ECO S394D: Probability and Statistics

Instructor: James Scott (james.scott@mccombs.utexas.edu)

Scheduled class time:
M T W Th, 6:30AM to 7:20 AM and 2:30 PM – 3:20 PM, (via Zoom, link on Canvas)
Fridays 1:30 PM to 2:30 PM (this will be a TA session)

Office hours:
Friday 6:30AM to 7:20 AM and 3:30 PM – 4:20 PM (via Zoom, link on Canvas)

Teaching Assistants

Jonathan Garita (jgth24@utexas.edu)
Office hours Tuesdays 3:30 PM to 5:30 PM, U.S. central time (via Zoom, link on Canvas)

Qingsong Pan (qingsongpan@utexas.edu)
Office hours Wednesdays, 9 PM to 11 PM U.S. central time (via Zoom, link on Canvas)

Background and objectives. This course is an intermediate level introduction to probability theory and statistics. The main goal is to prepare students for econometrics. Students taking the course should have a working knowledge of calculus and basic descriptive statistics. The emphasis will be on conveying the basic principles of statistical theory. This is fast-paced course with a condensed schedule, so it is essential to attend lectures regularly and keep up with the material.

Course materials. Your main point of reference will be the lecture notes from class, provided as slides via Canvas As a reference on many probability topics we will also use: "Introduction to Probability Lecture Notes," by Dimitri P. Bertsekas and John N. Tsitsiklis. This is posted as a PDF file on the course web page. As a reference for some of the material on data analysis, I have posted a set of course notes entitled "Data Science: A Gentle Introduction." You can find both of these in the "References" folder under "Files" within Canvas.

I will also draw on "All of Statistics: A Concise Course in Statistical Inference" by Larry Wasserman. You should not feel obligated to buy this. It is entirely optional. Since the material we're covering is pretty standard, basically any intermediate probability/statistics textbook can be used as a supplement.

Lectures. All lectures will be held via Zoom, with links for live viewing and recordings both posted on Canvas. Owing to the Covid-19 pandemic, our class will have students spanning 14 time zones. Therefore, I have done my very best to make sure that every student, regardless of time zone, has the opportunity to experience at least part of the class synchronously (i.e. in real time, as opposed to fully recorded.) Therefore, I've decided to hold lectures at the following times.

6:30 AM to 7:20 AM, U.S. Central time: Part 1 of that day's lecture. This starts at 12:30 PM

in London, 1:30 PM in Cairo, 5:00 PM in New Delhi, and 7:30 PM in Beijing.

2:30 PM to 3:20 PM, U.S. Central time: Part 2 of that day's lecture. It starts at 8:30 PM in London, 9:30 PM in Cairo, 1:00 AM in New Delhi, and 3:30 AM in Beijing.

All lectures will be recorded, with links posted on Canvas ASAP. Obviously anyone can join any of the "live" lectures they want to! But I've tried to time it so that:

- Everyone has a chance to attend at least half of the class "live" (i.e. synchronously rather than recorded).
- During the morning session, I can be awake and functioning in Texas while students based in Europe, Africa, and Asia can watch in their afternoon/evening, before most people want to be asleep.
- Students in the U.S. who (quite reasonably) aren't awake or in the mood for statistics at 6:30 AM can watch the recording of Part 1 during the "normal" class time slot of 1:30 PM to 2:30 PM US central time, and be ready to pick up Part 2 live at 2:30.
- Students in time zones east of the U.S. can watch the recording Part 2 of the previous day's lecture during their normal daytime hours, before Part 1 of that day's lecture in their afternoon/evening.

Communication. I will post all assignments, announcements, and class materials on our class's Canvas site.

Software. While this course is geared more towards the basic mathematics of probability and statistics, the homework will sometimes include some simulation or data-analysis exercises. We will R and RStudio for this purpose. I'll show you in class how to download and install them.

