

ECON 2250: Statistics for Economists

Jim Crozier

Fall 2022

E-mail: `rcrozier3@gatech.edu`

Repo: github.com/jimcrozier/gt_econ2250

Class Hours: MWF 12:30 - 1:20

Class location: Van Leer E361

Office Hours: Fridays 2-3PM

Office: 310 of the Rich Computer Center

TA: Christopher Kitchens, email: `christopher.kitchens@gatech.edu`

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 [e.g. 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. Introductory statistical coding python proficiency.
3. Short video presentation of notebooks skills.
4. Strengthen voice in expressing statistical analysis.

Texts

You are not required to purchase these texts, the online versions are fine (although they are both great reference books)

- Cunningham, Scott. 2021. Causal Inference: The Mixtapes. **Online version is fine.**
- Hastie, Tibshirani, Friedman. 2009. The Elements of Statistical Learning **Online version is fine.**

We will also use many online reference materials which will be listed in the github repo.

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>.

Notice that for homeworks, I am completely fine with students working together and using any online materials. That said in many of the homeworks you will each receive different data, and you are responsible for explaining the outcomes yourself.

All videos must be done by yourself, cover your own data (or data seed), and explain your own outputs. If your video presentation does not present your own results, or I do not feel that you can explain your code or results in a way that makes it clear that you understand how it is working you will not receive credit for that assignment.

canvas

Communication about the course and other course materials is available on Georgia Tech's Canvas system. I will upload additional course materials and post course announcements on the course page. You may access Canvas at <https://canvas.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.
- 25% of your grade will be determined by the midterm grade.
- 35% of your grade will be determined by a final exam.
- 5% extra credit (see below).

Your final grades transfer to the Georgia Tech grading system as follows:

Grading Rubric			
A	4.0	Excellent	90%-100%
B	3.0	Good	80%-89.99%
C	2.0	Satisfactory	70%-79.99%
D	1.0	Passing	60%-69.99%
F	0.0	Failure	< 60%

I reserve the right to increase all final grades equally if the material is more difficult than expected, but this is unlikely. At the end of the semester your final grade cannot be changed with extra credit or makeup opportunities; this is for fairness to all students.

Pass /fail policy

If you are taking the class pass/fail, you must achieve a C grade to earn a pass in the class. Please reach out to me at the beginning of the semester if you are taking the class pass/fail.

Homework Videos

Homeworks are due on Thursday morning at 10AM EST. Submit the video (youtube link) and colab notebook link to Canvas.

***** Because I am dropping your two lowest, I am not accepting late homework. Homework submitted at 10:01AM will not be accepted. *****

Each week that we do not have 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.

Our repo will have suggestions on how to record, but you will not be reviewed on production quality, just the results.

Homework Videos

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.

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**

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

Freedom of expression

As a faculty member at Georgia Tech, I respect your rights to the freedom of speech and expression. I am also committed to maintaining an orderly learning environment for all students and ensuring that all facilities are used in a way that facilitate teaching, learning, and research. Therefore, you should treat your peers and instructor respectfully in discussion. Disagreements are likely to happen. When they do, you are expected to disagree respectfully and to keep your discussion focused on evidence. Discussions in this class are expected to take place solely within the course. Thus, statements made during class should not be quoted on social media unless the individual being quoted has provided their express permission. This applies to the instructor, students, and classroom guests. This policy is meant to protect student privacy and create a safe environment to learn.

Academic integrity

“I commit to uphold the ideals of honor and integrity by refusing to betray the trust bestowed upon me as a member of the Georgia Tech community.” Georgia Tech aims to cultivate a community based on trust, academic integrity, and honor. Students are expected to act according to the highest ethical standards. For information on Georgia Tech’s Academic Honor Code, please visit <http://www.catalog.gatech.edu/policies/honor-code/> or <http://www.catalog.gatech.edu/rules/18/>. Academic integrity is extremely important to me. Any student suspected of cheating or plagiarizing on an exam or assignment will be reported to the Office of Student Integrity, who will investigate the incident and identify the appropriate penalty for violations.

Student-faculty expectations agreement

It is important to strive for an atmosphere of mutual respect, acknowledgement, and responsibility between faculty members and the student body. See <http://www.catalog.gatech.edu/rules/22/> for an articulation of some basic expectations that you can have of me and that I have of you. In the end, simple respect for knowledge, hard work, and cordial interactions will help build the environment we seek. Therefore, I encourage you to remain committed to the ideals of Georgia Tech while in this class.

Athletics absences

If you are a student athlete, you must obtain statements of approved absence from the Office of the Registrar (<https://registrar.gatech.edu/>) one week in advance of your planned absence. I will accommodate your schedule if I have enough time, so please get in touch early.

Accommodations for students with disabilities

If you are a student with learning needs that require special accommodation, contact the Office of Disability Services at (404) 894-2563 or <http://disabilityservices.gatech.edu/> as soon as possible to make an appointment to discuss your special needs and to obtain an accommodations letter. Please also e-mail me as soon as possible to set up a time to discuss your learning needs. Your accommodations cannot be put into place until you discuss them with me.

This syllabus is subject to change

At any time for any reason, I may need to update this syllabus. I will inform you immediately if I make any changes.

