MATHEMATICS AND STATISTICS (Div III)

STATISTICS

Chair of the Department of Mathematics and Statistics: C. Carlisle and Margaret Tippit Professor of Statistics, Richard De Veaux

- Colin C. Adams, Thomas T. Read Professor of Mathematics
- Bhagya Athukorallage, Visiting Assistant Professor of Mathematics
- Julie C. Blackwood, Associate Professor of Mathematics
- Xizhen Cai, Assistant Professor of Statistics
- Daniel Condon, Visiting Assistant Professor of Mathematics
- Richard D. De Veaux, Chair & C. Carlisle and Margaret Tippit Professor of Statistics
- Thomas A. Garrity, Webster Atwell Class of 1921 Professor of Mathematics
- Leo Goldmakher, Associate Professor of Mathematics
- Pamela E. Harris, Associate Professor of Mathematics, Faculty Fellow of the Davis Center and the Office of Institutional Diversity, Equity and Inclusion; affiliated with: The Davis Center, InstDiversity,Equity&Inclusion
- Stewart D. Johnson, Professor of Mathematics
- Bernhard Klingenberg, Professor of Statistics
- Susan R. Loepp, William R. Kenan, Jr. Professor of Mathematics
- Steven J. Miller, Professor of Mathematics
- Ralph E. Morrison, Assistant Professor of Mathematics
- Shaoyang Ning, Assistant Professor of Statistics
- Allison Pacelli, Professor of Mathematics
- Lori A. Pedersen, Lecturer in Mathematics
- Anna M. Plantinga, Assistant Professor of Statistics
- Cesar E. Silva, Hagey Family Professor of Mathematics
- Mihai Stoiciu, Professor of Mathematics
- Annie Tang, Visiting Lecturer of Statistics
- Laurie L. Tupper, Assistant Professor of Statistics
- Daniel B. Turek, Assistant Professor of Statistics
- Elizabeth M. Upton, Assistant Professor of Statistics
- Jenna Zomback, Visiting Assistant Professor of Mathematics

MAJOR IN STATISTICS

The major in Statistics is designed to meet three goals: to introduce some of the central ideas of information and data science, to develop problem-solving ability by teaching students to combine creative thinking with rigorous reasoning, and to develop interdisciplinary skills by applying statistics to an application area of interest.

REQUIREMENTS (10 courses plus colloquium)

The major in Statistics consists of ten courses plus a colloquium requirement. The major includes courses in mathematics, computer science and statistics. Students interested in continuing their study of statistics in graduate school should strongly consider taking Math 350/351 in addition to the other requirements.

Mathematics (2 courses)

MATH 150 or 151 Multivariable Calculus or equivalent high school course

MATH 250 Linear Algebra

Except in unusual circumstances, students planning to major in statistics should complete the calculus sequence (MATH 130, 140, 150/151) before the spring of the sophomore year, at the latest. MATH 150 is a prerequisite for STAT 201 and MATH 250 is a prerequisite for STAT 346.

Computer Science (1 course)

CSCI 134 Intro to Computer Science or CSCI 135 Diving into the Deluge of Data or CSCI 136 Data Structures and Advanced Programming or some other course in the Computer Science Department with prior approval of the Math/Stat department.

Core Courses (4 courses)

STAT 201 Statistics and Data Analysis, STAT 202 Introduction to Statistical Modeling or STAT 302 Applied Statistical Modeling

STAT 346 Regression and Forecasting

STAT 341 Probability

STAT 360 Inferential Statistics

Continuation (2 courses)

Any two courses among the 300 or 400 level courses in the department with a STAT prefix.

Capstone Course (1 course)

The capstone course is a 400-level STAT course taken in the senior year. Although the specific methodological emphasis of the course may vary from year to year, an in-depth project with both a written report and an oral presentation is typically part of the capstone course.

Pass/Fail policies during the Academic Year 2020-2021

Information about the Department of Mathematics and Statistics Pass/Fail policies during the Academic Year 2020-2021 can be found here.

Colloquium Requirement

Participation in the Department Colloquium, in which each senior major presents a talk on a mathematical or statistical topic of their choice. Each major must also attend at least 20 colloquia (15 during the Academic Year 2020-2021), and up to 5 attendances may be counted in their junior year. Up to 5 colloquia in mathematics or computer science may also be counted. Students engaged in study away may petition the department in advance to count up to 5 suitable colloquia attendances from their study away program.

PLACEMENT

Students with an AP Stat score of 5 or 4 are placed in the advanced introductory course Stat 202.

