Sect. 5)/November 24 (Sect. 3 & 4)
Note: No Class on Wednesday November 25 for Thanksgiving Holiday

• Conceptual Issues (Review Randomized Experiments versus Observational Studies), Confounding Factors, Fitting the model, Checks of regression assumptions, Model fit

STATA: Multiple regression estimates, use residuals to check regression assumptions

Required Readings:
M & M: Chapters 11.1-11.2
Lewis-Beck: pp. 47-54

Courseworks: Session 11 Handout, Homework 10

Session 12: Multiple Linear Regression II
December 1 (Sect. 3 & 4)/December 2 (Sect. 5)

• Dummy Independent Variables (Same slopes, different intercepts), Interaction effects (First-order interactions, Different slopes/different intercepts), Dummying out categorical variables

STATA: Creating and using dummy independent variables, creating and using interaction terms

Required Readings:
Lewis Beck: pp. 54-74

Courseworks: Session 12 Handout, Pseudo Homework 11, Final Review Key Topics Handout, Final Exam Practice Problems

Session 13: FINAL EXAM
December 8 (Sect. 3 & 4)/December 9 (Sect. 5)

SIPA U6500 – SAMPLE 2

Course Description

This course introduces methods of quantitative data analysis to students with little or no statistical background. The course has following objectives: to become familiar with fundamental concepts and principles of quantitative data analysis; to understand how to calculate basic descriptive and inferential statistics and interpret them; to learn a statistical software package to perform analysis of quantitative data; and to prepare students for more advanced statistical method courses. Emphasis will be placed on understanding underlying logics of statistical analysis and appreciating when, why, and how various statistical methods are used in di_erent research situations. The ultimate goal is to acquire statistical literacy and to become a critical and informed consumer of quantitative knowledge.

Texts
The required textbook for this course is:

Thurman, Paul W. 2008. MBA Fundamentals of Statistics. Kaplan.

The most important advantage of this book is in its accessibility: it explains and demonstrates the basic theory and practice of statistical data analysis in plain and easy-to-understand languages with minimum levels of mathematics and technicality. Although this book will serve most of our objectives in this course, for those students who want more in-depth and rigorous treatment of the matter, a supplementary textbook (which is NOT required for purchase to complete this course successfully) is assigned:

Agresti, Alan, and Barbara Finlay. 2009. Statistical Methods for the Social Sciences (4th ed.). Pearson Prentice Hall.

Textbooks are available for purchase at the Columbia Bookstore. I also have requested textbooks to be placed on reserve at Lehman library.

Statistical Software

Stata is the statistical software package we will be using throughout the course. Stata is available in all CUIT computer labs; if you want, you can also buy your own copy at a significant discount from their normal prices (see www.stata.com for the GradPlan package).

Stata is largely command-driven (as opposed to menu-driven): though this makes its learning curve a bit steeper, once you learn Stata, you will find it substantially faster, easier, and more flexible in use than its competitors. While Stata has excellent manuals and help-_les, for those who may need additional help, I find the following sources very useful:

Acock, Alan C. 2008. A Gentle Introduction to Stata (2nd ed.). Stata Press.
http://www.ats.ucla.edu/stat/stata/

Requirements and Evaluation

Students are expected to have competed the assigned reading before each class and actively participate in the discussion. In addition to weekly lectures, students are required to attend lab/recitation session every week. In these weekly sessions, TAs will answer general questions, help with homework assignments, and teach students how to manage and analyze data in Stata. Although you don't have to register for any one particular session, it is recommended to attend the same one every week. All labs meet in the 5th Floor SIPA computer lab (510A).

