STA 4210

Instructor

Rohit Patra

Office: Griffin Floyd 221

E-mail: rohitpatra@ufl.du

Office: Griffin Floyd 221

E-mail: rohitpatra@ufl.edu

Office hours:

Wednesday 4:00pm to 6:00pm

Zoom link for office hours: https://ufl.zoom.us/j/892339318

Synchronous interactive session**:**

Thursday from 12:50pm to 1:40pm

Zoom Link: https://ufl.zoom.us/j/892339318

Teaching Assistant

Cheng Zeng

Office: TBD

E-mail: czeng1@ufl.edu

Office hours:

1. Tuesday 1:00pm-2:00pm
2. Friday-- 2:00pm--4:00pm

Zoom link for Cheng's office hours:

https://ufl.zoom.us/j/96750005981

Course Prerequisites

STA 3024 or STA 3032 or STA 4321 or MAS 3114 or MAS 4105.

Course Lectures

Lectures covering the weekly material will be pre-recorded and posted on the course website at the beginning of each week. In addition to the recorded lectures, we will have one synchronous session on Thursdays from 12:50 pm to 1:40pm. This session will vary from week to week, but will typically involve more student interaction, implementing approaches in R, or working on problems that I assign. It is intended to be complementary to the recorded lectures, which cover the bulk of the course material.

Course Notes/Material

Notes for the week will be posted at the beginning of each week on the course website. These should contain nearly all of the material that we cover in class, however, I will present some additional material in the lectures that is not posted on the course website.

Software

We will be using the R software language throughout. R is free and should be easy to download on your personal computer. I also highly recommend running R through RStudio, though it is not a requirement. If you have any problems downloading R or RStudio, feel free to talk to myself or the TA. If you do not have access to a computer, please reach out to me via email.

Required Text

Kutner, Nachtsheim, Neter, and Li, Applied Linear Statistical Models, 5th ed., Volume I (Chapters 1–14, Appendix A) The Data Sets used in the class are available online in canvas tab for files (subfolder data)

Course Description

The course is primarily on the linear regression model, for which the main techniques are rooted in the method of least squares. Procedures are motivated by applications. Mathematical results are stated and explained, and occasionally derived. A course in mathematical statistics is helpful but not strictly required. The focus is on carrying out and explaining the methods. After some review of basic statistics, the course proceeds systematically through the simple regression model, the matrix formulation of this model, the multiple regression model, and a number of related tools such as model diagnostic measures, collinearity statistics, and variable selection procedures. Computations will be carried out in the R statistical programming language.

Homework

There will be a homework assignment every week (except the weeks after and during exams) and it will be due via Canvas submission at the end of that week (Sunday 11:59pm). The homeworks can be found canvas tab for files (subfolder Homework). The homeworks and their due dates are also listed on the weekly breakdown on the canvas homepage. The answers will have partial credit and each homework is worth the same number of points. Even though, the homeoworks will focus on the recent topics covered in the class, they are cumulative. No makeup homeworks will be offered. However, homeworks with the lowest **two** grades will be **dropped**. If you have tech troubles, please contact UF helpdesk or call them at 352-392-4357 (they are open 24x7).

Exams

You will have 2 synchronous exam and one take home project. On each exam, there will be some short-answer questions, both multiple-choice and fill-in-the-blank types. Other questions will require a written response; for these questions, it is not sufficient to simply give a numerical or one-word answer. There will be short calculations required; you need a calculator for the exams. On each exam, there will be at least one question which describes a regression scenario, gives R output from the analysis, and asks you to interpret the output. The exams will be proctored via Honorlock.

Project

Students will be expected to complete a written project at the end of the semester and present their findings to the class. The grade for the final project will consist of three main components: statistical modeling decisions and code and a written report. The first component is based on whether the student used appropriate statistical tools for the setting and objectives, whether the coding done was efficiently and correctly, and whether the conclusions are consistent with their analysis. The written report will state the objectives of the study, describe data collection, describe the statistical model used, explain any assumptions required by the analysis, and provide conclusions for the main study questions. There will also be some questions similar to Exam 1 and 2 that will need to be answered.

Grades

After each exam I will send you an email with your current grade. The grade will take into account all the exams, quizzes, and homeworks up to that point in time. It should be only taken as a rough guide on your current standing in the class. The final grade is based on the following composite score:

Homework	25%
Exam 1 (October 5)	20%
Exam 2 (November 5)	25%
Project	30%

Letter Grade Assignment

Grades will be assigned as follows: 90-100, A; 87-89.9, A-; 84-86.9, B+; 80-83.9, B; 77- 79.9, B-; 74-76.9, C+; 70-73.9, C; 67-69.9, C-; 64-66.9, D+; 60-63.9, D; 55-59.9, D-; 0- 55, F

The numeric scores will be rounded up to the nearest tenth.