NOTES

Substitutions, Study Abroad, and Transfer Credit: In some cases, and with prior permission of the Mathematics and Statistics Department, appropriate courses from other institutions may be substituted for the application and continuation requirements, but at least eight courses must be taken from the Department of Mathematics and Statistics at Williams.

These can, with prior permission, include courses taken away. Students with transfer credit should contact the department about special arrangements.

Double Counting: No course may count towards two different majors.

Early Senior Capstone Course: In exceptional circumstances, with the prior permission of the department, a student may be allowed to satisfy the Senior Capstone Course requirement in the junior year, provided that the student has completed at least three 300-level statistics courses before enrolling in the capstone course.

Planning Courses: Core courses are normally offered every year. Other 300 and 400 level statistics courses are offered on an irregular basis. Students should check with the department before planning far into the future.

Course Admission: Courses are normally open to all students meeting the prerequisites, subject to any course caps. Students with questions about the level at which courses are conducted are invited to consult department faculty.

FAQ

Students MUST contact departments/programs BEFORE assuming study away credit will be granted toward the major or concentration.

Can your department or program typically pre-approve courses for major/concentration credit?

Yes, in many cases, though students should be sure to contact the department.

What criteria will typically be used/required to determine whether a student may receive major/concentration credit for a course taken

while on study away?

Course title and description, and complete syllabus including readings/assignments.

Does your department/program place restrictions on the number of major/concentration credits that a student might earn through study away?

No.

Does your department/program place restrictions on the types of courses that can be awarded credit towards your major?

Yes. They have to be approved MATH/STAT courses.

Are there specific major requirements that cannot be fulfilled while on study away?

Yes. Colloquium requirement, Senior Seminar requirement.

Are there specific major requirements in your department/program that students should be particularly aware of when weighing study away options? (Some examples might include a required course that is always taught in one semester, laboratory requirements.)

Yes. The highly cumulative structure of the major.

Give examples in which students thought or assumed that courses taken away would count toward the major or concentration and then learned they wouldn't:

None to date.

THE DEGREE WITH HONORS IN STATISTICS

The degree with honors in Statistics is awarded to the student who has demonstrated outstanding intellectual achievement in a program of study which extends beyond the requirements of the major. The principal considerations for recommending a student for the degree with honors will be: Mastery of core material and skills, breadth and, particularly, depth of knowledge beyond the core material, ability to pursue independent study of statistics, originality in methods of investigation, and, where appropriate, creativity in research.

An honors program normally consists of two semesters (STAT 493 and 494) and a winter study (WSP 031) of independent research, culminating in a thesis and a presentation. During the Academic Year 2020-2021 the winter study requirement for the honors program in Statistics is waved. One of STAT 493 or STAT 494 can count as a continuation course, but not both. Neither counts as the 400-level senior capstone course.

An honors program in actuarial studies requires significant achievement on four appropriate examinations of the Society of Actuaries.

Highest honors will be reserved for the rare student who has displayed exceptional ability, achievement or originality. Such a student usually will have written a thesis or pursued actuarial honors. In all cases, the award of honors and highest honors is the decision of the Department.

STAT 101 (F)(S) Elementary Statistics and Data Analysis (QFR)

It is impossible to be an informed citizen in today's world without an understanding of data. Whether it is opinion polls, unemployment rates, salary differences between men and women, the efficacy of vaccines, etc, we need to be able to interpret and gain information from statistics. This course will introduce the common methods used to analyze and present data with an emphasis on interpretation and informed decision making.

Requirements/Evaluation: weekly homework, quizzes, exams, and a project

Prerequisites: MATH 102 (or demonstrated proficiency on a diagnostic test)

Enrollment Limit: 50

Enrollment Preferences: juniors and seniors

Expected Class Size: 35

Grading: yes pass/fail option, yes fifth course option

Unit Notes: students with MATH130 but no statistics should enroll in STAT161; students with MATH150 but no statistics should enroll in STAT201. Students with AP Stat 4/5 or STAT 101/161 should enroll in STAT 202.

Distributions: (D3) (QFR)

Quantative/Formal Reasoning Notes: It is a quantitative course.

Attributes: COGS Related Courses PHLH Statistics Courses

Fall 2022

LEC Section: 01 TF 1:10 pm - 2:25 pm Elizabeth M. Upton

Spring 2023

LEC Section: 01 TF 1:10 pm - 2:25 pm Elizabeth M. Upton

STAT 161 (F)(S) Introductory Statistics for Social Science (QFR)

This course will cover the basics of modern statistical analysis with a view toward applications in the social sciences. Topics include exploratory data analysis, linear regression, basic statistical inference, and elements of probability theory. The course focuses on the application of statistical tools to solve problems, to make decisions, and the use of statistical thinking to understand the world.

