Advanced Quantitative Methods I (API-209)
Harvard Kennedy School
Course Syllabus
Fall 2009
Final Version

Faculty:	Dan Levy	Faculty Assistant: Alex Kent
Office:	Littauer 115	Office: Rubenstein 310-C
Phone:	617-496-4281	Phone: 617-496-6435
E-mail:	dan_levy@harvard.edu	E-mail: alex_kent@hks.harvard.edu

---

Key people:

Teaching Fellows**:** **Matt Blackwell;** blackwel@fas.harvard.edu
Laura Faden; lfaden@fas.harvard.edu

Course Assistants**:** **Kit Rodolfa;** kit_rodolfa@hks11.harvard.edu
Olga Romero; olga_romero@hks10.harvard.edu
Onur Yildirim; Onur_yildirim@hks10.harvard.edu

Course Description:

The goal of this course is to prepare students to analyze public policy issues using statistics. It covers topics in the areas of probability theory, sampling, estimation, hypothesis testing, and regression analysis. While many students taking this class will have already taken courses in statistics and regression analysis, this course will probably place a much stronger emphasis than typical courses on conceptually understanding the statistical methods. Since the course is targeted to first-year students in the MPA-ID program, we will not shy away from using the mathematical tools needed to develop the conceptual understanding. But the emphasis of the course will be on the conceptual understanding and application of the tools rather than on the math or the mechanics behind the tools. So for example, when studying hypothesis testing, we will place a heavier emphasis on what the test is doing, when to use it and how to interpret its results, than on mechanical repetitions of the calculations involved in conducting the test.

The ultimate objective is that by the end of this course you will be able to:

(i) Conceptually understand the statistical methods covered in the course and how they can be applied to analyze a variety of public policy issues.
(ii) Interpret the results of statistical analyses and think critically about the potential issues that arise when trying to draw conclusions from such results.
(iii) Conduct statistical analyses using Excel and a statistical package called Stata.

Class Meetings:

Classes: Tuesdays and Thursdays, 1:10-2:30, L140.
Review sessions: Fridays, 8:40-10:00, Land
Fridays, 10:10-11:30, Land
[You only need to attend one session]

Office Hours:

Fridays, 2:00-4:30, L115.
Office hours for TFs and CAs will be posted on the course website by next week.

Prerequisites:

The main prerequisite for this course is an understanding of calculus at the level of a typical undergraduate course on the subject. However, as far as mathematics goes, the most important prerequisite is a certain level of "mathematical sophistication," i.e. comfort in dealing with mathematical constructs and arguments. Experience with computer programming is desirable.

Audience:

This course is intended for first-year students in the MPA/ID program. Students not in the MPA/ID program will be admitted only with permission of the instructor. Familiarity with the concepts indicated in the background section of the schedule (last column of table below) will be assumed. Students looking for a more mathematical course should consider EC2110 and EC2140 (offered by the Economics Department). Students looking for a less mathematical course should consider API-201 (offered by the Kennedy School).

Texts and materials:

Required Course Packets:

Two course packets contain readings required for the course. They can be acquired at CMO.

Key Textbooks:

- *Statistics and Econometrics: Methods and Applications*, by Ashenfelter, Levine, and Zimmerman, Wiley, 2003. [ALZ]. This book provides a good and concise coverage of the key concepts used in this course.
- *Introductory Econometrics: A Modern Approach*, by Jeffrey Wooldridge, South-Western, Third Edition, 2005. [W]. This textbook will be used mainly for regression analysis, and may be a useful book to have for API-210.

Background Textbooks (Optional):

- *Mathematical Statistics with Applications*, by Wackerly, Mendenhall and Scheaffer, Duxbury, Sixth Edition, 2002. [WMS]. We will use this textbook mainly for the first half of the course. It is recommended if you do not already have a good mathematical statistics textbook with you.
- *Probability and Statistics*, by DeGroot and Schervish, Addison Wesley, Third Edition, 2002: This textbook is similar to WMS both in style and mathematical sophistication, but it's organized slightly differently.
- *Statistics: A Tool for Social Research, by Joseph Healey, Wadsworth, Sixth Edition, 2002:* This textbook is at a lower mathematical level than WMS, but contains many more applications related to social science. It does not cover probability theory in any great detail.
- *Mind on Statistics, by Utts and Heckard, Thomson, Second Edition, 2004.* This book is at a much lower mathematical level than WMS and ALZ but provides more examples of the use of statistics in everyday life.

