

**Course Syllabus**

**COURSE OVERVIEW**

In this course, you will learn the basics of using statistical data to do research on economic and social issues.

Doing statistical research in the social sciences requires several skills: an understanding of some basic probability and statistical theory, knowledge of how that theory is applied in practice to different kinds of data and different kinds of issues, and basic computer skills necessary to obtain and analyze data.

Doing statistical research also requires the kinds of skills you would use in a history or philosophy project: formulating a well-defined question, surveying previous research on the topic, generating original ideas and insights that are supported by evidence and careful arguments, and presenting your analysis and conclusions to others in a well-written paper.

During the three hours per week of regular class meetings, we will cover some basic theory of probability and statistical inference, and then use that theory to understand the rationale for a variety of applied statistical methods. This aspect of the course will mostly follow a textbook and will be the basis of problem sets and exams.

During the regular class meetings, we will develop the fundamental concepts of interval estimation (which has to do with the notion of “margins of error”) and p-values (which have to do with the notion of “statistical significance”). The terms margin of error and statistical significance arise frequently in conversation and the media, but usually without any explanation of exactly what they mean. It turns out that their meanings are pretty subtle, and to define them precisely it is necessary first to develop some basic tools and concepts of probability and statistics. Developing these tools, and then defining margins of error and statistical significance, are central goals of this course. We will then see that a wide variety of statistical analyses that superficially appear very different from each other are unified in the sense that they are all built on the fundamental concepts of interval estimation (margins of error) and p-values (statistical significance).

One of the methods in which the concepts of interval estimation and p-values are encountered is linear regression. Regression is one of the statistical techniques most widely used in economics, and the course will culminate in an introduction to this topic. Economics 304, Introduction to Econometrics, includes extensive study of the linear regression model and a variety of extensions.

The objectives of the course related to conceiving of and conducting a statistical research project and presenting the results in an original paper will be pursued in a weekly lab. In the lab you will learn about the nuts and bolts of working with large data sets on a personal computer, as well as the conceptual processes involved in doing empirical social science research. Working in groups of about three or four students, you will choose topics for a research project that you will work on throughout the semester. The research project will be the basis of an original research paper that you will turn in at the end of the semester.

## TEXTBOOKS

There is one required book for this course:

Devore, J. L. and Berk, K. N. (2012). *Modern Mathematical Statistics with Applications*, 2<sup>nd</sup> edition. Springer Texts in Statistics. ISBN: 978-1-4614-0390-6.

It is important to get exactly this edition; make sure that the ISBN of the copy you buy matches the ISBN listed above.

## SUPPLEMENTARY MATERIAL AND MOODLE

Throughout the semester, important topics that are not covered in the textbook will be presented in class. For these topics, supplementary handouts or readings will be assigned. These materials will be posted on Blackboard.

## CALCULATORS

You will need a calculator for this class, but you should use it only for a very limited set of arithmetic operations: addition, subtraction, multiplication, division, finding square roots, raising numbers to integer powers, and calculating factorials.

You will need to use a calculator during the midterm and final exams for this class, but you will be allowed to use it only for the operations specified above.

## PROBLEM SETS

Problem sets (some involving work on a computer) will be assigned about once a week. The due dates for the problem sets are listed on the attached calendar of assignments.

You may skip one problem set of your choosing without penalty. (If you do all the problem sets, I will drop the one with the lowest grade when calculating your problem set average.)

No credit will be given for problem sets turned in after the beginning of class on the day they are due.

You are encouraged to work on the problem sets with other members of the class. Be sure, though, that when all is said and done, you have understood the material and can solve similar problems on your own. In particular, the final version of your answers that you write up to turn in must be done without any assistance from anyone. A document titled “Problem Set Guidelines,” which is posted on the Moodle site for this course, explains in more detail the ways in which you are and are not allowed to collaborate with other students on problem sets for this course.

**IMPORTANT NOTE:** Problem sets must be turned in at the beginning of the class on the day they are due. Problem sets turned in at any other time or in any other manner will be returned to you ungraded and will not receive credit. For instance, it is not permissible to hand the problem sets in directly to the TA or grader, leave them in my mail box, send them to me by campus mail, slide them under my office door, leave them with the economics department secretary, fax them to me, or e-mail an electronic version of them to me.

Please do not turn in problem sets before they are due. If you know that you will not be attending class on a day when a problem set is due, please give your problem set to someone else in the class to bring to class and turn in for you on the day it is due.

## **EXAMS**

There will be one mid-term exam and one final exam.

The mid-term will be given during the lab period (7:30-9:30 PM) on Tuesday, March 17. This will be a closed-book, closed-notes exam.

The final exam will be given during the lab period (7:30-9:30 PM) on Tuesday, April 28. This will be a closed-book, closed-notes exam, with one exception: you will be allowed to bring to the exam one page of notes that you have written in advance.

***Exams must be taken at the time they are scheduled.*** No make-up exams will be given. Please check your calendar now for any conflicts with the exam dates, and inform me immediately if any exist.

## **COMPUTER LAB AND GROUP RESEARCH PROJECT**

Starting on January 27, there will be a computer lab in Hilles 204 every Tuesday from 7:30-9:30 PM. The lab is an integral part of the course.

