THE OHIO STATE UNIVERSITY

STAT 3201
Introduction to Probability for Data Analytics (24998)
Autumn 2026

Syllabus - Instructor Information

Steven MacEachern

Professional Qualifications
PhD, Statistics, University of Minnesota

Bachelor's, Mathematics, Carleton College

Syllabus - Required or Recommended Readings

The required textbook for this course is

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

The text will be supplemented with additional material. For computing with R, the following free resource will prove useful

Introduction to Probability and Statistics using R, by Kerns; available online at https://ipsur.r-forge.r-project.org/book/download/IPSUR.pdf

Syllabus - Course Schedule

Tentative schedule for STAT 3201; MacEachern; Fall 2026

Text is Mathematical Statistics with Applications, by Wackerly, Mendenhall and Scheafer, 7th edition.

Week 1. Course introduction. Introduction to R. Installation of R and basic use (1.1-1.2). 2 lectures; class begins on Wednesday.

Week 2. More on R. Summarization of data – graphical and numerical summaries. From summarization of a sample to inference about a population (1.3-1.6). Basics of probability. Definitions, formal mathematics for probability (2.1-2.2). 3 lectures.

Week 3. Probability models for discrete data. Calculating a probability (2.3-2.5). 2 lectures. Monday is Labor Day, no class.

Week 4. Counting methods. Conditional probability (2.6-2.7). 3 lectures.

Week 5. Bayes' rule. Strategies for computing probabilities. Examples (2.8-2.13). 3 lectures.

Week 6. Discrete random variables. Probability mass function, mean and variance. Bernoulli distribution (3.1-3.3). 3 lectures.

Week 7. Exam 1. Named discrete random variables. Binomial distribution. Geometric distribution. Negative binomial distribution. Hypergeometric distribution. (3.4-3.7). 3 lectures; one lecture devoted to the exam. [Exam – Monday Oct 5]

Week 8. Poisson distribution. Continuous random variables. Probability density function. Computing probabilities (3.8; 4.1-4.2). 2 lectures; Friday is Fall Break, no class.

Week 9. Continuous random variables – mean and variance. Named continuous

distributions. Uniform distribution. Normal (Gaussian) distribution (4.3-4.5). 3 lectures.

Week 10. Named continuous distributions. Exponential distribution. Gamma distribution. Beta distribution (4.6-4.7). 3 lectures.

Week 11. Functions of random variables. Method of distribution functions. Method of transformation (6.1-6.4). 3 lectures.

Week 12. Catch up. Exam 2. 2 lectures; one lecture devoted to the exam. Wednesday is Veteran's Day. [Exam - Friday Nov 13]

Week 13. Sampling distributions. Simulations. The central limit theorem (7.1-7.3). 3 lectures.

Week 14. Bivariate probability distributions. Computing probabilities. Marginal and conditional distributions (5.1). 1 lecture; Wednesday and Friday are Thanksgiving Break, no class.

Week 15. A random sample (independent random variables) (5.2-5.4). 3 lectures.

Week 16. Covariance and correlation. Conditional expected values (5.7, 5.11). 2 lectures.

Final Exam. As scheduled by the registrar. In the regular classroom. Friday, December 11, from 12:00 noon to 1:45 PM.

Course Details - Instructor Contact Details

Email: maceachern.1@osu.edu

Course Details - Office Hours

Office hours: Monday, 11:30 - 12:30; Friday, 10:30 - 11:30; and by appointment

Office location: 440D Cockins Hall

Course Details - Additional Instructional Staff

The grader for this course is Feilin Feng, a graduate student in the Statistics Department. Full contact information will be provided later.

Course Details - Class Meeting Details

Class Meeting Details: 03:00 PM-03:55 PM - Evans Lab 2004 - MWF

Course Details - 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. Applications of probability and sampling to big-data settings are discussed. Prereq: Math 1152, 1161.xx, 1172, 1181, or equiv; or permission of instructor. Not open to students with credit for 4201 or Math 4530.
Prerequisite(s) / Corequisites: Prereq: Math 1152, 1161.xx, 1172, 1181, or equiv; or permission of instructor. Not open to students with credit for 4201 or Math 4530.

Course Details - More About This Course

This course serves several audiences -- the majors in Statistics and in Data Analytics, the minor in Statistics, and those who simply want to know a bit about probability and statistics. It is the first substantive course in the minor/major and

is part of a year long sequence. This semester focuses on probability, from the basic definition and historical problems, to discrete and continuous random variables, to varied techniques useful for summarizing and working with random variables, and beyond. Along the way, we will learn about parallel summarizations of data sets.

Paper and pencil work:

Much of the material in this course can only be internalized by reading the text, participating actively in lectures, and working problems "by hand". Reading assignments, lectures, and homework assignments are designed to help you build the skills that will be helpful in later courses and in your post-collegiate life! Investing time and effort now will be worthwhile in in the long run.

Computing:

The course will make use of the R computing environment. R is a statistical software package that is available for free download. I strongly encourage you to make use of RStudio, as this platform will be used heavily in future courses. This course will get you the tiniest start on R. You will use R more in subsequent courses in the minor and majors.