Feedback

I have created an assignment on Canvas. There you can submit anonymous feedback. You should be able to submit **as many feedbacks** as you want throughout the semester.

Make up Policy

Requirements for class attendance and make-up exams, assignments, and other work in this course as well as policies regarding absences, religious holidays, illness and student athletes are consistent with UF Attendance Policies

(https://catalog.ufl.edu/UGRD/academic-regulations/attendance-policies/)

Dropping and Withdraw

For late course drops and course withdrawals please visit https://catalog.ufl.edu/UGRD/academic-regulations/dropping-courses-withdrawals/

Incomplete

An incomplete grade may be assigned at the discretion of the instructor as an interim grade for a course in which the student has completed a major portion of the course with a passing grade, been unable to complete course requirements before the end of the term because of extenuating circumstances, and obtained agreement from the instructor and arranged for resolution of the incomplete grade in the next term. Instructors are not required to assign incomplete grades. For complete details please visit CLAS incomplete grade policies and forms.

(https://www.advising.ufl.edu/academicinfo/clas-policiesprocedures/incomplete-grades/)

Accommodating Students with Disabilities

Students requesting accommodation for disabilities must first register with the Dean of Students Office. The Dean of Students will provide documentation to the students who must then provide this documentation to the instructor when requesting information. You must submit this documentation prior to submitting any assignments for which you are requesting accommodation.

Academic Misconduct

Students are held accountable to the UF Honor Code.

(https://sccr.dso.ufl.edu/process/student-conduct-code/

Evaluations

Students are expected to provide feedback on the quality of instruction in this course by completing online evaluations at https://evaluations.ufl.edu. Evaluations are typically open during the last two or three weeks of the semester, but students will be given specific times when they are open. Summary results of these assessments are available to students at

https://evaluations.ufl.edu/results/.

Counseling and Wellness Center

Contact information for the Counseling and Wellness Center: http://www.counseling.ufl.edu/cwc/Default.aspx, 392-1575; and the University Police Department: 392-1111 or 9-1-1 for emergencies.

Additional resources

Any additional resources including academic support or information technology can be found at https://www.ufl.edu/about/offices-services/

Privacy statement regarding online lectures

Our class sessions may be audio-visually recorded for students in the class to refer back and for enrolled students who are unable to attend live. Students who participate with their camera engaged or utilize a profile image are agreeing to have their video or image recorded. If you are unwilling to consent to have your profile or video image recorded, be sure to keep your camera off and do not use a profile image. Likewise, students who un-mute during class and participate verbally are agreeing to have their voices recorded.

If you are not willing to consent to have your voice recorded during class, you will need to keep your mute button activated and communicate exclusively using the "chat" feature, which allows students to type questions and comments live. The chat will not be recorded or shared.

As in all courses, unauthorized recording and unauthorized sharing of recorded materials is prohibited.

Privacy statement regarding use of Canvas

Canvas accessibility and privacy statements can be found in the following link. https://docs.google.com/document/d/1EQ3PlY5PwyhLEeVCXr2ZR8s27V9Jl8e2-zKTU7NVGX0/copy

Course objectives

Here is a list of course objectives.

1. Be able to use **R**
2. With two quantitative variables, be able to carry out correlation and simple regression analyses, and to correctly interpret such analyses.
3. Be able to carry out and interpret inference procedures for simple linear regression.
4. Know the simple and multiple linear regression models, and be able to state and explain the standard methods of estimation for these models.
5. Know the multiple linear regression model in its matrix form, including all the common variations on this model (e.g., continuous predictors, categorical predictors, square and interaction terms).
6. Know what is the general F-test; given a particular multiple regression scenario, be able to apply the general F test.
7. Be able to carry out and interpret inference procedures for the various types of multiple regression model, including the quadratic regression model.
8. Know what are the purposes of diagnostic methods in simple and multiple regression; be able to carry out several common diagnostic procedures and interpret them.
9. Know several measures of model performance (R^2, adjusted R^2, and Akaike's Information Criterion), how to compute and interpret them for a multiple regression model.
10. Know what multicollinearity is, why it is an issue in multiple regression, and how to analyze and deal with its presence.