

Stat E104: Introduction to Quantitative Methods for Economics and Other Quantitative Fields

Tentative Syllabus for Spring Semester 2016

This syllabus is required reading for the course. You will be expected to understand the policies and assignments discussed in the syllabus.

This course has been modified from previous semesters. Please review this syllabus in detail before you commit to taking the course.

PREFACE

This is the fifth time I have taught an extension school/distance learning course. We are still working on details but this document contains the general schedule of course topics, the dates for exams/project, and optional section information. Please contact me at michaelparzen@gmail.com with any questions or concerns.

Course Lectures: The campus lecture is MWF from 11am-12pm (feel free to attend if you are in town). My lectures will be videotaped and placed online, usually a few hours later the same day.

Office Hours: I will have video office hours each week-details forthcoming.

Weekly Section: There is an optional weekly section where a teaching assistant reviews topics from lecture, goes over additional example problems and answers questions about the homework. This will be held online at a time/date not determined yet. This section will also be videotaped and available online for viewing whenever is convenient.

Teaching Assistant office hours: There will be at least two weeks of video conferencing enabled office hours each week-details forthcoming.

Teaching Assistants: Kaitlin Hagan is the head teaching assistant for the course. Depending on enrollment we might have more teaching assistants. You will be randomly assigned a teaching assistant for grading purposes, but everyone will be handling the Thursday night sections and are available for your questions.

Availability: My teaching staff and I want to be as accessible as possible, given the constraints of a distance class. We welcome you to be in contact with us as much as possible and to let us know of any issues or problems as soon as possible.

Course Mechanics

Instructor

Michael Parzen

Office: Science Center, Room 300b

Campus Office hours: Tues/Thur 9-10:00 am or by appointment

Email: mparzen@fas.harvard.edu or michaelparzen@gmail.com

Head TF

Kaitlin Hagan, stat104kaitlin@gmail.com

The Head TF will be your first contact after your assigned TF for grading issues, section issues and questions about the homework and exams. You are always welcomed and encouraged to contact the instructor directly with any questions or concerns

Course website

<https://canvas.harvard.edu/courses/8322>

Optional texts

There will be several textbooks we might refer to during the course. All books are available electronically from hollis.harvard.edu if you are logged in as a student. DO NOT PURCHASE THESE BOOKS-as a Harvard student you have free access to electronic versions of all of them. The books include (don't be alarmed by this list-this is as much for my reference as for yours).

- [*Statistics for Business and Financial Economics*](#) by Cheng-Few Lee
- [*Modern Mathematical Statistics with Applications*](#) by Jay Devore
- [*The Manga Guide to Statistics*](#) by Shin Takahashi
- [*Head First Statistics*](#) by Dawn Griffiths
- [*Schaum's outline of theory and problems of business statistics*](#) by Leonard J. Kazmier
- [*Schaum's outline of theory and problems of statistics and econometrics*](#) by Dominick Salvatore.
- [*Exploratory Data Analysis in Business and Economics*](#) by Thomas Cleff

Computing

The course will use Stata, a statistics package favored by economists. Stata is available free of charge from Harvard University's Instructional Computing Group and details concerning how to install Stata on your machine (mac or pc) will be given in class. You will also need a hand-held calculator with log, exponential and square-root functions.

Lecture Slides

Eighteen slides per page will be made available on the course website. A packet of slides may be purchased-see website for details.

Grading

Your overall percentage score for the course will be computed using the weights described below. We will also drop some of your lowest scores as indicated in the table.

Grade Component	Weight	Dropped	Date
Homework	15%	1	Weekly
Individual Regression Project	5%	0	May 3
Exam 1	20%	0	Feb 29
Exam 2	20%	0	April 4
Final Exam	40%	0	TBD

Homework (15%). Ten homework assignments (approximately) will be assigned. The assignments will be posted on the course website. You will submit homework in pdf format to the course website. This course largely focuses on teaching you how to interpret and analyze data. We expect you to use correct grammar and spelling while providing clear and concise explanations.

Homework will be due Sunday by 10pm Boston time in pdf format submitted electronically to the course website. Homework assignments may be downloaded from the course website. No late homework will be accepted for any reason.. Homework is not only a fairly substantial portion of your grade, but it is vital to your success in this class. Working with other students on homework is allowed and encouraged, as long as you hand in your own work, and do not simply copy someone else's work. Solutions to all problems from each assignment will be posted on the course website.

