ECON 2300: Statistics for Economists, Fall 2022
University of Pennsylvania

Syllabus

(Tentative: August 15, 2022)

Instructor:

Wayne Gao

Email: waynegao@upenn.edu

Office: PCPSE 630

Office Hours: Monday 2:30-3:30pm & Tuesday 10-11am

Lectures:

Time: Tuesday & Thursday, 1:45pm-3:15pm

Location: TBA

Teaching Assistants:

TBA

Recitation Sections:

TBA

Course Description:

This course focuses on data description, probability, and statistics, as relevant for economics. Topics include economic data sources, descriptive statistics, probability distributions and population moments, sampling distributions and sample moments, statistical estimation, confidence intervals, hypothesis testing, and an introduction to linear regression. Focus is on both theoretical and practical issues involved in the substantive interpretation of economic data using statistical and econometric techniques. Empirical case studies are discussed throughout. Students will be guided to learn the statistical computing language R, and will be required to carry out various statistical analyses in R.

This course fulfills the College's Quantitative Data Analysis requirement.

Prerequisites:

ECON 0100 AND ECON 0200 AND MATH 1400 AND (MATH 1410 OR MATH 1510)

To do well in this course you will need to be comfortable with algebra, manipulating sums, differentiation and partial differentiation, solving unconstrained optimization problems, and integration.

Textbook:

The official textbook for this course is "*Introductory Statistics* (4th edition)", by Sheldon M. Ross, Elsevier. Digital copies of the textbook are freely available online via the Penn Library or the following webpage (PennKey login required): https://www.sciencedirect.com/book/9780128043172/introductory-statistics . While I suggest that you complete the assigned readings, my lecture slides, which will be posted online at the start of each week, are the final authority on course material. In particular, you are not responsible for material in the textbook unless it is also covered in lecture, but you are responsible for material from lecture even if it is not covered in the textbook.

Required Software: R

We will use the statistical package R via a front-end called RStudio throughout the course. Both R and RStudio are open source and free. First, download and install R from http://cran.r-project.org/. Second, download and install RStudio by visiting https://rstudio.com/products/rstudio/download/ and clicking "Download" under the free RStudio Desktop version. You might need to further choose the right version for your operating system (Windows/Mac). Make sure you download and start using R as the homework assignments will involve coding and running some data analysis on R .

You will be *guided to learn* R *by yourself*. The lectures, recitations and homework assignments will provide the necessary guidance, tutorials or instructions about the R basics and the specific R commands used in this course. However, this course will *not* provide a comprehensive and systematic coverage of computer programming with R, for the following reasons: (1) The focus of this course is on probability theory and statistical method, and we have a limited amount of time for instruction. (2) The heterogeneity in the students' past exposure to coding also makes it inefficient to teach coding in lectures. (3) The best way to learn coding (in R) is "learning by doing", and the process of figuring out how to code on your own is in many ways more important and useful than the "correct codes" per se.

That said, *students are highly encouraged to seek help from the Instructor, the TAs, and other students on coding.* Often times, the questions and problems students encountered in coding

tend to be very specific and detailed. Thus "one-on-one tutoring" during office hours, along with discussions among students (e.g. on Ed Discussion), are often much more effective and efficient than lectures.

In addition, an invitation link for *free* access to *DataCamp*, an online learning platform for computer programming and data analysis, will also be sent out on Canvas. The platform allows you to learn and practice R interactively on your web browser or mobile device. Those students who are interested in learning computer prorgamming in R more systematically are encouraged to take advantage of DataCamp. Homework assignment may contain R exercises on DataCamp.

Finally, the following textbooks/references are freely available online:

- *Hands-On Programming with R*, by Garrett Grolemund, available at https://rstudio-education.github.io/hopr/

- *R for Data Science*, by Hadley Wickham and Garrett Grolemund, available at https://r4ds.had.co.nz/

Course Website: Canvas

We will use *Canvas* to make course announcements, post course material, answer questions about course material and respond to private messages from individual students regarding personal issues. All written communication for ECON 103 should be sent through *Canvas*, *not* to the instructors' personal email accounts.

Discussion Board: Ed Discussion

You are highly encouraged to the discussion board, *Ed Discussion*, which is accessible via *Canvas* for Q&A about course material. By asking your questions and answering others' questions on *Ed Discussion*, you create a positive externality: other students benefit from your questions/answers and you benefit from theirs. The instructor and TA's will actively moderate *Ed Discussion* both to answer questions and approve (or correct) answers written by your fellow-students. In order to encourage participation, and reward the benefit that students bring to the discussion, I will award bonus points for participation in Ed Discussion. This could be up to 5% of the final grade. Bonus points will be awarded for constructive questions, answers, and notes on Ed Discussion, as well as in-class/online student presentations about assigned Ed Discussion questions during lectures. Even if you post anonymously to other students (but not to instructors: if you post anonymously to the instructors as well, I will not be able to recover your identity and award you bonus points) on Ed Discussion, we can still award you bonus points for online participation. Bonus points are discretionary.

Departmental Course Policies:

All Economics Department course policies are in force in ECON 2300 even if not explicitly listed on this syllabus. See https://economics.sas.upenn.edu/undergraduate/course-information/course-policies for full details.

