



THE OHIO STATE UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

SYLLABUS: STAT 3201

INTRODUCTION TO PROBABILITY FOR DATA ANALYTICS

AUTUMN 2023

Course overview

Instructor

Instructor: Subhadeep Paul

Email address: paul.963@osu.edu

Lectures: Mondays, Wednesdays and Fridays, 3:00 -- 3:55 PM at Evans Lab 2004.

Office hours: Cockins Hall 231, Wednesdays 4:15 – 5:15 PM.

Teaching Assistant and Tutoring

Jiwon Hong (hong.1069@osu.edu)

Tutoring and office hours will be coordinated through Data Analytics Learning Center (DALC) (<https://data-analytics.osu.edu/dalc>) and

Mathematics and Statistics Learning Center (MSLC) (<https://mslc.osu.edu/>). More details on hours coming soon.

Course description

An introduction to probability and its role in statistical methods for data analytics. Equal emphasis is placed on analytical and simulation-based methods for quantifying uncertainty. Approaches to assessing the accuracy of simulation methods are discussed. Students should have some prior knowledge of basic programming. Applications of probability and sampling to big-data settings are discussed.

Topics we will cover:

1. Introduction to Probability

- a. Introduction to probability and counting methods
 - b. Conditional probability, independence, Bayes' Theorem
 - c. Discrete random variables and probability distributions
 - d. Simulation and Monte Carlo estimation
2. Discrete probability distributions
 - a. Bernoulli and Binomial
 - b. Negative Binomial, Geometric, Hypergeometric, and Poisson
3. Continuous probability distributions
 - a. Density and distribution functions
 - b. Uniform, Normal, Gamma, Exponential, Beta distributions
4. Functions of random variables
5. Sampling distributions and Central limit theorem
6. Marginal and conditional distributions

Prerequisite or corequisite:

Prerequisites: Math 1152 or 1161.xx or 1172 or 1181 or equivalent, and CSE Placement Level A or equivalent.

Course learning outcomes

Upon successful completion of the course, students will be able to

1. Quantify uncertainty about events using mathematical descriptions of probability
2. Quantify uncertainty about events using simulation methods
3. Assess the quality and accuracy of simulation-based descriptions of uncertainty
4. Update a description of uncertainty based on new information
5. Identify appropriate probability models for experiments/data and summarize expected outcomes from such models
6. Use correlation and conditional expectation to describe the relationship between two random variables.
7. Quantify uncertainty about summary statistics for large data sets

Course materials

Required

Mathematical Statistics with Applications (7th edition) by Wackerly, Mendenhall and Sheaffer.

This book is available to you (unless you have opted out) through CarmenBooks.

Course material will be supplemented with the freely available books:

R programming for Data Science by Roger D Peng

<https://bookdown.org/rdpeng/rprogdatascience/Links to an external site.>

R for Data Science by Garrett Golemund and Hadley Wickham

<https://r4ds.had.co.nz/>

I will also post incomplete lecture notes in Carmen before each class, which will be interactively completed in class through annotation. You are **not required** to study the incomplete lecture notes beforehand.

Course technology

For help with your password, university e-mail, Carmen, or any other technology issues, questions, or requests, contact the OSU IT Service Desk. Standard support hours are available at <https://ocio.osu.edu/help/hours>, and support for urgent issues is available 24x7.

- **Self-Service and Chat support:** <http://ocio.osu.edu/selfservice>
- **Phone:** 614-688-HELP (4357)
- **Email:** 8help@osu.edu
- **TDD:** 614-688-8743

Baseline technical skills necessary for this course

- Basic computer and web-browsing skills
- Navigating Carmen

Technology skills necessary for this specific course

- Zoom
- Scanning a written solution and submitting a PDF file through Carmen.