Individual Regression Report (5%). A report of a multiple linear regression analysis will be individually prepared. More details of this project will be handed out towards the middle of the course. If you are taking the course for graduate credit, you will have to do a more open ended project. More details about this around the midway part of the lecture, though feel free to contact me if you want to start earlier.

Exam 1 (20%). This is an evening exam on February 29 (see below for exam taking options) lasting 1 hour. Two double sided 8.5x11 inch pages of notes are allowed, as well as a multifunction calculator. This exam will cover material from Lectures 1-14.

Exam 2 (20%). This is an evening exam on April 4 (see below for exam taking options) lasting 1 hour. Two double sided 8.5x11 inch page of notes are allowed, as well as a multifunction calculator. This exam will cover material from Lectures 15-26.

Final Exam (40%). This 3 hour exam will be given at the assigned final exam time (not known yet). Six double sided 8.5x11 inch pages of notes are allowed, as well as a multifunction calculator.

The Exam portion of your final grade will be the maximum of (20%exam1 +20%exam2+30%final), (10%exam1 +20%exam2+40%final) or (20%exam1 +10%exam2+40%final).

How Distance Students Take Exams

Please review the information at <http://www.extension.harvard.edu/exams-grades-policies/exams> and <http://www.extension.harvard.edu/exams-grades-policies/exams#online>

If you live within the six-state New England area (Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont) you must come to campus to take your exam. If your primary address throughout the term is not in New England you may come to campus or you must find a qualified proctor near your home to administer the exam in absentia in a proctored setting in your home area according to the following procedures.

Finding a qualified proctor

You are responsible for finding a qualified proctor to administer each exam. The proctor must be a teacher, professor, or administrator in a nearby secondary school, college, university, or testing center. Proctors cannot be family members, friends, acquaintances, co-workers, or librarians.

For assistance finding a proctor, contact a guidance counselor, an academic advisor, or the dean of students or registrar's office at a nearby college, university, or high school . You may also use a testing service such as Sylvan Learning Centers, Huntington Learning Centers, and the Consortium of College Testing Centers to proctor your exams. Testing services charge fees for their proctoring services. The Consortium of College Testing Centers has 255 testing centers in 42 states and districts that provides proctoring services to distance students near their homes. Any fees incurred are your responsibility.

Submit your proctor information through online services no later than a week before the campus exam. Before doing so, confirm that your proctor has a business e-mail address, reliable access to the Internet, the ability to download mp3 files (some exams have audio components) and to download and print PDFs, and access to a fax machine or scanner (for returning completed exams to the Extension School). Exam materials are not sent to proctors who do not have a business e-mail address.

Submitting the proctored exam form

- Complete and submit a proctored exam form online for each exam (even if you use the same proctor) no later than one week before the on-campus exam date.
- Before submitting the form, find a qualified proctor (see above) with whom you have arranged the date, time, and location for their off-campus exam. If you do not submit a proctored exam form or your form is late, you may forfeit the opportunity to take your exam and will be subject to the instructor's make-up policies for mid-term exams and Extension School make-up policies for final exams.
- E-mail confirmations will be sent to you and your proctor upon receipt of the proctored exam form. If confirmation is not received within 24 hours, send an e-mail to distance_exams@dcemail.harvard.edu or call (617) 495-0977 Monday through Friday, 9 am to 5 pm eastern time.

Proctored exam approval and procedures

When the proctored exam request has been approved, a second e-mail that will be sent to you and your proctor that includes details about exam policies and procedures. Exam materials are available to proctors via a password-protected, secure website for a specific length of time. Proctors must download the exam materials within the specified timeframe and administer the exams to the students.

You must complete exams within the specific period of time and no later than the on-campus exam date. Proctors must return all exam materials by fax or e-mail (scanned materials) immediately following completion of the exam. Detailed instructions—including the secure web address, fax number, and return e-mail address—are included in the e-mail sent to proctors shortly before the date of the exam.

If you or your proctor do not follow the exam procedures, you may forfeit the exam.