Other Material:

- *Stata software*. Completing some homework assignments will require the use of a statistical package called Stata. If you don't have easy access to a computer with Stata, you may want to consider purchasing this software package from CMO. For details about using Stata, please visit http://www.ksg.harvard.edu/mpaid/secure/stata.htm

Handouts:

Handouts will be distributed throughout the course. The main objective of the handouts is to facilitate the process of taking notes so that students can fully engage in class. They are not meant to substitute for class attendance or for studying the assigned reading material. Handouts will contain blank spaces for you to fill in during class, usually in response to questions.

Grading: The class grade will be based on the following criteria:

10% - Problem sets
15% - Class participation and engagement
15% - Final Exercise
25% - Midterm exam
35% - Final exam

Problem Sets (10%)

Problem sets will be assigned almost every week. They will give you hands-on experience with the analytic techniques introduced in class. You should plan to spend approximately 8-10 hours on each problem set. Problem sets will be posted on the course website, as will suggested answers. They will be graded on check-plus/check/check-minus basis.

Problem sets not received before the deadline will be considered late. There will be no credit for late assignments. The lowest problem set grade will be dropped when calculating the average grade for the problem sets.

Under the Kennedy School Academic Code, the problem sets for this course are "Type II" assignments. You are encouraged to work in a study group, but **must submit their own hand- or type-written solutions**. Examples of assignments that are not in accordance with the KSG academic code include photocopies or reprints of substantially identical assignments, printouts of substantially identical Excel tables or graphs, and copies of solutions from previous years. The Kennedy School Academic Code is available at: http://www.hks.harvard.edu/var/ezp_site/storage/fckeditor/file/pdfs/degree-programs/registrar/academic_code.pdf

Instructions for submitting problem sets:

- Put them on specially designated piles located near one of the entrances (near one of the CAs)
- Submit them before class begins. Assignments handed in 20 or more minutes after class begins will be considered late.
- Indicate on the top right corner of the first page the names of the students you worked with.

Class participation and engagement (15%)

I strongly believe that student participation can substantially enrich the learning experience for both the students and the instructor. In this spirit, class participation is encouraged. Effective class participation requires that you read the assigned readings *before* coming to class. You are encouraged to ask questions and to share with the class any relevant insights you may have from your work experience or from previous exposure to these topics. I only ask that the questions and comments be brief and related to the topic at hand. Given that this is a large class, I will sometimes need to defer questions for a future class or office hours.

For a select number of classes, I will ask that you post on the class website a very brief comment (no longer than 150 words). There are two kinds of postings: pre-class and post-class. The pre-class postings are about the assigned reading(s) and are due by 4:00 AM the morning of class. The post-class postings are about the class and will be due 24-48 hours after the class. Comments should be posted on the class website and I encourage you to build on one another's comments.

The class participation and engagement grade will depend on three things: (1) your participation in class, (2) your engagement with the course outside class, (3) your postings on the course website (quantity and quality of postings will both count).

Final Exercise (15%)
The final exercise will require applying some of the statistical tools learned in class using a real data set. It will be due on December 3rd. More details will be provided later in the course.

Regrade Policy
Requests for reconsideration of grades on exams are not encouraged, and will be accepted only in writing, with a clear statement of what has been mis-graded, and within one week of receiving your graded exam. Please submit your full exam so grading on all questions can be reconsidered.

All course activities, including class meetings, homework assignments, and exams are subject to the HKS Academic Code and Code of Conduct.

Letter Grades
Letter grades will be determined according to the Dean's Recommended Grade Distribution (available at http://www.hks.harvard.edu/degrees/registrar/faculty/grades#heading_04)

Tentative Schedule:

Tentative schedule is attached. Adjustments may need to be made and will be distributed as soon as they are available. ALZ refers to readings from Ahenfelter, Levine and Zimmerman. WMS refers to readings from Wackerly, Mendenhall and Schaeffer. W refers to readings from Wooldridge.

READINGS

The New York Times. "Mammogram talks prove indefinite." January 24, 1997, pp. A1, A15.