Necessary equipment

- Computer: current Mac (OS X) or PC (Windows 10+) with high-speed internet connection
- Webcam: built-in or external webcam, fully installed
- Microphone: built-in laptop or tablet mic or external microphone

Necessary software

- This class requires you to use the statistical software package called R (The R Project for Statistical Computing; <http://www.r-project.org/>). This software package is available as Free Software.
 - You can download R for Windows, Mac, and Linux, from the CRAN archive at <https://cran.r-project.org>.
 - An in-depth introduction to R is available at <http://cran.r-project.org/doc/manuals/R-intro.pdf>
 - Hands-on tutorials are available in the Swirl system, which you can learn about at <http://swirlstats.com/>. In particular, “R Programming: The basics of programming in R” is an appropriate first tutorial for students who have never used R.
- An easier to use interface to R is available in the software package RStudio. This package is available for Windows, Mac, and Linux and can be downloaded for free from <http://rstudio.org>. **Note that RStudio requires R to be installed.**
- [Microsoft Office 365 ProPlus](#) All Ohio State students are now eligible for free Microsoft Office 365 ProPlus through Microsoft’s Student Advantage program. Each student can install Office on five PCs or Macs, five tablets (Windows, iPad® and Android™) and five phones.
 - Students are able to access Word, Excel, PowerPoint, Outlook and other programs, depending on platform. Users will also receive 1 TB of OneDrive for Business storage.
 - Office 365 is installed within your BuckeyeMail account. Full instructions for downloading and installation can be found <https://ocio.osu.edu/kb04733>.

Course delivery

The course will be **primarily delivered in-person**. Occasionally some lectures will be presented online either live over zoom during the lecture time or recorded lectures will be posted on the class website in carmen (**up to 25% of the total lecture hours**). Please visit the following link for more information on OSU’s in-person teaching modality.
<https://teaching.resources.osu.edu/glossary/term/modality-mode>

Each week we will cover approximately 165 minutes of content in total. You will be responsible for attending in-person and any live online/recorded videos as well as studying the material that is assigned. All incomplete and annotated lecture slides will be posted on the class website in Carmen. In addition to the lectures, assignments will also be posted on the class website. You will be given ample time to complete the assignments.

The instructor will hold weekly office hours in his office at Cockins Hall 231. The times are given above.

Grading and faculty response

Grades

Assignment or category	Percentage
Homework	40
Exam 1	20
Exam 2	20
Final Exam	20
Total	100

Grading scale

93–100: A
 90–92.9: A-
 87–89.9: B+
 83–86.9: B
 80–82.9: B-
 77–79.9: C+
 73–76.9: C
 70–72.9: C-
 67–69.9: D+
 60–66.9: D
 Below 60: E

Assignment information

Homework: Homework will be assigned regularly (about 5-6 assignments over the semester). It will consist of numerical problems motivated by data analytics applications and small computer programming/simulation problems. You are encouraged to work together on the problems, but each student must hand in their own work, written in their own words. Do not copy any part of another student's homework including computer output. Use of homework solutions distributed in previous offerings of the course or available on the web constitutes academic misconduct and will be handled according to university rules. **All homework must be submitted online as a PDF file through the class website (carmen).** For written problems you may take a picture of or scan your written solutions and upload them as a PDF file. For R based problems please submit typed solutions. Please be sure that the questions are clearly labeled, all supporting work (including computer code) can be easily identified, and that all figures/tables are referenced and interpreted in the text.

Please note late submission of assignments will be penalized at a rate of 10% of points received in the assignment per day they are late. The late penalty may be waived if prior exception has been sought. if you are unable to complete an assignment on time and need extra time, please get in touch with me as soon as possible so we can discuss your situation.

Exams: There will be **two in-class exams and a final exam**. Coverage includes lecture material, assigned reading, and homework. **All exams are in-person and closed book/closed notes.** The two in-class exams will take place at the usual lecture room during regular class hours. The final exam location and time are university scheduled and will be communicated as the information become available. Further details will be given in advance of each exam. Statistical tables will be provided as needed. Calculators may be used, but no communication devices are allowed (e.g. mobile phones). You may use one 8.5x11 inch handwritten sheet of paper (both sides) with formulas. Makeup exams require a valid excuse and official proof if I am notified in advance or as soon as possible.