Requirements/Evaluation: Weekly homework, quizzes, two midterms and a final exam (midterms include take-home components), and a data analysis project. Students will need to become familiar with the statistical software STATA.

Prerequisites: MATH 130 (or equivalent); not open to students who have completed STAT 101 or equivalent

Enrollment Limit: 40

Enrollment Preferences: Economics majors, sophomores

Expected Class Size: 40

Grading: yes pass/fail option, no fifth course option

Unit Notes: Students with calculus background should consider STAT 201. Students without any calculus background should consider STAT 101. Students with AP Stat 4 or 5 should consider Stat 202. Please refer to the placement chart on the Math&Stat department website for more information.

Distributions: (D3) (QFR)

Quantative/Formal Reasoning Notes: Reasoning with data

Attributes: PHLH Statistics Courses

Fall 2022

LEC Section: 01 MWF 9:00 am - 9:50 am Anna M. Plantinga

LEC Section: 02 MWF 10:00 am - 10:50 am Anna M. Plantinga

Spring 2023

LEC Section: 01 MWF 10:00 am - 10:50 am Richard D. De Veaux

STAT 197 (F) Independent Study: Statistics

Directed 100-level independent study in Statistics.

Requirements/Evaluation: decided by the department

Prerequisites: permission of department

Enrollment Limit: 20

Enrollment Preferences: decided by the department

Expected Class Size: 1

Grading: yes pass/fail option, yes fifth course option

Distributions: (D3)

Fall 2022

IND Section: 01 TBA Richard D. De Veaux

STAT 198 (S) Independent Study: Statistics

Directed 100-level independent study in Statistics.

Requirements/Evaluation: decided by the department

Prerequisites: permission of department

Enrollment Limit: 20

Enrollment Preferences: decided by the department

Expected Class Size: 1

Grading: yes pass/fail option, yes fifth course option

Distributions: (D3)

Spring 2023

IND Section: 01 TBA Richard D. De Veaux

STAT 201 (F)(S) Statistics and Data Analysis (QFR)

Statistics can be viewed as the art and science of turning data into information. Real world decision-making, whether in business or science is often based on data and the perceived information it contains. Sherlock Holmes, when prematurely asked the merits of a case by Dr. Watson, snapped back, "Data, data, data! I can't make bricks without clay." In this course, we will study the basic methods by which statisticians attempt to extract information from data. These will include many of the standard tools of statistical inference such as hypothesis testing, confidence intervals, and linear regression as well as exploratory and graphical data analysis techniques. This is an accelerated introductory statistics course that involves computational programming and incorporates modern statistical techniques.

Requirements/Evaluation: weekly homework and projects, midterm exams, and a final exam.

Prerequisites: MATH 150 or equivalent; not open to students who have completed STAT 101 or STAT 161 or equivalent

Enrollment Limit: 30

Enrollment Preferences: Prospective Statistics majors, students for whom the course is a major prerequisite, and seniors

Expected Class Size: 30

Grading: yes pass/fail option, yes fifth course option

Unit Notes: Students with AP Stat 4/5 or STAT 101/161 should enroll in STAT 202. Students with no calc or stats background should enroll in STAT 101. Students with MATH 140 but no statistics should enroll in STAT 161.

Distributions: (D3) (QFR)

Quantative/Formal Reasoning Notes: Students will learn to interpret, choose, carry out, and communicate analyses of data.

Attributes: COGS Related Courses PHLH Statistics Courses

Fall 2022

LEC Section: 01 MWF 12:00 pm - 12:50 pm Stewart D. Johnson

LEC Section: 02 MWF 11:00 am - 11:50 am Stewart D. Johnson

Spring 2023

LEC Section: 01 MWF 9:00 am - 9:50 am Anna M. Plantinga

STAT 202 (F)(S) Introduction to Statistical Modeling (QFR)

Data come from a variety of sources: sometimes from planned experiments or designed surveys, sometimes by less organized means. In this course we'll explore the kinds of models and predictions that we can make from both kinds of data, as well as design aspects of collecting data. We'll focus on model building, especially multiple regression, and talk about its potential to answer questions about the world -- and about its limitations. We'll emphasize applications over theory and analyze real data sets throughout the course.

Requirements/Evaluation: Homework problems; quizzes; exams; a final project (on a topic that interests you!). Participation matters! Engagement with your peers is an important part of learning, of being a statistician in the Real World...and of your evaluation in this course. While your assignments will be submitted (and graded) individually, you'll be responsible for giving and receiving peer feedback, contributing to class discussions, and working together with classmates on practice problems.