Academic Integrity:

All suspected violations of the code of academic integrity as set forth in the Pennbook will be reported to the Office of Student Conduct. Confirmed violations will result in a failing grade for the course.

Grading:

Grades for this course will be determined based on 10 homework assignments, 7 quizzes, 2 midterms, a comprehensive take-home final exam, and Ed Discussion bonus points:

$$
\begin{aligned}
\text{Course Score} = \ &(20\% \times \text{Homework}) + (5\% \times \text{Online Quizzes}) \\
&+ (25\% \times \text{Midterm 1}) + (25\% \times \text{Midterm 2}) + (25\% \times \text{Take-Home Final}) \\
&+ (3\% \times \text{Bonus})
\end{aligned}
$$

Homework:

Homework assignments will be posted on Canvas each Tuesday, starting from the second week of the semester. Homework will be collected and graded. Group work is encouraged, but you will have to submit your own answers. When calculating your homework average, I will drop your two lowest scores and weigh the remaining homeworks evenly.

Quizzes:

There will be seven short online quizzes over the course of the semester, which you will be able to take through *Canvas/Quizzes*. You can take each quiz at anytime of your choice duirng a 24-hour window, but only one attempt will be allowed. You will need to finish the quiz in 10 minutes once it starts, so please be ready before you start the quiz on Canvas. The quiz is open-book, but the work should be your own. Communication with any other person is prohibited. Unless otherwise indicated, each quiz will cover all the material before, and including, the last lecture with higher weight given to material that was not tested in previous

quizzes. When calculating your quiz average, I will drop your two lowest scores and weight the remaining quizzes evenly. There will be no makeup quizzes so be sure to use your two "free skips" carefully.

Exams:

There will be two 80-minute in-class midterm exams and a take-home final exam. Each midterm is worth 25% and the take-home final is also worth 25% of your grade.

The midterms will be closed-book, and no cheat sheets are allowed. You may write in pencil or pen on your exam. You may use a calculator. We will check ID cards at each midterm. Your exams will be photocopied before being returned to you. For those who have approved accommodations from the ,

The take-home final exam is open-book, but the work should be your own. Any form of communication with any other person is strictly prohibited. The take-home final will involve programming and data analysis in R.

Regrade Requests:

Exam regrade requests must be made in writing within a week of receiving your graded exam. As we re-grade the entire exam, your score could rise or fall. You may not discuss your answers with the Teaching Assistants or the instructor before submitting a regrade request. To deter regrade-related cheating, we will photo-copy the exams of all or a random subset of students before returning the exams.

Course Curve:

We typically try to target an average GPA in the range between 3.0 and 3.2, or slightly above a B average. In a nutshell, I will give about 30% As and A-s, 40-50% Bs and 20-30% Cs. If necessary, I will curve the grades so that they approximately fall into these ranges. I reserve grades below a C-minus for those cases in which a student fails to attain a minimum level of basic competence in statistics, an absolute rather than relative standard.

Accommodations for Students with Disabilities:

If you need a disability accommodation for this course, please make sure to register in advance through the Weingarten Center online portal (https://wlrc.vpul.upenn.edu/). The time limit for Canvas quizzes will be automatically extended for those with approved extension for timed tests. However, for the in-person midterms, you will need to take the exams with any approved

accommodations at the Weingarten Center. Please note that this arrangement is *not automatic*, and requires that you make appropriate appointments with the Weingarten Center in advance.

Course Schedule (Tentative)

Date	Day	Lecture	HW Posted	Quiz
Aug. 30	Tue	Introduction		
Sep. 01	Thu	Summary Statistics: I		
Sep. 06	Tue	Summary Statistics: II	HW1	
Sep. 08	Thu	Basic Probability I		Q1
Sep. 13	Tue	Basic Probability II	HW2	
Sep. 15	Thu	Basic Probability III		
Sep. 20	Tue	Discrete Random Variables I	HW3	Q2
Sep. 22	Thu	Discrete Random Variables II		
Sep. 27	Tue	Discrete Random Variables III	HW4	
Sep. 29	Thu	Discrete Random Variables IV		Q3
Oct. 04	Tue	**In-Class Midterm 1**		
Oct. 06	Thu	**Fall Break: No Lecture**		
Oct. 11	Tue	Continuous Random Variables	HW5	
Oct. 13	Thu	Normal Random Variables		
Oct. 18	Tue	Sampling Distribution		
Oct. 20	Thu	Estimation I		Q4
Oct. 25	Tue	Estimation II	HW6	
Oct. 27	Thu	Confidence Intervals I		
Nov. 01	Tue	Confidence Intervals II	HW7	Q5
Nov. 03	Thu	Hypothesis Testing I		
Nov. 08	Tue	Hypothesis Testing II	HW8	
Nov. 10	Thu	**In-Class Midterm 2**		Q6
Nov. 15	Tue	Linear Regression I		
Nov. 17	Thu	Linear Regression II		
Nov. 22	Tue	Linear Regression III	HW9	
Nov. 24	Thu	**Thanksgiving Break: No Class**		
Nov. 29	Tue	Linear Regression IV		
Dec. 01	Thu	Linear Regression V	HW10	Q7
Dec. 06	Tue	Concluding Remarks		
Dec. 08	Thu	Reserved Lecture		
		Take-Home Final Exam: Time TBA		