

# Math 254 - Statistical Modeling and Applications

Spring 2025, MWF 10:00-10:50 Christian A. Johnson (CJ) 305

## Instructor

Erin Tripp  
[etripp@hamilton.edu](mailto:etripp@hamilton.edu)

## Office Hours

CJ 107  
MWF 11-12, 2-3  
[Or by appointment](#)

## Course description

Math 254 is a second course in statistics for students who have already been exposed to the basic ideas of statistical inference. We focus on both parametric and nonparametric models, emphasizing how and why they work (and how to tweak them) rather than push-button results. Students will work with real, messy data using the statistical programming language R. There are three main units:

1. Review of and Alternatives for Statistical Inference
  - classical confidence intervals vs. bootstrap intervals
  - classical tests for means and proportions vs. randomization tests
  - chi-squared tests vs. Fisher's exact test
2. Regression
  - multiple linear regression
  - logistic regression
  - cross-validation
  - variable selection and shrinkage methods
3. Machine Learning
  - neighbor-based regression and classification
  - decision trees and random forests
  - support vector machines (time permitting)
  - clustering algorithms (time permitting)

## Textbook

Most topics will be drawn from two **free** online books:

1. [OpenIntro Statistics](#), 4th edition, by D. Diez et al., 2019
2. [An Introduction to Statistical Learning \(with Applications in R\)](#), 7th printing, by G. James et al., 2017

*See the Other Resources tab for additional references.*

## Software

RStudio (<https://posit.co/downloads/>)

We will be using the open source R IDE RStudio for labs and other assignments.

## Evaluation

*Assignment, % of total grade, Description*

<b>Cooperative Labs</b>	15%	You and 1-2 randomly chosen classmates will work through a guided exploration of a data set.  Labs will be submitted as PDFs on Gradescope.
<b>Homework</b>	15%	Practice problems. Work with whomever you wish, but everyone must submit their own solution. The lowest two scores will be dropped.  Homework will be submitted as images or PDFs on Gradescope.
<b>Midterm Exam</b>	25%	Take-home, open note exam with a two-hour time limit, to be taken during a 48-hour period.
<b>Project</b>	15%	Individuals or small groups will research a dataset of interest, provide a short written summary along with basic statistical analysis and visualization, and give a short oral presentation on their chosen data.
<b>Presentation</b>	5%	Short project presentations to be given in the last week of classes.
<b>Final Exam</b>	25%	Take-home, open note exam with a three-hour time limit, to be taken during the exam block if possible

## Blackboard and Google Drive

The Blackboard site will be an invaluable resource to find information and material for the course including: the syllabus and course calendar, assignments, exam practice problems,

Gradescope and Lumen information, updated course grades, and links to other campus and online resources.

I will use both Blackboard and Google Drive to share documents with you. Links to folders and files in Drive will be included in Blackboard, so always look there! You just need to be logged in with your Hamilton account to be able to access files.

## **Quantitative and Symbolic Reasoning (QSR) Intensive Course**

This course meets the College's expectations for the QSR-Intensive designation. Therefore, this class will:

- Include material in at least one of the following three categories:
  - Statistical Analysis. The use of statistical analysis to describe data and make inferences.
  - Mathematical Representation. The use of mathematical models such as those based on graphs, equations, and geometric objects to represent patterns, relationships, and forms.
  - Logic and Symbolic Reasoning. The use of formal logic or symbolic reasoning such as in the following examples: the proper construction of a computer program or a formal proof; the analysis of language in linguistics; or the study of music theory.
- Include four or more graded assignments (tests, quizzes, problem sets, labs, oral presentations, exhibits) in at least one of the three categories described above. QSR projects will be substantial and will be distributed across the semester.
- Provide explicit instruction in problem solving or data analysis techniques and strategies specific to the level and content of the course.

## **SSIH Requirement**

This course can be applied to the SSIH Requirement. The goal is to help students gain an understanding of structural and institutional hierarchies based on one or more of the social categories of race, class, gender, ethnicity, nationality, religion, sexuality, age, and abilities/disabilities. Throughout the course we will discuss these topics through examples, but to get credit as an SSIH course, the student's final project should address one or more of these categories in greater detail.

## **AI use**

No AI tools may be used for coursework unless explicitly stated in the assignment. This includes ChatGPT, Google Gemini, CoPilot, etc.

## **Honor code**

Our College has a strict Honor Code. Students may (should!) work together on homework, but must write their own submissions. Exams should be completed alone. All resources used, including other students, should be cited appropriately.

## **Late Work**

Your lowest two homework grades will be dropped. This policy is meant to automatically cover most situations in which students ask for extensions, without you having to email me. If you have a serious, long-term situation which will affect your ability to complete three or more assignments, please reach out, but otherwise, late work will not be accepted for any reason.

## **Accommodations**

The College will make reasonable accommodations for students with documented disabilities. Eligible students who would like to make a formal request for this course must contact me during the first two weeks of class.