Prerequisites: MATH 140 and STAT 101/161/201/AP Statistics 4/5, or permission of instructor.

Enrollment Limit: 40

Enrollment Preferences: Prospective Statistics majors and more senior students

Expected Class Size: 25

Grading: yes pass/fail option, no fifth course option

Unit Notes: Students with a 4 on the AP Stats exam should contact the department for proper placement. Students with STAT 201 are strongly encouraged to take STAT 346 or other 300-level statistics electives.

Distributions: (D3) (QFR)

Quantative/Formal Reasoning Notes: This course uses mathematical tools and computing programs to create models, make predictions, assess uncertainty, and describe data. We'll also emphasize choosing appropriate mathematical tools and interpreting their results in a real-world context.

Attributes: PHLH Statistics Courses

Fall 2022

LEC Section: 01 TF 1:10 pm - 2:25 pm Xizhen Cai

LEC Section: 02 TR 9:55 am - 11:10 am Daniel B. Turek

Spring 2023

LEC Section: 01 TF 1:10 pm - 2:25 pm Xizhen Cai

LEC Section: 02 TR 9:55 am - 11:10 am Daniel B. Turek

STAT 297 (F) Independent Study: Statistics

Directed 200-level independent study in Statistics.

Requirements/Evaluation: decided by the department

Prerequisites: permission of department

Enrollment Limit: 20

Enrollment Preferences: decided by the department

Expected Class Size: 1

Grading: yes pass/fail option, yes fifth course option

Distributions: (D3)

Fall 2022

IND Section: 01 TBA Richard D. De Veaux

STAT 298 (S) Independent Study: Statistics

Directed 200-level independent study in Statistics.

Requirements/Evaluation: decided by the department

Prerequisites: permission of department

Enrollment Limit: 20

Enrollment Preferences: decided by the department

Expected Class Size: 1

Grading: yes pass/fail option, yes fifth course option

Distributions: (D3)

Spring 2023

IND Section: 01 TBA Richard D. De Veaux

STAT 302 (S) Applied Statistical Modeling (QFR)

Data may come from various sources and studies with different purpose of analysis. Statistical modeling provides a unified framework to embrace different data types, and focuses on the goals of understanding relationships, assessing differences and making predictions. We will explore different types of statistical models (linear regression, ANOVA, logistic regression etc), and focus on their conditions, the interactive modeling process, as well

as the statistical inference tools for drawing conclusions from them. Throughout the course, real datasets will be modeled for interesting questions about the world, and the limitations will be addressed as well.

Class Format: Hybrid format

Requirements/Evaluation: weekly homework assignments, quizzes, exams and a course project.

Prerequisites: One of the following: i) STAT 201; ii) MATH 140 and STAT 101/161/AP Statistics 4/5; iii) Permission of instructor

Enrollment Limit: 19

Enrollment Preferences: Students interested in statistics who have background in calculus and intro stat. Students cannot take STAT 302 either simultaneously or after STAT 346.

Expected Class Size: 15

Grading: yes pass/fail option, no fifth course option

Distributions: (D3) (QFR)

Quantative/Formal Reasoning Notes: It is an intermediate statistics class with prerequisites that are QFR courses

Not offered current academic year

STAT 310 (F) Data Visualization (QFR)

This course is about preparing, visualizing, reporting and presenting different types of data. We will start with creating common plots (e.g., barcharts, histograms, density plots, boxplots, time series and lattice plots), but also discuss visualizing results of statistical models, such as linear or logistic regression models. We will use the ggplot library in R but then switch to the plotly library for interactive graphs with mouse-over and click events. Using R's shiny and DT libraries, we will learn how to create and publish web-apps and dashboards that explore datasets and support online filtering. We will end the class with creating web apps that contain multiple graphs or maps which react to user inputs (such as selecting which variables to plot) or provide real time monitoring of streaming data. Throughout, we will use version control software (Github) to organize and keep track of our code. This course will be taught in a semi-flipped style. While the instructor will introduce certain topics, students will often be responsible for reading material ahead of time and then work individually or in pairs to reproduce material or implement it on their own data.

Requirements/Evaluation: Grading will almost entirely be based on class participation, individual and team-work, project presentations and the student's portfolio.

Prerequisites: Stat 201/202/302; Good knowledge of R

Enrollment Limit: 15

Enrollment Preferences: Preference may be given to stats majors who need the course in order to graduate, but then random selection.

Expected Class Size: 15

Grading: no pass/fail option, yes fifth course option

Distributions: (D3) (QFR)

Quantative/Formal Reasoning Notes: This course teaches how to organize and present data graphically, but also how to critique existing data visualizations.