Course Details - Course Learning Outcomes

By the end of this course, students should successfully 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 variables

7. Quantify uncertainty about summary statistics for large data sets

Course Details - Course Supplies

The text.

A calculator capable of doing simple calculations (including a square root). The calculator should not have internet connectivity. You do not need to buy a special "statistical calculator".

Course Details - Detailed Course Schedule

See above for the tentative course schedule.

Tentative midterm exam dates:

Monday, October 5

Final exam date (as scheduled by the registrar):

Friday, December 11, Noon - 1:45 PM

All exams will be in person in the regular classroom.

Course Details - Grading Policy

Your numerical grade for this course will be computed on the basis of:

Class Participation: 5%

Homework Assignments: 10%

Midterm Exams (two): 50% (25% each)

Final Exam: 35%

Your two lowest homework scores will be dropped before computing your homework grade. If you miss a homework assignment, you will receive a zero for the assignment (likely one that you will drop).

Grades are tied to performance. Grades will not be lower than those on the 90/80/70/60 scale. Grade cutoffs will, in all likelihood, be lower than 90 for an A (of some sort) and lower than 80 for a B (of some sort). Similarly, cutoffs for C and D grades are likely to be lower than 70 or 60. I reserve the right to make adjustments to grades if circumstances warrant.

Course Details - Grading Scale

See above. Grades are based on performance, not a set point scale.

Course Details - Other

Assignments:

The reading assignments will not be graded. But read the book. It's a good one.

You will have homework assignments at a pace of roughly one assignment per week. The baseline expectation is for you to spend two hours outside of class on this course for each nominal hour in class. Depending on your background and the pace at which you work, your actual time may be longer or shorter than this expectation.

Homework assignments will contain a mix of problems. Some will get you comfortable with the basic calculations and mechanics of working with probability. Most of these will come from the textbook. Others will stretch you a bit. Problems that involve R are custom-written. The assignments will be

available on Carmen. Solutions to the assignments will be made available after the assignments are due.

The midterm exams will be closed book, closed notes. Bring a non-internet-connected calculator for brief numerical calculations. You will be provided with a formula sheet for the exam. Each exam will take one class period. Tentative dates for the exams are Monday, October 5 and Friday November 13. These dates are tentative. They may change, depending on the pace of lectures.

Assignments will be due, on Carmen, as a pdf. The choice of format is to make grading more efficient and to make it easier to provide you with feedback. For late assignments, one (business) day late is a deduction of 10%; two days late, of 20%; beyond two days late, the assignment will not be accepted.

Attendance:

The class is designed around the lectures. Attendance is strongly encouraged as is class participation. Over time, I have tracked the performance of regular class attenders and non-attenders. There is a substantial difference in performance for these two groups. Show up, read the book, do the homework (and a bit more, if time and inclination permit), and you will do well in class. If you have questions, ask in class or see me after class or during office hours.

Note that participation is part of your grade. Attendance is one component of participation.

If you are unable to attend due to illness or for other reasons, let me know. If, for example, you test positive for COVID or believe that you have symptoms, it is best that you take care of health concerns before worrying about this class. Let me know of the situation promptly and we will make special arrangements.

Course Details - AI Use Policy

The main purpose of this course is for you to learn the material and to increase your understanding of Statistics. For the homework assignments, you can talk

with fellow students and you can work on the problems with them. You may even wish to look on the web or use AI tools when working on the homework. This is acceptable, but your write-up should be your own – not a direct copy of a classmate's work, of a solution found online, or a write-up produced by AI. The exams will be in-class. For them, you will work solely on your own, so make sure that you have command of the material. For exams, a basic calculator will be permitted but cell phones will not be permitted. Tables will be provided as needed.

Note that this policy of open discussion and use of AI on the homework is specific to this course. Other instructors and other courses may well have a different policy. Make sure that you understand the policy for each course that you take.

Also note: this is one class in the Statistics and Data Analytics majors' curriculum. The curriculum has been designed to give you the breadth of skills and knowledge needed to function as an effective professional statistician / biostatistician / data scientist. In addition to individual courses, to progress to a degree, you must take (and pass) a comprehensive examination or two that covers much of the curriculum. The best route to success in the program – and for your future career – is to absorb much from the courses while taking them and to then fill in gaps in knowledge as needed.

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. I will report serious academic misconduct to the committee (Faculty Rule 3335-5-487). For additional information, see the Code of Student Conduct http://studentlife.osu.edu/csc/.

Syllabus - University Policies

Office of Undergraduate Education - Standard Syllabus Statements

- Academic Misconduct
- Religious Accommodations
- Accommodations for Disability and Illness
- Intellectual Diversity
- Grievances and Solving Problems
- Creating an Environment Free from Harassment, Discrimination, and Sexual Misconduct
- Copyrights
- Content Warnings
- Mental Health Support
- Resources for Military-Connected Students
- Campus-wide Class Cancellations