The Economist, "More Haste. Rapid diagnostic tests" June 16, 2007.

* Mathematical Statistics with Applications, by Wackerly, Mendenhall and Scheaffer, Duxbury, Sixth Edition, 2002. Chapter 1 (pages 1-16), Sections 10.1-10.2 (pages 460-467), 10.6 (pages 482-485)

* Stokey and Zeckhouser, A Primer for Policy Analysis, *Chapter 12*

The Economist, "Signifying Nothing?" January 31st 2004

* Wooldridge, *Introductory Econometrics: A Modern Approach*, South-Western, Third Edition, 2005, Sections 3.3-3.4 (p. 89-111), Sections 4.1-4.5 (p. 123-165)

* Shadish, William R., Thomas D. Cook, and Donald T. Campbell. 2002." Statistical Conclusion Validity and Internal Validity," and "Construct Validity and External Validity." Chapters 1, 2, and 3 in *Experimental and Quasi-Experimental Designs for Generalized Causal Inference* (Boston: Houghton Mifflin). pp. 33-47, 53-55, 83-96.

Dugger, Celia. World Bank Challenged: Are Poor Really Helped? The New York Times; July 28, 2004.

The Economist, "Try it and See it" (2002)

* Orr, Larry L. 1999. Social Experiments: Evaluating Public Programs with Experimental Methods (Thousand Oaks, CA: Sage Publications), pp. 103-115.

Les Roberts, Riyadh Garfield, Jamal Khudhairi, Gilbert Burnham; "Mortality before and after the 2003 invasion of Iraq: cluster sample survey" Lancet (November 20, 2004)

*: Reading is in course packet. Rest of the readings can be accessed electronically through links on the course website.

Date	Class	General Topic	Specific Topics (Tentative List)	Problem Set Due	Readings	Background Material
3-Sep	1	COURSE OVERVIEW	• Review of syllabus and logistics •Key concepts/topics: Population, sample, estimator, sampling distribution, hypothesis testing, causality, regression, program evaluation		ALZ [Ch. 1], WMS [Ch. 1]	Basic concepts in descriptive statistics: Mean, Variance, Std Deviation, Mode, Covariance, Correlation, Conditional Mean [ALZ: Ch. 4, 5.7-5.8; WMS: 1.3, 1.4]
8-Sep	2	PROBABILITY: Bayes Rule	• Introduction (motivation; relation to statistical inference) •Conditional Probability •Laws of Probability •Bayes Rule•Mammogram application		NYT Mammogram (1997), The Economist (2007), ALZ (2.5-2.7],	Key concepts in probability: Experiment, Event, Set, Axioms of Probability [ALZ: 2.1-2.4 or WMS: 2.3, 2.4]
10-Sep	3	PROBABILITY: Decision Analysis	• Decision Analysis and Decision Trees •Choosing the Preferred Course of Action •The Value of Information •Drawing Inference from Imperfect Tests•Allowing for Risk Aversion•General Framework	PS 1	Stokey and Zeckhouser (Ch 12)	•Random Variables and Probability Distributions (pdf, cdf) • Expected Value and Variance of Random Variables [ALZ: 3.1-3.5; WMS: 3.1-3.3, 4.1-4.3; W: B1, B3], Conditional Expectation
15-Sep	4	ESTIMATION: Sampling Distribution	•Sampling •Sampling Distribution •The Central Limit Theorem • Illustrations of the Central Limit Theorem•Normal Approximation to the Binomial Distribution	PS2	ALZ [Ch. 6]	Normal Distribution [ALZ: 3.7.1]; Binomial Distribution [ALZ: 3.6.2]
17-Sep	5	ESTIMATION: Estimators and Key Properties	•Estimators •Confidence Intervals • Point Estimators •Properties (Unbiasedness, Consistency, and Relative Efficiency)		ALZ [7.1-7.4; 8.1-8.2]	
22-Sep	6	HYPOTHESIS TESTING: Introduction	• Conceptual Framework •Intuition Behind Hypothesis Testing •Steps involved in Hypothesis Testing • p-value	PS3	WMS (10.1-10.2, 10.6)	
24-Sep	7	HYPOTHESIS TESTING: t test	•Hypothesis tests involving the normal distribution •Hypothesis tests involving the t distribution•Hypothesis tests in the context of regression analysis		ALZ [8.3]	Main distributions used for hypothesis testing: Normal, t, Chi-Square, and F [W: C.6]
29-Sep	8	HYPOTHESIS TESTING: Chi-Square	•Introduction •Goodness of Fit application •Testing Independence •Advantages and Limitations •Practical Significance vs. Statistical Significance	PS4	The Economist (2004)	
1-Oct	9	HYPOTHESIS TESTING: ANOVA	•Introduction •Conceptual Framework •Example •Relationship between t test and F test•Multiple tests•Two other uses of the F test			
6-Oct	10	HYPOTHESIS TESTING: Final Overview	• Type I and Type II Errors•Specifiying Hypotheses•Examples: which hypotheses to test? Which test to use?	PS5		
8-Oct		**MIDTERM EXAM**				
13-Oct	11	REGRESSION: Introduction	• Introduction •Causality and the Counterfactual • Validity		Shadish et al. (2004), Dugger (2004), The Economist (2002)	
15-Oct	12	REGRESSION: Bivariate Regression I	• Population Regression Function •Sample Regression Function •Ordinary Least Squares		ALZ [Ch 9]	