Assignments and grading. There are five weekly problem sets, each worth 20% of your grade. The weekly problem sets are posted in a single file on Canvas, and your write-up for each week's problems is due on the Friday of that week by 11:59 PM U.S. Central time. So, for example, the first week's problem set is due on the Friday of the first week, and so on each successive week. You will submit your assignments electronically through Canvas. Each assignment should be submit as a single file, or as a link.

In the past, students have found it pretty simple to just take pictures of their written work with a phone, and then stitch the individual pictures together into a single file using any appropriate piece of software (e.g. Mac Preview, Microsoft Word, Google docs. etc). You can also write your solutions in a Google doc (including images, e.g. phone pictures of handwritten work) and submit that link via Canvas. If you do this, make sure you change the permissions so that anyone with the link can view the document.

In grading the homework, we will be looking at the following:

- Did you make an honest, concerted attempt at each problem?
- For a math problem: did you show all your steps? If you were stuck, did you show evidence of attempting to get yourself unstuck, e.g. by writing down results you know that you think might be relevant?
- For a data-analysis problem: did you attempt to address all parts of the question?

Did you include figures/tables where appropriate? Did you write up your solution with some semblance of narrative flow and professionalism (good), or did you just copy and paste a bunch of R code without much in the way of explanation (bad)?

- And finally, of course, did you make any significant technical mistakes?

The solutions will be discussed in class each week.

There are no exams in this course.

Course outline

I. Basics of probability

Introduction: Random experiment. Set operations, Kolmogorov's rules of probability. Interpretations of probability. Basic rules.

Joint, conditional, and marginal probability. Rule of total probability. Independence, compounding, and confounding. Bayes' rule.

Random variables and their distributions. Discrete random variables and their PMFs: binomial, Poisson. Continuous random variables and the CDF/PDF: normal distribution. Transformations of random variables.

Properties of distributions: expectation (mean), variance, conditional expectation, conditional variance, covariance and correlation.

The bivariate normal distribution. Regression to the mean.

II. Introduction to data analysis

Basic data visualization

Least squares: fitting equations to data.

Nonlinear equations: polynomials, exponential growth/decay, power laws

If time: Model for groups: dummy variables, interactions.

III. Statistical theory

Random sampling, i.i.d. random variables. Population parameters and sample statistics (estimators). The concept of a sampling distribution. Bootstrapping estimators. Confidence intervals.

Hypothesis testing: basic concepts. Null and alternative hypotheses, confusion matrix, Type 1 and 2 errors. Alpha level vs. p-values. Permutation tests.

Large-sample properties of estimators: Convergence in probability. The Law of Large Numbers. Convergence in distribution. The Central Limit Theorem. Inference based on asymptotic normality. Consistency.

General methods for constructing estimators: maximum likelihood, method of moments.

University policies:

Honor code: "The core values of the University of Texas at Austin are learning, discovery, freedom, leadership, individual opportunity, and responsibility. Each member of the University is expected to uphold these values through integrity, honesty, trust, fairness, and respect towards peers and community."

Any student with a documented disability (physical or cognitive) who requires academic accommodations should contact the Services for Students with Disabilities area of the Office of the Dean of Students at 471-6259 (voice) or 471- 4641 (TTY for users who are deaf or hard of hearing) as soon as possible to request an official letter outlining authorized accommodations.

Students who violate University rules on scholastic honesty are subject to disciplinary penalties, including the possibility of failure in the course and dismissal from the University. Since dishonesty harms the individual, fellow students, and the integrity of the University, policies on scholastic dishonesty will be strictly enforced.

Occupants of buildings on The University of Texas at Austin campus are required to evacuate buildings when a fire alarm is activated. Alarm activation or announcement requires exiting and assembling outside. Familiarize yourself with all exit doors of each classroom and building you may occupy. Remember that the nearest exit door may not be the one you used when entering the building. In the event of an evacuation, follow the instruction of faculty or class instructors. Do not re-enter a building unless given instructions by the following: Austin Fire Department, The University of Texas at Austin Police Department, or Fire Prevention Services office.

Students requiring assistance in evacuation shall inform their instructor in writing during the first week of class.

Behavior Concerns Advice (BCAL): 512-232-5050.