

# Statistics 3201: Introduction to Probability for Data Analytics

## Spring 2020 Course Syllabus

**Lectures:** 11:30am – 12:25pm on Mondays, Wednesdays, Fridays in Pomerene Hall 250.

**Instructor:** Prateek Sasan

**Graduate Teaching Assistant:** Azriel Krongauz

**Office Hours:**

- Instructor: Tuesday 10.30-11.30am and Thursday 9-10am in Pomerene 161. Individual appointments outside of office hours must be requested via email and will not be available on short notice: students must provide a list of their available meeting times for the next three to five days.
- Graduate Teaching Assistant: Refer to Data Analytics Learning Center Document

**Email:** Please begin subject with “STAT 3201”; *use my OSU email sasan.1@osu.edu* and avoid using the Carmen email tool. In order to protect your privacy, all course email correspondence must be conducted using your valid OSU name. # email account: any email from a non-OSU account will not be answered. I will attempt to answer emails within 24 hours, however, due to the large volume of emails this may not always be possible. Before writing an email, check whether the question has already been answered in the syllabus, the notes, or the textbook. Also please consider whether your question would be best answered in person during office hours. I will sometimes receive questions via email regarding homework problems or clarifications: if your question may be helpful to other students, I will it anonymously but verbatim along with my answer in the Discussions section on the class Carmen page.

- Instructor: sasan.1@osu.edu
- Graduate Teaching Assistant: krongauz.1@osu.edu

**Extra help:** Graduate teaching assistants (GTAs) for Stat 3201, 3202, 3301, 3302, 3303 and 4620 will hold their office hours in the Data Analytics Learning Center (DALC) in Pomerene 151. The hours during which the GTA for our course will hold office hours can be found at the top of the syllabus. You can meet with the GTA for our course in the DALC during his or her office hours to discuss questions you have about the course material, homework assignments, R, etc.

You are welcome to stop by the DALC when it is open, even if it is not currently being staffed by the GTA for our course, e.g. if you are looking for a place to study or work on an assignment for one of the supported courses. If the DALC is staffed by a GTA for another Statistics course when

you stop by, he or she will help you if possible, but may not be able to answer all of your questions.

A complete list of hours during which the DALC will be staffed by GTAs for Statistics Department courses can be found at <https://data-analytics.osu.edu/dalc>.

In rare situations due to last minute emergencies, the GTA assigned to the DALC may not be able to attend his or her office hours. If the DALC is closed when the schedule indicates it should be open, we recommend waiting for a few minutes. If no one shows up in a reasonable amount of time, please email your instructor to let us know about the problem. You can also contact your GTA to see about arranging a make-up time to meet.

**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.

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

**Prerequisites:** Prerequisites: Math 1152 or 1161.xx or 1172 or 1181 or equivalent, and CSE Placement Level A or equivalent; or permission of the instructor.

**Course Website:** Important announcements, course materials, homework problems, computing references, and other information about the class are posted on Carmen ([carmen.osu.edu](http://carmen.osu.edu), login with your web ID).

**Textbook:** The required textbook for this course is:

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

**Course Materials:** A reading list from the required textbook will be provided for each lecture. Please read the sections of the textbook that will be covered, and any additional material posted on Carmen prior to the class.

**Homework Assignments:** Homework will be assigned approximately biweekly. It will consist of mostly textbook-style problems, problems motivated by data analytics applications, and small computer simulation problems. Question numbers referenced in the homework are from the textbook edition listed above. If you are using a different edition/version of the textbook, it is your responsibility to check that you have solved the correct questions. No points will be awarded for answering a question other than the one being assigned. You are encouraged to work together on the problems, but each student must hand in his or her own work, written in his or her own words. **Do not copy any part of another student's homework including computer code or 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. **Sharing or disseminating solutions, or in any way knowingly enabling others to commit academic misconduct also constitutes academic misconduct, and will be reported. A hard copy of the homework solutions should be submitted at the beginning of class on the due date.** The solutions may be handwritten or typed, except any R code and output, which must always be typed. 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. **Electronic versions of homework submissions will not be accepted unless permission from the instructor is obtained in advance.** If advance permission is not obtained, and the reason is not a provable emergency, the instructor reserves the right to consider the homework as late. In other words, please do your best to get hard copies of your assignments to me on time. For example, consider sending a copy to a friend in the class and have them print it and submit it. **Please staple all submissions as pages of loose homework often become lost. Neither I nor the grader will accept responsibility for any lost pages if the homework is not stapled.**

**Exams:** There will be three in-class exams. Coverage includes lecture material, assigned reading, and homework. Tentative dates are provided on the weekly lesson plan. 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 8.5×11 inch handwritten sheet of paper (both sides) with formulas for all exams. Make-up exams require a valid excuse and official proof if the instructor is not notified in advance or as soon as possible. A make-up exam must be taken within a week of the missed exam. Exceptions to this policy are permitted only in extreme situations such as serious injury immediately prior to an exam or severe illness requiring hospitalization.

**Project:** A class project will consist of a written report and will require use of the R software.

**Computing:** We will be using the R statistical computing software. R may be downloaded for free from <http://www.r-project.org/>. Many students prefer to use the interface RStudio, available for free at <http://www.rstudio.com>.

**Attendance:** Regular attendance and class participation is required. Please let the instructor know via email if you plan to miss several lectures. Though attendance will not be taken daily, please remember that I fully am aware of which students consistently miss class.

**Grading:** In order to obtain full credit on homework and exam problems you need to show a justification or full work. Answers without supporting work will not receive full credit. The following is a breakdown of the final course grade:

Homework: 25%  
Exam 1(Monday, Feb 17): 20%  
Exam 2(Friday, Mar 27): 20%  
Final Exam: 25%  
Project: 10%

The following rubric will be used to compute the final letter grade: A: 93 – 100, A-: 90 – 92.9, B+: 87– 89.9, B: 83 – 86.9, B-: 80-82.9, C+: 77-79.9, C: 73 – 76.9, C-: 70–72.9, D+: 67 – 69.9, D: 60–66.9, E: below 60. The instructor reserves the right to make appropriate changes to the above if necessary. However, as usual there are no exceptions nor arbitrary grade adjustments for individual students, nor grade guarantees of any kind, for any reason.

**Special Considerations:** If a situation exists or arises that you think may hinder your progress in this class, you must notify me as soon as possible.

**Advising:** For questions related to prerequisites and course suggestions, please contact Academic Planning Specialist Brooke O’Leary (oleary.158@osu.edu).

**Academic Misconduct:** 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/>.

**Disability Services:** The University strives to make all learning experiences as accessible as possible. 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](mailto:slds@osu.edu); 614-292-3307; [slds.osu.edu](http://slds.osu.edu); 098 Baker Hall, 113 W. 12th Avenue.

**Mental Health Statement:** 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](http://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](http://suicidepreventionlifeline.org).