20-Oct	13	REGRESSION: Bivariate Regression II	• Randomized Experiments •Bivariate Regression with a Dummy Variable			
22-Oct	14	REGRESSION: Multiple Regression - Estimation I	• Introduction•Multiple Regression Analysis - Inference •Sampling distribution of OLS Estimators •The t test•The F test		ALZ [11.1-11.3], W [3.1-3.2]	
27-Oct	15	REGRESSION: Multiple Regression - Estimation II	•Omitted Variable Bias - Theory •Omitted Variable Bias - Examples	PS 6	ALZ [13.3], W [3.3-3.5]	
29-Oct	16	REGRESSION: Multiple Regression - Estimation III	• Introduction•OLS in Matrix Algebra Notation•Goodness of fit•Variance of OLS Estimators•Efficiency of OLS: The Gauss-Markov Theorem•Heteroskedasticity		ALZ [13.3], W [3.3-3.5]	
3-Nov	17	REGRESSION: Hypothesis Testing	•Introduction • Sampling Distribution of OLS Estimators•The t test • The F test	PS 7	ALZ [11.6, 12.3], W [4.1-4.6]	
5-Nov	18	REGRESSION: Functional Form - Dummy Variables	• A Closer Look at Dummy Variables•Using Dummy Variables for Multiple Categories • Interactive Dummy Variables•Dummy Variables as Dependent Variables•Appendix - Probit and Logit		ALZ [12.1-12.2]	
10-Nov	19	REGRESSION: Functional Form - Other Non-Linearities	•Introduction•Using Logs: The Three Cases•Quadratics•Summary of Main Functional Forms Using Logarithms•The Natural Logarithm Function•Quadratics in Regression Analysis•Summary Table of Interpreting Regression Coefficients	PS 8	ALZ [9.5, 13.4]	
12-Nov	20	REGRESSION: Other Topics	TBA			
17-Nov	21	REGRESSION: Designing an Empirical Study	•Assessing Hurricane Mitch - Selecting a design to evaluate the impact of PATH	PS 9		
19-Nov	22	SAMPLING: Statistical Power	• Introduction •Framework •Factors that Affect Statistical Power •Minimum Detectable Differences		Orr (1999)	
24-Nov	23	SAMPLING: Other	TBA	PS10		
26-Nov		**NO CLASS: Thanksgiving**				
1-Dec	24	SAMPLING: Sampling Design	•Introduction•Problems in Political Polling•Probability Sampling •Issues in Sampling•Probability Sampling•Illustration: Eestimating the number of civilian deaths in post-war Iraq•Appendix#1 - Survey Design•Appendix#2 - Nonprobability Sampling		Roberts et al (2004)	
3-Dec	25	REVIEW		Final Exercise		
14-Dec		**FINAL EXAM (2 pm - 5 pm)**				

ALZ refers to readings from Ashenfelter, Levine and Zimmerman. WMS refers to readings from Wackerly, Mendenhall and Schaeffer. W refers to readings from Wooldridge.