In the first few weeks of the lab, you will learn the basics of Stata, the statistical software we will be using in this course. Several labs will involve computer exercises that you will need to work on before and/or after the lab. For purposes of grading, each of these exercises will be treated as the equivalent of one problem set.

Early in the semester, you will be asked to form small groups of students to work together in the lab throughout the semester. Each group will choose a topic, problem, or set of questions that is of interest to the members, and that might be illuminated by the analysis of some relevant data. Over the course of the semester, you will then go through the entire process of writing an original research paper on your chosen topic: reviewing the existing literature on the topic, finding suitable data from a reliable source; downloading the data, extracting the relevant parts, and converting them into a format Stata can read; cleaning and processing the data to prepare them for analysis; analyzing the data; comparing your findings to theories and empirical evidence reported in the existing literature on the topic; drawing appropriate conclusions; and finally writing up your study in a professional-quality report of about 14-16 pages (excluding figures, tables and appendixes).

For groups with one or more members who are seniors, the final paper will be due on Saturday, May 9, at 5:00 PM.

For groups with no seniors, the final paper will be due Friday, May 15 at noon.

In addition to the final report, several installments for the project—shorter write-ups of your progress at various stages of your work—will be due as the semester progresses. Detailed instructions for each of these installments, as well as for the final paper, will be distributed; the due dates are listed on the Calendar of Assignments for this course.

## **ATTENDANCE**

***Regular attendance at this class is very important.*** Material not covered in the book will frequently be presented in class, and some problems not in the book will be discussed and assigned in class. Adjustments to the content of assignments and the dates they are due may also be announced in class. You are responsible for all course material, assignments and announcements given in class.

Attendance at the labs is essential for all the reasons above, and because for most of the semester you will be collaborating with several other students on your research project. It is important to be in lab so that you can participate fully in the work of your group.

Chronic absenteeism, whether from class or from lab, will be grounds for a reduction of your final course grade, regardless of the grades you earn for the work you have done for the class.

## **GRADE DETERMINATION**

Subject to the above comment on attendance, your final grade will be determined by the following weighting of the problem sets, the computer lab project, and the exams:

Problem Sets: 10 percent

Computer Lab: 20 percent

Mid-Term Exam: 35 percent

Final Exam: 35 percent

## **ACCOMMODATIONS FOR STUDENTS WITH DISABILITIES**

Students who think they may need accommodations in this course because of the impact of a disability are encouraged to meet with me privately early in the semester. Students should also contact Rick Webb, Coordinator, Office of Disabilities Services ([rwebb@haverford.edu](mailto:rwebb@haverford.edu), 610-896-1290) to verify their eligibility for reasonable accommodations as soon as possible. Early contact will help to avoid unnecessary inconvenience and delays.

## **COMMUNICATION BY E-MAIL**

E-mail is usually not a good way to communicate with me about matters having to do with this course. Simply because of the volume of e-mail I receive, I may not be able to read an e-mail message you send me until two or three days after I receive it, and it may take several more days for me to respond.

The best way to handle most questions is to come see me in person during my office hours. In an effort to make it as easy as possible for you to see me, I have scheduled office hours on three different days each week. Nonetheless, the fact that I ask you to come to my office hours to discuss questions you have about the class does require you to start assignments earlier and plan your time more carefully than would be the case if I were able to reply promptly to questions sent by e-mail.

You should not use e-mail to communicate about urgent or time-sensitive issues. To give one example, if you send me an e-mail the day before an exam is scheduled to tell me that for some reason you will not be able to attend the exam, you should assume that I will not read the message until after the exam.

If you need to get in touch with me about an urgent or time-sensitive matter, please contact Kim Minor, the economics department administrative assistant, at 610-795-1701, or [kminor@haverford.edu](mailto:kminor@haverford.edu). She is available during regular business hours, and usually knows how to contact me in an emergency.

### **OFFICE HOURS**

My office is Stokes 203B. My e-mail address is [rball@haverford.edu](mailto:rball@haverford.edu) and my phone number is 610-896-1437.

This semester I will have office hours at the following times:

Monday 1:45-3:15 PM

Thursday 1:45-3:15 PM

Friday 1:45-3:15 PM

# CALENDAR OF ASSIGNMENTS

	Problem Sets	Exams	Lab Project
Monday, January 26	Problem Set #1		
Monday, February 2	Problem Set #2		
Monday, February 9	Problem Set #3		
Friday, February 13			Topic and data ideas
Monday, February 16	Problem Set #4		
Monday, February 23	Problem Set #5		
Monday, March 2	Problem Set #6		
Friday, March 6			Project proposal
<b><i>Spring Break: Monday, March 9 through Friday, March 13</i></b>			
Monday, March 16	Problem Set #7		
Tuesday, March 17		Midterm	
Tuesday, March 24	Problem Set #8		
Monday, March 30	Problem Set #9		
Monday, April 6			Data report
Tuesday, April 7	Problem Set #10		
Monday, April 13	Problem Set #11		
Friday, April 20	Problem Set #12		
Friday, April 24			Preliminary draft
Monday, April 27	Problem Set #13		
Tuesday, April 28		Final	
<b><i>Last Day of Classes: Friday, May 1</i></b>			
Saturday, May 9			Final paper (groups with seniors)
Friday, May 15			Final paper (groups without seniors)