

STAT 3302

Statistical Modeling for Discovery II

- Spring 2024
- 3 credit hours
- In-person lectures
 - Time: Monday, Wednesday, Fridays 1:50pm – 2:45pm
 - Location: Pomerene Hall (PO) 250
 - Lectures will not be recorded

Course overview

Instructor

- Instructor: Vincent Q. Vu
- E-mail: vqv@stat.osu.edu
- Office hours
 - Time: Wednesdays, 4:00pm – 5:00pm
 - Location: Cockins Hall (CH) 428B

TA

- TA: David Angeles
- E-mail: angeles.6@buckeyemail.osu.edu

Description

This course continues to investigate statistical models for data analysis and discovery in big-data settings. The regression methods developed in STAT 3301 are extended to data settings with binary and multi-category outcomes. An introduction to some of the most commonly used

statistical methods for exploring and analyzing multivariate data is provided. Interpretation and communication of the results of analyses is emphasized.

Expected learning outcomes

By the end of this course, students should successfully be able to:

1. Build, fit and interpret statistical models for binary outcomes
2. Understand the difference between nominal and ordinal outcomes and build regression models that are appropriate for each
3. Recognize the types of questions that can be answered by regression models for multi-category data and structure models to answer those questions
4. Comprehend the statistical principles that underlie basic methods of multivariate data analysis

Prerequisites

STAT 3301 (Statistical Modeling for Discovery I) and MATH 2568 (Linear Algebra); or permission of instructor.

Course materials

Textbooks

The following two textbooks are required for this course:

- A. J. Dobson and A. G. Barnett (2018), **An Introduction to Generalized Linear Models**, Fourth Edition, CRC Press. *Available online at <https://library.ohio-state.edu/record=b8615141~S7>*
- A. C. Rencher and W. F. Christensen (2012), **Methods of Multivariate Analysis**, Third Edition, Wiley. *Available online at <https://library.ohio-state.edu/record=b7149844~S7>*

I will highlight other useful references as the course progresses.

Software

You are expected have to install on your personal computer (or have access to) up to date versions of the following software:

- **R** <https://cloud.r-project.org/> (The latest version as of January 8, 2024 is 4.3.2)
- **RStudio** Desktop IDE <https://posit.co/download/rstudio-desktop/> (The latest version as of January 8, 2024 is 2023.12.0+369)

In addition you will need to install the following R packages: `dobson`, `tidyverse`, `broom`, and `rmarkdown`. In R, use the command `install.packages('PACKAGENAME')` to install a package. For example, use `install.packages('dobson')` to install `dobson`.

Assignments, grading, and attendance

Assignments

Assignment category	Percentage
Homework	20%
Midterm Exam 1 (February 16, 2024)	20%
Midterm Exam 2 (March 22, 2024)	20%
Project	10%
Final Exam (April 26, 2024; 4:00pm – 5:45pm)	30%

Homework

Homework will be assigned on an approximately biweekly basis and made available on Carmen. Your answers must be submitted on Carmen as a **single PDF**. Some homework problems will require computing and data analysis in R. You produced from an R Markdown file template provided with the assignment. Homework that that do not require R may be handwritten (electronically, or on paper and scanned) but they must be combined together with the rest of your submission into a **single PDF**. You can use Adobe Acrobat (see <https://it.osu.edu/adobe>) to combine PDF files before submitting to Carmen.

I encourage you to for study groups to discuss and work on homework problems in groups. However, your solutions must be written independently and in your submission, include the names of the people with whom you've discussed the homework. Do not copy, refer to, or look at solutions from a previous offering or from solutions posted online, official or unofficial. Keep in mind the university policies on academic misconduct.

Exams

There will be two in-class midterm exams and a cumulative final exam. All exams will be closed book/closed notes. You will not need a calculator. Computers, smartphones, and similar devices are not allowed to be used. If there is an exceptional circumstance, e.g. medical or family emergency, then please contact me as soon as possible before the exam date to make potential arrangements.

Project

There will be a project that you will complete in groups of 4 students. The group leader and members will be assigned automatically. The project will consist of finding a dataset, formulating questions that can be answered with the data, and performing an appropriate analysis to answer the questions posed.

Proposals for the project will be due just before spring break. The complete project will be due before the final exam. Further details, including deadlines, will be given as the semester progresses and announced on Carmen. There are limited options to change your group, under some circumstances if you want to. Email me so we can discuss and find a way.

Late assignments

If you absolutely need turn in an assignment late and have a valid excuse, please see me for the necessary arrangements. However, you must notify me in advance in such a situation. Exceptions to this policy will be permitted only in extreme situations such as serious injury immediately prior to an assignment being due or severe illness requiring hospitalization.

Grading scale

Points	Grade
93–100	A
90–92	A-
87–89	B+
83–86	B
80–82	B-
77–79	C+
73–76	C
70–72	C-
67–69	D+
60–66	D

Points	Grade
00-60	E

Class attendance

You are expected to attend all lectures.