Tentative Dates for exams:

Exam 1: Friday, Oct 06, 2023, during class time.

Exam 2: Monday, Nov 06, 2023, during class time.

Final Exam (University scheduled): Friday, Dec 08, 2023, 12:00 PM - 1:45 PM.

Faculty feedback and response time

I am providing the following list to give you an idea of my intended availability throughout the course. (Remember that you can call **614-688-HELP** at any time if you have a technical problem.)

Grading and feedback

For large weekly assignments, you can generally expect graded homeworks back within **14 days**.

E-mail

I will reply to e-mails related to course logistics within **72 hours on school days**. However, if you have questions on understanding the course materials or regarding the homework assignments please visit my Zoom office hours, the TA's office hours and Data Analytics Learning Center.

Attendance, participation, and discussions

Students may miss class, for a variety of reasons including those related to COVID-19. As much as possible, please stay in contact with the instructor so that we can discuss accommodations should they be needed.

Student participation requirements

- **Logging in: AT LEAST ONCE PER WEEK**
Be sure you are logging in to the course in Carmen each week, including weeks with holidays or weeks with minimal online course activity. (During most weeks you will probably log in many times.) If you have a situation that might cause you to miss an entire week of class, discuss it with me *as soon as possible*.
- **In-person lectures: REGULAR ATTENDANCE EXPECTED** Students will be expected to participate, discuss, and answer questions during in-person lectures.
- **Office hours: OPTIONAL OR FLEXIBLE**
All office hours are optional.

Other course policies

Health and safety

The Ohio State University Wexner Medical Center's Coronavirus Outbreak site (<https://wexnermedical.osu.edu/features/coronavirus>) includes the latest information about COVID-19 as well as guidance for students, faculty and staff. Guidelines and requirements for campus safety from the University's COVID-19 Transition Task Force were published on July 1

on the Safe and Healthy website (<https://safeandhealthy.osu.edu>).

Potential disruptions to instruction

- As much as is possible, students will have access to material online if they are unable to attend class because of positive diagnosis, symptoms, or quarantine required following contact tracing.
- If the instructor is unable to be present in person because of positive diagnosis, symptoms, or quarantine following contact tracing alternative arrangement of instruction will be made. Details will be given on the course website

Student academic services

Student academic services offered on the OSU main campus

<http://advising.osu.edu/welcome.shtml>.

Student support services

Student support services offered on the OSU main campus <http://ssc.osu.edu>.

Academic integrity policy

Policies for this course

- **Exams:** You must complete the exams yourself, without any external help or communication.
- **Written assignments:** Your written assignments, including discussion posts, should be your own original work. You are allowed to discuss with your peers and work together on the assignments. However, the submission must be written in your own words.
- **Reusing past work:** In general, you are prohibited in university courses from turning in work from a past class to your current class, even if you modify it. If you want to build on past research or revisit a topic you've explored in previous courses, please discuss the situation with me.
- **Falsifying research or results:** All research you will conduct in this course is intended to be a learning experience; you should never feel tempted to make your results or your library research look more successful than it was.

Ohio State's 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/>.

Copyright disclaimer

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 <http://titleix.osu.edu> or by contacting the Ohio State Title IX Coordinator at titleix@osu.edu

Accessibility accommodations for students with disabilities

The university strives to make all learning experiences as accessible as possible. In light of the current pandemic, students seeking to request COVID-related accommodations may do so through the university's request process, managed by Student Life Disability Services. 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. SLDS contact information: slds@osu.edu; 614-292-3307; <http://slds.osu.edu>; 098 Baker Hall, 113 W. 12th Avenue.

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 through the 24/7 National Suicide Prevention Hotline at 1-800-273- TALK or at suicidepreventionlifeline.org

Disclaimer

This syllabus should be taken as a fairly reliable guide for the course content. However, you cannot claim any rights from it and in particular we reserve the right to change due dates or the methods of grading and/or assessment if necessary. Any changes will be communicated to you through official course announcements.

Course schedule (tentative)

See excel calendar file in the Syllabus tab.