Schedule (subject to change)

- **Week 1 (2022/08/22 - 2022/08/26):** Load data, summary stats, simple graphs, making videos.
 - 2022/08/22: Introductions, syllabus, colab intro.
 - 2022/08/24: Colab deep dive, python intro.
 - 2022/08/26: Data types, summary statistics.
- **Week 2 (2022/08/29 - 2022/09/02):** Continuous vs discrete, counts, descriptive stats.
 - 2022/08/29: Colab and summary statistics type reveiw.
 - 2022/08/31: Colab deep dive, python intro, discreet vs continuous.
 - 2022/09/01: HW1 due: Loading data into colab and making first video.
 - 2022/09/02: Summary stats tables: mean, median, and standard deviation.
- **Week 3 (2022/09/05 - 2022/09/09):** Dependent vs independent, light causality.
 - 2022/09/05: LABOR DAY BREAK
 - 2022/09/07: Reveiw mean, median, standard deviation.
 - 2022/09/08: HW2 due: mean, median, standard deviation.
 - 2022/09/09: Dependent vs indendependent variable, deep dive a data frame.
- **Week 4 (2022/09/12 - 2022/09/16):** Probability review.
 - 2022/09/12: Reveiw depdentent vs independent, discuss probability.
 - 2022/09/14: Rules of probability.
 - 2022/09/15: HW3 due: depvar and covariates, light probabiltly exercises.
 - 2022/09/16: Review of probability rules.

- **Week 5 (2022/09/19 - 2022/09/23):** Probability review 2: Conditional Probability.
 - 2022/09/19: Conditional probability, bayes rule.
 - 2022/09/21: More of that, it's tricky.
 - 2022/09/22: HW4 due: some fun stuff with conditional probability and functions.
 - 2022/09/23: Pull together probability review.
- **Week 6 (2022/09/26 - 2022/09/30):** Summation notation.
 - 2022/09/26: We are going to have to spend some time on notation, here is where we dive in to the summation operator.
 - 2022/09/28: Some work on the error term to the sum of squares.
 - 2022/09/28: HW5 due: turn that math into code using the summation operator and functions.
 - 2022/09/30: review of stats proofs and summation notation. We are going to need it quite a bit.
- **Week 7 (2022/10/03 - 2022/10/07):** Correlation and model introduction.
 - 2022/10/03: (progress report) Introduction to model building and correlation of independent and dependent variables.
 - 2022/10/07: No semester is complete with correlation vs causation, this is that lecture.
 - 2022/09/28: HW6 due: correlation vs causation the notebook.
 - 2022/10/09: Digging further into correlation.
- **Week 8 (2022/10/10 - 2022/10/14):** Review and Midterm: conditional prob to correlation.
 - 2022/10/10: Review probability rules.
 - 2022/10/12: Review correlation.
 - 2022/10/14: MIDTERM EXAM - you got this.
- **Week 9 (2022/10/17 - 2022/10/21):** Average Treatment Effect.
 - 2022/10/17: FALL BREAK
 - 2022/10/19: Review of midterm material. Discuss average treatment effect.
 - 2022/10/21: Average treatment effect further discussion.
- **Week 10 (2022/10/24 - 2022/10/28):** Linear regression.
 - 2022/10/24: Error term, sum of squared residual, and best fit line.
 - 2022/10/26: Linear regression summation notation.
 - 2022/10/27: HW7 due: linear regression model in sklearn.
 - 2022/10/28: Further discussion on linear regression model assumptions.
 - 2022/10/29: WITHDRAWAL DEADLINE
- **Week 11 (2022/10/31 - 2022/11/04):** Linear regression 2.

- 2022/10/31: Review of simple regression and model assumptions.
- 2022/11/02: Linear model predictions.
- 2022/11/03: HW8 due: Linear model predictions.
- 2022/11/04: More talk of linear model assumptions.
- **Week 12 (2022/11/07 - 2022/11/11):** Standard error.
 - 2022/11/07: Dispersion, and why it matters. We're really starting to pull everything together now.
 - 2022/11/09: Calculating standard errors.
 - 2022/11/10: HW9 due: linear model dispersion notebook.
 - 2022/11/11: More time on standard errors.
- **Week 13 (2022/11/14 - 2022/11/18):** Test for significance.
 - 2022/11/14: What is significance?
 - 2022/11/16: Calculating significance and expressing in your results.
 - 2022/11/17: HW10 due: tie it all together with model and significance testing.
 - 2022/11/18: Further digressions on testing significance.
- **Week 14 (2022/11/21 - 2022/11/25):** Continuous vs discrete modeling.
 - 2022/11/21: Bonus material: continuous vs discrete model, and why you would care.
 - 2022/11/23: THANKSGIVING BREAK
 - 2022/11/25: THANKSGIVING BREAK
- **Week 15 (2022/11/28 - 2022/12/02):** Bring it all together.
 - 2022/11/28: Review week (nothing new): Working together we'll take a continuous and discrete problem: tell summary stats
 - 2022/11/30: Continue above with build a linear model, work out the math, code it up
 - 2022/12/02: Continue the above show model fit and significance tests.
- **Week 16 (2022/12/05 - 2022/12/09):** Final Exam
 - 2022/12/28: Review
 - 2022/12/30: READING DAY
 - 2022/12/02: Final exam