Getting help

For more information, e-mail distance_exams@dcemail.harvard.edu anytime, or call (617) 495-0977 Monday through Friday, 9 am to 5 pm, Eastern standard time.

Missed Exams

Harvard's policy on missed exams may be found here :

(http://static.fas.harvard.edu/registrar/ugrad_handbook/current/chapter2/attendance_absences_etc.html)

Students who miss an exam due to a religious observance are entitled to a make-up exam. For any other excused absence (medical, sports or certain club activities) the remaining exam percentages will be reweighted and a make-up will not be offered.

Collaboration

You may discuss homework problems with other students (and with the instructor and TFs, of course), but you must write your final answer yourself, in your own words. Solutions prepared “in committee” or by copying or paraphrasing someone else’s work are not acceptable. All computer output you submit must come from work that you have done yourself. You may not consult with other students on the individual regression report.

Regrading

Clerical errors will be corrected without any hassle. Other regrade requests must be submitted **in writing** within a week of the items return. To discourage “grade grubbing,” the **entire** item will be subject to regrading (even if the regrade request is not honored).

Other learning resources

Office hours will be held by the course instructor, and each teaching fellow. In addition, additional walk-in office hours are provided by study.net for all the introductory statistics courses at Harvard. Details of when study.net meets will be given in class.

Students with disabilities:

Harvard Extension School is committed to creating an accessible academic and campus community.

If you have a disability, we ensure that you have equal opportunity to participate in, contribute to, and benefit from our academic and residential programs. Information regarding services and accessibility may be found at this link: <http://www.extension.harvard.edu/resources/disability-services>

Student Responsibilities and Academic Integrity:

Harvard College is a community dedicated to scholarship, leadership, and service and to the principles of honesty, fairness, respect, and accountability. Citizens of this community commit to reflect upon and uphold these principles in all academic and non-academic endeavors, and to protect and promote a culture of integrity. Cheating on exams and quizzes, plagiarism and copying others' work on homework assignments and projects, lying about an illness or absence and other forms of academic dishonesty are a breach of trust with classmates and faculty, and will not be tolerated. Such incidences will result in a grade of zero for all parties involved as well as being reported to the Extension School Administrative Board. Please review Harvard’s Student Responsibilities and Academic Dishonesty policies at

<http://www.extension.harvard.edu/exams-grades-policies/student-responsibilities>

Course Description, Objectives and Tips for Success

Consider the following recent findings about bread:

- More than 98 percent of convicted felons are bread users.
- Fully HALF of all children who grow up in bread-consuming households score below average on standardized tests.
- Bread is made from a substance called "dough." It has been proven that as little as one pound of dough can be used to suffocate a mouse. The average American eats more bread than that in one month!
- Newborn babies can choke on bread.
- Bread has been proven to be addictive. Subjects deprived of bread and given only water begged for bread after as little as two days.
- Bread is often a "gateway" food item, leading the user to "harder" items such as butter, jelly, peanut butter, and even cold cuts.
- In the 18th century, when virtually all bread was baked in the home, the average life expectancy was less than 50 years; infant mortality rates were unacceptably high; many women died in childbirth; and diseases such as typhoid, yellow fever, and influenza ravaged whole nations.
- More than 90 percent of violent crimes are committed within 24 hours of eating bread.
- Bread is baked at temperatures as high as 400 degrees Fahrenheit! That kind of heat can kill an adult in less than one minute.
- Many bread eaters are utterly unable to distinguish between significant scientific fact and meaningless statistical babbling.

As you probably already know (from this article and also from watching TV and reading magazines and newspapers), statistics can be dangerous. They can be used to manipulate public opinion, sell you products or services you don't need, change policy decisions, and significantly affect our lives. Therefore, it is important that we have a good understanding of where they come from (in addition to the subject matter they supposedly represent). Many Americans probably feel that statistics are nothing more than fancy lies (so goes the saying "Lies, damned lies, and statistics!") and statisticians are the wily spin doctors. Conversely, others feel that if data can be put into statistical form, it must be true. "Numbers don't lie!" Whatever your feelings or experiences with statistics, it is important to view statistical claims with a critical and knowledgeable eye. That means we need to understand enough about statistics (what they mean, how they are generated, and what their limitations are) to know when they are appropriate and when they are not.

