STAT 5200

Applied Statistics I

Lectures: MWF 2:30pm–3:20pm, 74 SH

Labs: Friday 8:30am–9:20am, 41 SH

Fall 2024

Instructor

Dr. Andrew M. Thomas

Department of Statistics and Actuarial Science

Office: 205 SH

andrew-thomas@uiowa.edu

DEO

Prof. Kung-Sik Chan

Office: 241 SH

kung-sik-chan@uiowa.edu

Prerequisites

STAT 3101 (Introduction to Mathematical Statistics II) or equivalent and knowledge of matrix algebra.

Co-requisites

A co-requisite indicates a UI course that you must have satisfactorily completed or be enrolled in at the same time as the listed course.

STAT 4100 (Mathematical Statistics I) or STAT 5100 (Statistical Inference I).

Textbook

Introduction to Regression Modeling by Bovas Abraham and Johannes Ledolter, (ISBN: 0534420753)

Software

We will use the **R** language for our analyses in this class. **R** is open-source statistical software and very useful for analyses involving linear models. It is available for free at: https://www.r-project.org/. It is also installed in the lab. You will also learn how to use **R** within a Linux operating system.

RStudio is a nice *integrated development environment* for **R** (https://posit.co/download/rstudio-desktop/), which will likely make your experience with **R** (and doing the homework) much easier.

Other software:
- **RMarkdown:** https://rmarkdown.rstudio.com/lesson-1.html
- **Tidyverse:** https://r4ds.had.co.nz/

Learning Objectives

We will cover chapters 1 through 7 of the textbook in the first 11 weeks or so. We will then cover chapters 9 and 11—especially logistic regression—and additional textbook topics if time allows. The following topics will be covered:

- Simple and multiple linear regression
- Linear model formulation, checking, and selection
- Linear model estimation, explanation, and inference
- Matrix algebra and random vectors for linear regression
- Missing data, transformations and categorical predictors (e.g. ANOVA)
- Simple and multiple logistic regression
- Interpretation and presentation of linear & logistic regression analysis
- Model selection procedures for linear & logistic regression
- **R** for linear and logistic regression
- Producing data analysis/visualization reports using **R** and **RMarkdown**

The focus of the course will be in applications, as the title suggests. Computing, interpreting, and communicating these results will be the primary objective of the course. However, we will cover some results and calculations with mathematical rigor as this is a graduate-level course.

Course Website

Lecture notes, lab handouts, homework assignments and solutions, review materials, grades, etc...

Office Hours

Office hours are a great time and place to discuss any questions that you have about the homework, exams, or other questions about the course. I encourage everyone to stop by at some point in the semester! Office hours will be held at the following times in my office **205 SH** or Zoom (unless there is an official university holiday, in which case *there will be no office hours*):

Monday: 1pm–2pm (205 SH)
Wednesday: 1pm–2pm (205 SH)
Thursday: 10am–11am (Zoom)

I am available to meet in-person or on Zoom (by appointment) if you are not able to drop in at these times.

Grading

Homework	10%	Approximately 8–10 homeworks will be assigned.
Quizzes	15%	2 or 3 quizzes will given in class.
Midterm	25%	
Project	15%	
Final	35%	

Exams will be curved as necessary. Final grades will be awarded based on the following ranges:

A		B		C		D	
A+	>97%	B+	>87%	C+	>77%	D+	>67%
A	>93.5%	B	>83.5%	C	>73.5%	D	>63.5%
A-	>90.5%	B-	>80.5%	C-	>70.5%	D-	>60.5%

Any final grade achieving less than 60.5% (inclusive) will be awarded an F.

Homework and Labs

Homework will be assigned (mostly) weekly and typically due the week after they are assigned. **The lowest <u>two</u> homework scores will be dropped**.
Homework will be relatively brief and consist of short mathematical proofs, computations (for which you will often be able to use **R**), and data analysis. Usually your lab assignment will be turned in along with your homework as well.
Homework will be submitted online on ICON.

A recommended way to construct homework will be to use **RMarkdown** to produce a word document or pdf. However, use of **RMarkdown** is not required and you need not even type up your homework as long you include the required code, output, and visualizations requested. It is expected that you include explanations in words for calculations and data analysis that you do in your homework, even if *the problem does not specifically request it.*

Exams

There will be one in-class midterm, scheduled tentatively for **Wednesday, October 23rd, 2024**. A single two-sided, 8.5" x 11" cheat sheet (handwritten) will be allowed for all exams.