- Wed. 4:10{6:00
- Wed. 6:10{8:00
- Thurs. 4:10{6:00
- Fri. 11:00{12:50

With respect to grading, an assessment of your performance will come from 400 possible points:

- Five homework assignments (30 points each)
- Midterm exam (100 points)
- Final group project (150 points)

Homework assignments are due in class, a week after it is distributed. Late homeworks and electronic versions will NOT be accepted. They may be discussed/solved in teams, but each of you must individually submit your own works for grading. Solutions will be provided in the labs and via CourseWorks. Note that PAs are responsible for grading homeworks and thus any dispute regarding their grading should be resolved with PAs before being brought to the instructor.

Final group project provides students with the opportunity to conduct independent research using statistical tools learned in this course. You should form a team of 3-4 people and write a research paper as a group. If time permits, we will reserve the last class for short presentations of students' work. Students are also required to submit a 1-2 page proposal describing the research question and data sources. These proposals are intended only to make sure that your choice of research question and data sources is reasonable and feasible, and thus will not be graded.

Final letter grades will be assigned based on relative ranking (not necessarily on total points accumulated), with the average grade across each section aimed at B+. Although subject to change based on final total scores, approximately:

- 25% of students will receive A+, A or A-
- 65% will receive B+, B, or B-
- 10% will receive C+ or lower

Statistics Tutoring

A statistics tutor is available to everyone enrolled in this course. The purpose of tutoring is to provide students with more individualized help (in a exible schedule) on preparing homeworks and exams and/or understanding the statistics in general. Please contact Karin Bennett to set up appointments.

Academic Integrity Statement

The School of International and Public A_airs does not tolerate cheating and/or plagiarism in any form. Those students who violate the Code of Academic & Professional Conduct will be subject to the Dean's Disciplinary Procedures. The Code of Academic & Professional Conduct can be viewed online at:

http://sipa.columbia.edu/resources_services/student_affairs/academic_policies/deans_discipline_policy.html

Please familiarize yourself with the proper methods of citation and attribution. The School provides some useful resources online; we strongly encourage you to familiarize yourself with these various styles before conducting your research:

http://sipa.columbia.edu/resources_services/student_affairs/academic_policies/code_of_conduct.html

Violations of the Code of Academic & Professional Conduct should be reported to the Associate Dean for Student Affairs

Session 1:

Topic:

Introduction and Descriptive Statistics

Required Readings:

Thurman - Ch.1

Agresti - 3.1-3.3

Session 2:

Topic:

Probability Distributions

Assignments:

Homework #1 Distributed

Required Readings:

Thurman - Ch.3
Agresti - 4.1-4.3

Session 3:

Topic:
Sampling Distribution of the Mean
Assignments:
Homework #2 Distributed
Required Readings:
Thurman - Ch.4
Agresti - 4.4-4.6

Session 4:

Topic:
Confidence Interval for Population Means
Assignments:
Homework #3 Distributed
Required Readings:
Thurman - Ch.5 & 6
Agresti - 5.1-5.4

Session 5:

Topic:
Single-Sample Hypothesis Tests
Assignments:
Homework #4 Distributed
Required Readings:
Thurman - Ch.7
Agresti - 6.1-6.5

Session 6:

Topic:
Two-Sample Hypothesis Tests
Required Readings:
Thurman - Ch.8
Agresti - 7.1-7.4

Session 7:

Topic: Midterm Exam

Session 8:

Topic:
Association between Two Variables
Assignments:
Group Project Proposal Due
Required Readings:
Agresti - 3.5 & 8.1-8.2

Session 9:

Topic:
Simple Linear Regression
Required Readings:
Thurman - Ch.9
Agresti - 9.1-9.6

Session 10:

Topic:
Multiple Regression
Assignments Due:
Homework #5 Distributed
Required Readings:
Thurman - Ch.10
Agresti -

10.1-10.3 & 11.1-11.4

Session 11:

Topic:
Categorical Variables and Interaction Effects
Required Readings:
Thurman - Ch.10
Agresti - 11.1-11.5

Session 12:

Topic:

Model Selection and Diagnostics

Required Readings:

Agresti - 14.1-14.3

Session 13:

Topic:

Group Project Presentations

Assignments:

Final Group Paper Due