This course introduces students to the discipline of statistics as a science of understanding and analyzing data. Throughout the semester, students will learn how to effectively make use of data in the face of uncertainty: how to collect data, how to analyze data, and how to use data to make inferences and conclusions about real world phenomena.

Note that this course does not employ the "Las Vegas Effect"-what happens in this class should not stay in this class (unless I say something really stupid). The skills you learn in this class will be invaluable to your success in other Harvard classes and in your future employment. We don't want to teach you just what you need to know-you'll be obsolete in five years! We want you to come away with understanding how to think objectively about information, knowing how to summarize and communicate that information as well as use it effectively in making sound economic decisions.

The course goals are as follows:

- 1) Recognize the importance of data collection, identify limitations in data collection methods, and determine how they affect the scope of inference.
- 2) Use statistical software to summarize data numerically and visually, and to perform data analysis.
- 3) Have a conceptual understanding of the unified nature of statistical inference.
- 4) Apply estimation and testing methods to analyze single variables or the relationship between two variables in order to understand natural phenomena and make data-based decisions.
- 5) Model numerical response variables using a single explanatory variable or multiple explanatory variables in order to investigate relationships between variables.
- 6) Interpret results correctly, effectively, and in context without relying on statistical jargon.
- 7) Critique data-based claims and evaluate data-based decisions.

Tips for success:

- 1) The class lecture notes will be available at least 24 hours before each class. Read them before coming to class, and then review again after the class is over.
- 2) Be an active participant during lectures and section.
- 3) Ask questions - during class or office hours, or by email. Ask me, the TFs, and your classmates.
- 4) Do the problem sets - start early and make sure you attempt and understand all questions.
- 5) Start your project early and allow adequate time to complete them.
- 6) Give yourself plenty of time to prepare a good cheat sheet for exams. This requires going through the material and taking the time to review the concepts that you're not comfortable with.
- 7) Do not procrastinate - don't let a week go by with unanswered questions as it will just make the following week's material even more difficult to follow.

Tentative Course Schedule

Class 1: The Nature of Statistics

- Compare/Contrast Stat 100/101/102/104
- Review Important Class Information
- General Introduction

Part 1: Getting the Information You Need

Class 2: Getting Good Data

- Samples versus Populations
- Random Sampling
- The Valid Survey
- How to Sample Badly

Part 2: Descriptive Statistics

Class 3: Descriptive Statistics, Part I

- Graphical Methods and Frequency Tables
- Center of a Distribution
- Quartiles
- Spread of a Distribution
- The Boxplot and Stata

Class 4: Descriptive Statistics, Part II

- Linear Transformations of Data
- Chebyshev's Inequality and the Empirical Rule
- Detecting Outliers
- Skewness and Kurtosis
- Transforming Skewed Data

Class 5: Measures of Association and Portfolios

- Covariance of Two Variables
- Correlation of Two Variables
- The Risk and Return of Stock Portfolios

Class 6: Introduction to Regression

- The Equation of a Line
- Fitting a Line to Data
- Using Regression for Prediction
- Market Models, an Example from Finance

Part 3: Probability and Probability Distributions

Class 7: Basic Probability Theory and Conditional Probability

- What is Probability?
- The Random Experiment, Outcomes and Events
- Probability Rules
- Joint Probability and Contingency Tables
- Conditional Probability

Class 8: Independence and 2x2 Tables

- Independence
- The Multiplication Rule
- Working With 2x2 Tables
- Randomized Response

Class 9: Decision Theory

- The Key Elements of a Decision
- Decisions Based on Extreme Values
- Decision Trees and Expected Monetary Value

Class 10: University Holiday

Class 11: Discrete Random Variables and Probability Distributions

- What is a Random Variable?
- Discrete versus Continuous Random Variables
- Probability Distributions for Discrete Random Variables
- Expected Value and Variance for Discrete Random Variables
- The Binomial Distribution

Class 12: Continuous Probability Distributions

- Continuous Random Variables
- Probability Distributions for Continuous Random Variables
- The Uniform Distribution
- The Normal Distribution
- Reverse Look-Up

