ECON 2250: Statistics for Economists

Jim Crozier

Fall 2022

E-mail: `rcrozier3@gatech.edu`
Office Hours: TBD
Office: TBD
TA: TBD

Repo: **github.com/jimcrozier/gt_econ2250**
Class Hours: TBD

Course Description

In this class we will cover an introduction into statistical computing which is a cornerstone to economic research. The course will be about one half statistical concepts like counts, summary statistics, and the basics of building statistical models, and one half practical coding skills that are necessary for estimating these concepts. I assume that the incoming student has a general knowledge of mathematical principals, most importantly familiarity with manipulating algebraic equations as well as basic probability theory and intuition. I also assume that the student has a basic understanding of python and notebooks. Although each of these subjects will be taught from the ground up, if this is the first encounter with any of these concepts then I expect a pretty steep learning curve.

Course Objectives

1. Introductory statistics proficiency (counts, summary stats, linear models, significance tests).

2. Data munging and basic SQL syntax.

3. Introductory statistical coding python proficiency.

4. Short video presentation of notebooks skills.

5. Strengthen voice in expressing statistical analysis.

Texts

- Cunningham, Scott. 2021. Causal Inference: The Mixtapes. **Online version is fine.**

- Hastie, Tibshirani, Friedman. 2009. The Elements of Statistical Learning **Online version is fine.**

Honor Code and Plagiarism

You are expected to follow the Georgia Institute of Technology Honor Code at all times. For any questions involving these or any other Academic Honor Code issues, please consult me or http://www.honor.gatech.edu.

Grading Policy

- **40%** of your grade will be determined by your homework videos. There will be a total of 10 homework videos graded (see below) and I will drop the lowest two grades and average the others.

- **20%** of your grade will be determined by the midterm grade.

- **30%** of your grade will be determined by a final exam.

- **10%** of your grade will be determined by a final project.

- **5%** extra credit (see below).

Homework Videos

Each week that we do not an exam (see schedule) you will have to submit a *no longer than* 2 minute video of your notebook assignment for that week, along with the notebook. You are welcome to work with others on the homework assignment, but your video and your code/notebook will need to be your own. The videos will reviewed by the TA and given an 0-100% grade. This is homework so my intention here is that a good shot at the work, and admitting what you can't get should be a 100%. The general idea here is to help you build your voice of your analysis. The grading rubric for the assignment are:

- **50%** does the notebook achieve the directions, or at least make a solid stab at and explain what they couldn't get in their video.

- **40%** does the video express the ideas of the analysis in a straightforward that explains what the author believes and how confident they are in the results. Just screen cast the notebook and explain what you did, and even what you didn't understand.

- **10%** is the video under 2 minutes. Not that faster is always better, but there are a lot of videos to grade and the point of this whole thing is learn to trust your voice, explain how much you believe it, and also state stuff that you didn't get. If you can do that in 30 seconds, all the better for all of us.

Exams

Midterm and final will be cumulative and cover statistical concepts and test the ability to read output of statistical code and explain the outputs in well formulated arguments that address central tendency and dispersion.

Final Project

Your final project will be a slightly larger version of your homework. You will each be assigned a dataset that you will need to present summary statistics on, and a brief statistical analysis. Again, a short video to present will be created with a limit of up to 5 minutes for the presentation. The grading rubric will be the same as homework. You final project should exhibit knowledge of central tendency, dispersion, and most importantly telling a story (even if that story is we cannot find any interesting results).

Extra Credit

The repo for this class will include links to *open source* videos, lectures, papers, datasets, code that link to each lecture. If you *CREATE A **PULL REQUEST** * to add an interesting contribution you can earn 1 point to your grade, up to 5 points.

*Open source is a minimum of MIT license. I will NOT merge a PR if the content does not have a OS license.

Special Accommodations

If you need any special accommodations due to a physical or learning disability, please let me know during the first week of class. In order to receive the requested accommodations you will need to obtain a form from the Access Disabled Assistance Program for Tech Students (ADAPTS) and give me this form. The ADAPTS Office is located in the Smithgall Student Services Building, Suite 210 and the website is http://www.adapts.gatech.edu. Also, if you will be missing any classes for religious holidays or other events, let me know as soon as you know you will be missing class. You will still be required to know the material from that class period.

Schedule (subject to change)

Week 01, 08/29 - 09/02: Load data, summary stats, simple graphs, making videos.

Week 02, 09/05 - 09/09: Continuous vs discrete, counts, descriptive stats.

Week 03, 09/12 - 09/16: Dependent vs independent, light causality.

Week 04, 09/19 - 09/23: Probability review.

Week 05, 09/26 - 09/30: Probability review 2: Conditional Probability.

Week 06, 10/03 - 10/07: Summation notation.

Week 07, 10/10 - 10/14: Correlation and model introduction.

Week 08, 10/17 - 10/21: Review and Midterm: conditional prob to correlation

Week 09, 10/24 - 10/28: Average Treatment Effect

Week 10, 10/31 - 11/04:   Linear regression.

Week 11, 11/07 - 11/11:   Linear regression 2.

Week 12, 11/14 - 11/18:   Standard error.

Week 13, 11/21 - 11/25:   Test for significance.

Week 14, 11/28 - 12/02:   Continuous vs discrete modeling.

Week 15, 12/05 - 12/09:   Bring it all together.

Week 16, 12/12 - 12/16:   Final: ATE and linear reg.