Academic policies

Academic integrity policy

It is the responsibility of the Committee on Academic Misconduct to investigate or establish procedures for the investigation of all reported cases of student academic misconduct. The term “academic misconduct” includes all forms of student academic misconduct wherever committed; illustrated by, but not limited to, cases of plagiarism and dishonest practices in connection with examinations. Instructors shall report all instances of alleged academic misconduct to the committee (Faculty Rule 3335-5-487). For additional information, see the Code of Student Conduct <http://studentlife.osu.edu/csc/>.

Accessibility accommodations for students with disabilities

Requesting accommodations

The university strives to maintain a healthy and accessible environment to support student learning in and out of the classroom. If you anticipate or experience academic barriers based on your disability (including mental health, chronic, or temporary medical conditions), please let me know immediately so that we can privately discuss options. To establish reasonable accommodations, I may request that you register with Student Life Disability Services. After registration, make arrangements with me as soon as possible to discuss your accommodations so that they may be implemented in a timely fashion.

If you are isolating while waiting for a COVID-19 test result, please let me know immediately. Those testing positive for COVID-19 should refer to the [Safe and Healthy Buckeyes site](#) for resources. Beyond five days of the required COVID-19 isolation period, I may rely on Student Life Disability Services to establish further reasonable accommodations. You can connect with them at slds@osu.edu; 614-292-3307; or slds.osu.edu.

Religious accommodations

It is Ohio State's policy to reasonably accommodate the sincerely held religious beliefs and practices of all students. The policy permits a student to be absent for up to three days each academic semester for reasons of faith or religious or spiritual belief.

Students planning to use religious beliefs or practices accommodations for course requirements must inform the instructor in writing no later than 14 days after the course begins. The instructor is then responsible for scheduling an alternative time and date for the course requirement, which may be before or after the original time and date of the course requirement. These alternative accommodations will remain confidential. It is the student's responsibility to ensure that all course assignments are completed.

Your mental health

As a student you may experience a range of issues that can cause barriers to learning, such as strained relationships, increased anxiety, alcohol/drug problems, feeling down, difficulty concentrating and/or lack of motivation. These mental health concerns or stressful events may lead to diminished academic performance or reduce a student's ability to participate in daily activities. The Ohio State University offers services to assist you with addressing these and other concerns you may be experiencing. If you or someone you know are suffering from any of the aforementioned conditions, you can learn more about the broad range of confidential mental health services available on campus via the Office of Student Life's Counseling and Consultation Service (CCS) by visiting ccs.osu.edu or calling 614-292-5766. CCS is located on the 4th Floor of the Younkin Success Center and 10th Floor of Lincoln Tower. You can reach an on call counselor when CCS is closed at 614-292-5766 and 24 hour emergency help is also available 24/7 by dialing 988 to reach the Suicide and Crisis Lifeline. If you wish to add language on sexual misconduct/relationship violence, the university recommends using the following:

Copyright for instructional materials

The materials used in connection with this course may be subject to copyright protection and are only for the use of students officially enrolled in the course for the educational purposes associated with the course. Copyright law must be considered before copying, retaining, or disseminating materials outside of the course.

Statement on Title IX

Title IX makes it clear that violence and harassment based on sex and gender are Civil Rights offenses subject to the same kinds of accountability and the same kinds of support applied to

offenses against other protected categories (e.g., race). If you or someone you know has been sexually harassed or assaulted, you may find the appropriate resources at <https://titleix.osu.edu> or by contacting the Ohio State Title IX Coordinator at titleix@osu.edu.

Course schedule

This schedule is approximate. Refer to the Carmen course page for up-to-date assignment due dates.

Week	Dates	Topics	Assignment
1	1/8–1/12	Introduction; Review of Binomial Model and Maximum Likelihood Estimation	
	1/15	No class (Martin Luther King, Jr. Day)	
2	1/17–1/19	Simple Logistic Regression: Model Formulation	
3	1/22–1/26	Estimation of Parameters; Multivariable Logistic Regression	HW1
4	1/29–2/2	Model Evaluation and Diagnostics in Logistic Regression	
5	2/5–2/9	Model Building in Logistic Regression	HW2
6	2/12–2/14	Binomial Regression	Exam 1
7	2/19–2/23	Poisson Regression	HW3
8	2/26–3/1	Models for Ordinal and Multicategory Data	
9	3/4–3/8	Introduction to Multivariate Data	HW4, Proposal
	3/11–3/15	No class (Spring break)	
10	3/18–3/22	Multivariate Numerical Summaries	Exam 2
11	3/25–3/29	The Multivariate Normal Distribution	HW5
12	4/1–4/5	More on the Multivariate Normal Distribution	
13	4/8–4/12	Principal Components Analysis	HW6
14	4/15–4/19	Linear Discriminant Analysis	Project
15	4/22	Catch-Up, Review, and Further Directions	
	4/26	Final Exam (4:00pm – 5:45pm)	