Class 13: Jointly Distributed (Discrete) Random Variables

- Jointly Distributed Discrete Random Variables
- Covariance and Correlation for Random Variables
- Expected Value and Variance of the Sum of Random Variables
- Conditional Distributions
- Conditional Expectation
- Independence

Class 14: Introduction to Simulation

- Using Stata for Simulation
- Using Google Docs for Simulation
- Simulating from Distributions
- Modeling Processes with Random Inputs

Class 15: Review of Classes 1-14

Class 16: Exam 1: will cover material from Lectures 1-14

Class 17: Distribution Minutiae

- The Poisson Distribution
- The Distribution of Sums of Normals
- Normal Probability Plots
- The t distribution

Part 4: Inference from Samples to the Population

Class 18: Sampling Distributions

- Sampling From a Population
- Sampling Distribution of the Sample Mean
- The Central Limit Theorem

Class 19: Point Estimation

- Sampling Distribution of the Sample Proportion
- General Concept of Point Estimation
- Estimates and Estimators
- Properties of Estimators

Class 20: Confidence Intervals for Means and Proportions

- What is a Confidence Interval?
- Confidence Interval for the Population Mean, known variance
- Confidence Interval for the Population Mean, unknown variance

Class 21: More Confidence Interval Topics

- Confidence Interval for the Population Proportion
- Estimating the Sample Size
- The Bootstrap

Class 22: University Holiday

Class 23: University Holiday

Class 24: University Holiday

Class 25: Introduction to Hypothesis Testing, Part I

- Concepts of Hypothesis Testing
- Type I and Type II Errors
- Hypothesis Testing the Population Mean Using Confidence Intervals
- Hypothesis Testing the Population Proportion Using Confidence Intervals

Class 26: Introduction to Hypothesis Testing, Part II

- Alpha Levels and Significance
- The Test Statistics
- Hypothesis Testing the Population Mean Using a Test Statistic
- Hypothesis Testing the Population Proportion a Test Statistic

Class 27: Introduction to Hypothesis Testing, Part III

- What is a p-value?
- Hypothesis Testing the Population Mean Using a P-Value
- Hypothesis Testing the Population Proportion Using a P-Value
- Testing Normality Using a P-Value

Class 28: Comparing Two Groups

- Review of Sums of Random Variables
- Confidence Intervals and Hypothesis Tests for Two Population Proportions
- Confidence Intervals and Hypothesis Tests for Two Population Means
- Matched Data

Class 29: Chi Square Tests

- Chi-Square Goodness of Fit Test
- Chi-Square Test for Independence

Class 30: Review of Classes 17-29

Class 31: Exam 2: will cover material from Lectures 17-29

Part 5: Modeling

Class 32: Simple Linear Regression Redux

- The Linear Regression Model (again)
- The Explanatory Power of a Linear Regression Equation

Class 33: Understanding the Regression Output

- Confidence Intervals for the Regression Coefficients
- Prediction Intervals
- The 2000 Election-a Regression Example

Class 34: Regression Hypothesis Testing

- Hypothesis Testing in Regression Models
- Worked Through Examples

Class 35: Multiple Regression

- The Multiple Regression Model
- Standard Assumptions for the Multiple Regression Model
- The Explanatory Power of a Multiple Regression Equation
- Confidence Intervals and Hypothesis Tests for Individual Regression Parameters

Class 36: Extending the Multiple Regression Model, Part I

- Variable Selection
- Dummy Explanatory Variables

Class 37: Extending the Multiple Regression Model, Part II

- Regression Diagnostics
- Residual Analysis
- Outliers and Nonlinearity

Class 38: Extending the Multiple Regression Model, Part III

- Normality
- Heteroscedasticity
- Multicollinearity
- Worked Through Examples

Part 6: Selected Advanced Topics

Class 39: Analysis of Variance

- Observational Studies
- Randomized, Comparative Experiments
- Components of the Variance
- The One-Way Analysis of Variance

Class 40: Nonparametric Analogues of Parametric Tests

- Mean versus Median
- The Sign Test
- The Mann-Whitney Test
- Kruskal-Wallis Nonparametric ANOVA

Class 41: Fallacies and Traps and Course Review

- Correlation is Not Causation
- Practical versus Statistical Significance
- Coincidences
- Course Review