The final exam will be comprehensive and, in principle, cover all of the information that we learned over the course of the semester.

Make-up policy: If you have an appropriate, documented reason why you must miss an exam, please reach out and provide me with such documentation and reason for your absence so that you may take it at a different time. Students who miss an exam without an appropriate, documented reason will receive a 0 on said exam.

For information on rescheduling the final because of final exam conflicts, please see the university policy: https://registrar.uiowa.edu/makeup-examination-policies. Please ensure that you submit this request by Tuesday, October 1st, 2024.

Quizzes

There will be 2–3 short quizzes to help you keep up with the material covered in class. The quizzes will be closed book and I will announce the date and content around one week before each quiz. The make-up policy for quizzes is same as that of exams.

Project

For the project you can either identify an appropriate dataset for which multiple linear regression/logistic regression is applicable or create an interesting simulation study (for example severe multicollinearity, small sample size relative to the number of predictors, nonnormal errors for linear regression etc...) and generate fake data under it. The goal would be to analyze the real dataset using some of the statistical methods taught in class or compare the performance of different methods using simulation studies. You may include methods that have not been taught in class, *but then you would be expected to have a good understanding of those methods*.

Some good resources for datasets are the DASL Library (The Data and Story Library) at http://lib.stat.cmu.edu/DASL/, the UCI Machine Learning Repository at http://archive.ics.uci.edu/ml/, and the **MASS** library in **R**. Please do not analyze data that have already been used in lectures, labs, or homework assignments. For completion of your project, you will need to submit

1. A project proposal
2. Slides for an oral presentation

as well as complete a final oral presentation.

The final oral presentation should address the following:
- Objective of the project
- Brief description of the data (source, variable key etc. or simulation set up)
- Statistical methods used for data analysis/simulation study
- Summary of results (presented as figures or tables when applicable)
- Discussion of findings and your interpretation, any shortcomings, and/or future work
- Appendix for **R** code

Once you have identified the dataset or simulation topic, please submit a brief project proposal (maximum one-page description—but can have extra pages for visualizations, figures, etc...) by **Monday, November 11th, 2024.** I am happy to talk with you about any questions regarding the project. The slides will be due on **Monday, December 9th, 2024,** and the presentations (approximately 10-15 minutes each) will most likely be on December 4th and December 6th, 2024.

Attendances, Class Participation, and Absences

Attendance is not necessary, and I plan on uploading lecture notes online. However, I should say that you will be much more well-equipped to succeed if you attend class! Students are encouraged to ask questions and come to office hours, as well as discuss homework problems with other students. Active participation in lecture is also encouraged. If there is something you don't understand, let me know—there's bound to be someone else who feels the same way.

University regulations require that students be allowed to make up examinations that have been missed due to *illness, religious holy days, military service obligations (including service-related medical appointments), or other unavoidable circumstances or University-sponsored activities*. Students with UI-authorized activities must discuss their absences with the instructor as soon as possible. Religious obligations must be communicated within the first three weeks of classes. Students are requested to use the **absence form** in ICON under Student Tools.

Academic Honesty and Misconduct

All students in CLAS courses are expected to abide by the CLAS Code of Academic Honesty. If you have any doubts about what constitutes a violation of the CLAS Code of Academic Honesty, or any other issue related to academic integrity, please contact me.

Student Complaints

Students with a complaint about a grade or a related matter should first discuss the situation with the instructor (me) and finally with the Director or Chair of the school, department, or program offering the course.

Graduate students should contact the CLAS <u>Associate Dean for Graduate Education and Outreach and Engagement</u> when additional support is needed.

Drop Deadline for this Course

You may drop an individual course before the deadline; after this deadline you will need collegiate approval. You can look up the <u>drop deadline for this course</u> here. When you drop a course, a "W" will appear on your transcript. The mark of "W" is a neutral mark that does not affect your GPA. Directions for adding or dropping a course and other registration changes can be found on the <u>Registrar's website</u>. Graduate students should adhere to the <u>academic deadlines</u> and policies set by the Graduate College.

