Dr. Christian Rojas

Office Hours: by Appointment
rojas@umass.edu

Resource Economics 701
Fall 2021

Description

This course will focus on probability theory and statistical inference, the foundations of econometric analysis. The department now uses this course (ResEcon 701) as the first course in our econometrics series. This course is followed by two econometrics courses (ResEcon 702 and 703).

Probability theory is the building block that will allow you to understand estimation and statistical inference. In this course, we'll first focus on univariate distributions and then move on to multivariate ones, with care taken to differentiate discrete and continuous random variables. We will then move on to estimation and inference. We will not be able to cover all the material related to inference, but we will be able to have a good idea of different estimation methods and the basics of estimators' properties.

COVID Policy

Students are expected to follow all COVID-related policies on campus. At the start of the Fall 2021 semester, there are two policies in effect that deserve special mention. First, students are required to follow the COVID-19 Daily Self-checklist. Students who are ill must not attend class, and they will be offered reasonable accommodations for missed work. Second, students must follow the masking policy while it remains in effect. Your instructor will remind you to wear a mask that securely covers your mouth and nose, and direct you to a nearby mask station if you do not have one. If you are unwilling to mask, you will be asked to leave the class. If you do not leave the class, the instructor will end the class, and the Dean of Students office will be informed that you have disrupted class and violated the Guidelines for Classroom Civility and Respect. Anyone with a mask exemption must provide prior written notice to the instructor and must maintain at least six feet of distance from faculty and other students at all times.

Objectives

The goal is to present probability theory and statistical inference concepts at an intuitive and practical level. For this, I will make efforts to provide you with practical experience working with datasets and the statistical software R. A goal at the end of the class is that you will be able to code, from scratch, (simple versions of) estimation routines such as OLS and maximum likelihood. One objective of the course is to use the statistical software's ability to carry out simulations and to use matrix algebra to illustrate key statistical concepts. A byproduct of this is that you will obtain a skill (knowledge of R) that will be useful in other courses (not only those with a statistical focus) as well as, most likely, later in your professional life.

Texts and Software

The main text is ***Mathematical Statistics with Applications***, by Wackerly, Mendenhall and Scheaffer (WMS). I will also refer to ***Learning and Practicing Econometrics***, by Griffiths, Hill and Judge (GHJ) (you may want to get this book as it will be used later in other econometrics courses; i.e., 702), as this text will allow to bridge a gap between the purely statistical treatment of WMS and that of an econometrics. The topics in this class will largely be those of a first course in statistics, but we plan to look at the material, as much as possible, from an economist's and practitioner's angle. If you have a previous econometrics text, you'll find most topics covered in the first few chapters or perhaps an appendix. Reading multiple references, reading different presentations/explanations for this material can be helpful.

R is an open source software that can be downloaded and installed on any computer. The only requirement to

get started with this for this course is to install both R and the R Studio (the interface that makes the R experience easy and powerful). We will be spending a considerable amount of time in R, and the main purpose will be pedagogical. But as time progresses you will notice that the time you will invest in learning/using R will be well spent. R is a powerful and constantly evolving software that is broadly used across disciplines and widely accepted in both academics and industry as the leading statistical analysis tool. Thus, chances are that you will become a frequent user of R both later in your studies as well as in your professional life. In short, becoming proficient in R is a transferrable and highly marketable skill.

Prerequisites

You must be familiar with basic statistical concepts, linear (matrix) algebra and calculus. You do not need to have previous knowledge of R. I will provide guidance as to how to get started in R and support during the semester as to how to go about programming. However, a large portion of your success will depend on your own initiative to struggle with and solve coding issues yourself - as is expected from graduate students. A great advantage of R is that it is an open source software for which there is a wealth of online information and support.

Examinations

There will be a mid-semester exam and a final exam; the final is by nature comprehensive but will primarily focus on topics following the mid-semester exam. The mid-semester exam will be in-class and will take place on October 25 (details to follow). The two exams will comprise 40% of your final grade (larger weight given to the exam on which you do best).

Problem Sets and In-Class Work

There will be several problem sets (~8) and in-class assignments that. Problem sets will comprise theoretical and/or applied material and will comprise 25% of your grade. I don't want you to word process your problems sets, but your handwriting must be immaculate.

For the in-class assignments will be asked to work on a dataset/code; you will work individually on these assignments, but you can engage in collaboration with other students in the class (I do, however, strongly suggest that you try to solve the problem(s) yourself through different means before asking the solution to someone else – the easy way will not teach you much!). The in-class assignments throughout the semester have a total weight of 35%. We will use the second half of our scheduled class time to work on these.

Grading

I anticipate the following thresholds for assigning final grades based on my prior experience: greater than 90 is A; 85 to 90 is A-; 80 to 85 is B+; 75 to 80 is B; 70 to 75 is B-; 65 to 70 is C+ and 60 to 65 is C.

More on Computer Software

As indicated, R will be an important pedagogical tool. Thus, the plan for each class is to have its first half dedicated to the theory/concepts, and the other half to work on an exercise (I call these in-class exercises) that you have to submit for credit (usually at the end of class or, if class is not enough, at the end of the day). Since the software will be a large component of the class and that this software is usually not known by most students that enter our program, I will be spending a significant amount of time introducing you to it and will be providing support in class during the assigned exercises. In addition, our TA (Mr. Augusto Espin) will provide support as needed. I should point out, however, that there are **many** online resources for R and that I expect a lot of proactive behavior from you as should be the case in any graduate class.

Content of the Course

The outline on the following pages is my planned list of topics.

I. Introduction

- Empirical Modeling: WMS: Chapter 1
- Economics, Economic Data, and Inference: GHJ (Ch. 1, and 2.1 – 2.3);

II. Probability Theory

Basic concepts of probability, calculating probabilities, laws of probability, events and random variables

- WMS: Chapter 2

III. Discrete Random Variables and Probability Distributions

Basic concept, probability distribution for a discrete random variable, expected values, examples of discrete probability distributions.

GHJ: 2.5, 2.6
WMS: Chapter 3

IV. Continuous Variables and Probability Distributions

Basic concept, probability distributions, expected values, examples of continuous probability distributions.

GHJ: 2.4 – 2.6
WMS: Chapter 4

V. Additional Topics: Multivariate Probability Distributions and Functions of Random Variables.

Some selected topics from bivariate and multivariate probability distributions, marginal and conditional distributions, covariance, variance, multinomial distributions.

WMS: Chapters 5 and 6.

VI. Sampling Distributions and the Central Limit Theorem.

Basic concept, sampling distributions and the normal distribution, central limit theorem.

GHJ: Chapter 3
WMS: Chapter 7

VII. Estimation.

Introduction/basic concept, classes of estimators, evaluating estimators and confidence intervals.

GHJ: Chapter 3
WMS: Chapters 8 and 9

VIII. Hypothesis Testing.

Basic concept, elements of a test, common tests, errors, power of a test, relationship to confidence intervals, p-values.

GHJ: Chapter 4
WMS: Chapter 10