Date and Time of the Final Exam

The <u>final examination date and time</u> will be announced by the Registrar generally by the fifth week of classes and it will be announced on the course ICON site once it is known. **Do not plan your end of the semester travel plans until the final exam schedule is made public. It is your responsibility to know the date, time, and place of the final exam.** According to Registrar's final exam policy, students **have a maximum of two weeks after the announced final exam schedule** to request a change if an exam conflict exists or if a student has more than two exams in one day (see the <u>policy</u> here).

Communication: UI Email

Students are responsible for all official correspondences sent to their UI email address (uiowa.edu) and must use this address for any communication with instructors or staff in the UI community. For the privacy and the protection of student records, UI faculty and staff can only correspond with UI email addresses.

ChatGPT and AI Policy

Use of AI for the completion of assignments is *discouraged* but not strictly prohibited. Where AI-content generators such as ChatGPT are used, I expect there to be attribution made to said content. For example, even a brief sentence that some code was generated by ChatGPT and modified by the student can suffice.

Without proper attribution, this constitutes academic dishonesty. If you are unsure about whether the use of an AI tool may constitute academic dishonesty, you are welcome to contact me, though my advice is to simply not use it.

Free Speech and Expression

The University of Iowa supports and upholds the First Amendment protection of freedom of speech and the principles of academic and artistic freedom. We are committed to open inquiry, vigorous debate, and creative expression inside and outside of the classroom. Visit the Free Speech at Iowa website for more information on the university's policies on free speech and academic freedom.

Accommodations for Students with Disabilities

The University is committed to providing an educational experience that is accessible to all. If a student has a diagnosed disability or other disabling condition that may impact the student's ability to complete the course requirements as stated in the syllabus, the student may seek accommodations through Student Disability Services (SDS). SDS is responsible for making Letters of Accommodation (LOA) available. **The student must provide an LOA to the instructor as early in the semester as possible, but requests not made at least two weeks prior to the scheduled activity for which an accommodation is sought may not be accommodated.** The LOA will specify what reasonable course accommodations the student is eligible for and those the instructor should provide. Additional information can be found on the SDS website.

Absences for Religious Holy Days

The university is prepared to make reasonable accommodations for students whose religious holy days coincide with their classroom assignments, test schedules, and classroom attendance expectations. Students must notify their instructors in writing of any such Religious Holy Day conflicts or absences within the first few days of the semester or session, and no later than the third week of the semester. If the conflict or absence will occur within the first three weeks of the semester, the student should notify the instructor as soon as possible. See Policy Manual 8.2 Absences for Religious Holy Days for additional information.

Non-discrimination statement

The University of Iowa prohibits discrimination in employment, educational programs, and activities on the basis of race, creed, color, religion, national origin, age, sex, pregnancy (including childbirth and related conditions), disability, genetic information, status as a U.S. veteran, service in the U.S. military, sexual orientation, gender identity, or associational preferences. The university also affirms its commitment to providing equal opportunities and equal access to university facilities. For additional information on nondiscrimination policies, contact the Senior Director, Office of Civil Rights Compliance, the University of Iowa, 202 Jessup Hall, Iowa City, IA 52242-1316, 319-335-0705, daod-ocrc@uiowa.edu. Although not required, students have the option to share their pronouns and chosen/preferred names in class and through MyUI. Instructors and advisors can find information about a student's chosen/preferred name in MyUI.

Mental Health Resources and Student Support

Students are encouraged to be mindful of their mental health and seek help as a preventive measure or if feeling overwhelmed and/or struggling to meet course expectations. Students are encouraged to talk to their instructor for assistance with specific class-related concerns. For additional support and counseling, students are encouraged to contact University Counseling Service (UCS). Information about UCS, including resources and how to schedule an appointment, can be found at counseling.uiowa.edu. Find out more about UI mental health services at mentalhealth.uiowa.edu.

Student Care and Assistance provides assistance to University of Iowa students who are experiencing a variety of crisis and emergency situations, including but not limited to medical issues, family emergencies, unexpected challenges, and sourcing basic needs such as food and shelter. More information on the resources related to basic needs can be found at basicneeds.uiowa.edu/resources/. Students are encouraged to contact Student Care & Assistance in the Office of the Dean of Students (Room 135 IMU, dos-assistance@uiowa.edu, or 319-335-1162) for support and assistance with resources.

Additional Links

[Classroom Expectations](#)

[Class Recordings](#)

[Conflict Resolution](#)

[Sexual Harassment/Misconduct and Supportive Measures](#)