Not offered current academic year

STAT 315 (S) Applied Machine Learning (QFR)

How does Netflix recommend films based on your viewing history? How does Facebook group its users and send out targeted ads? How did Google select from thousands of search terms to predict flu? Machine learning (ML) is a rapidly growing field that is concerned with algorithms and models to find patterns in data and solve these practical problems at the intersection between statistics, data science and computer science. This course provides a broad introduction to ideas and methods in machine learning, with emphasis on statistical intuitions and practical data analysis. Topics including regularized regression, SVM, supervised/unsupervised learning, text analysis, neural networks will be covered. Students will use R extensively throughout the course while getting introduced to some ML tools in Python.

Class Format: Hybrid format. Students cannot take both STAT 315 and STAT 442. Only one of the two can be taken for credit.

Requirements/Evaluation: weekly homework, one class project, and two or three exams

Prerequisites: MATH 140, and STAT 201/202, or equivalent; or permission of instructor. Students cannot take both STAT 315 and STAT 442. Only one of the two can be taken for credit.

Enrollment Limit: 15

Enrollment Preferences: Seniors.

Expected Class Size: 15

Grading: yes pass/fail option, yes fifth course option

Distributions: (D3) (QFR)

Quantative/Formal Reasoning Notes: This is a statistics class with a focus on mathematical, computational, and data analysis skills as well as appropriate practical application of analysis methods

Not offered current academic year

STAT 319 (S) Statistical Computing (QFR)

This course introduces a variety of computational and data-centric topics of applied statistics, which are broadly useful for acquiring, manipulating, visualizing, and analyzing data. We begin with the R language, which will be used extensively throughout the course. Then we'll introduce a variety of other useful tools, including the UNIX environment, scripting analyses using bash, databases and the SQL language, alternative data formats, techniques for visualizing high-dimensional data, and text manipulation using regular expressions. We'll also cover some modern statistical techniques along the way, which are made possible thanks to advances in computational power. This course is strongly computer oriented, and assignments will be project-based.

Requirements/Evaluation: based primarily on projects, homework, and exams

Prerequisites: STAT 201 or 202 and CSCI 134, 135, or 136

Enrollment Limit: 30

Enrollment Preferences: juniors and seniors, Statistics majors

Expected Class Size: 15

Grading: yes pass/fail option, yes fifth course option

Distributions: (D3) (QFR)

Quantative/Formal Reasoning Notes: This course uses statistical tools and programming techniques to acquire data, create visualizations, and make future predictions.

Not offered current academic year

STAT 335 (S) Biostatistics and Epidemiology (QFR)

Epidemiology is the study of disease and disability in human populations, while biostatistics focuses on the development and application of statistical methods to address questions that arise in medicine, public health, or biology. This course will begin with epidemiological study designs and core concepts in epidemiology, followed by key statistical methods in public health research. Topics will include multiple regression, analysis of categorical data (two sample methods, sets of 2x2 tables, RxC tables, and logistic regression), survival analysis (Cox proportional hazards model), and a brief introduction to regression with correlated data.

Class Format: Hybrid format

Requirements/Evaluation: evaluation will be primarily based on weekly homework, two midterm exams, a final exam, and a data analysis project

Prerequisites: STAT 201, STAT 202 and MATH 140, or permission of instructor

Enrollment Limit: 15

Enrollment Preferences: Junior and senior statistics majors; public health concentrators

Expected Class Size: 15

Grading: yes pass/fail option, yes fifth course option

Distributions: (D3) (QFR)

Quantative/Formal Reasoning Notes: This is a statistics course with a focus on quantitative methods relevant to public health studies.

Attributes: PHLH Statistics Courses

Not offered current academic year

STAT 341 (F)(S) Probability (QFR)

Cross-listings: STAT 341 MATH 341

Secondary Cross-listing

The historical roots of probability lie in the study of games of chance. Modern probability, however, is a mathematical discipline that has wide applications in a myriad of other mathematical and physical sciences. Drawing on classical gaming examples for motivation, this course will present axiomatic and mathematical aspects of probability. Included will be discussions of random variables (both discrete and continuous), distribution and expectation, independence, laws of large numbers, and the well-known Central Limit Theorem. Many interesting and important applications will also be presented, including some from classical Poisson processes, random walks and Markov Chains.

Requirements/Evaluation: homework, classwork, and exams

Prerequisites: MATH 150 and MATH 250 or permission of the instructor

Enrollment Limit: 30

Enrollment Preferences: Priority will be given to Mathematics majors and to Statistics Majors.

Expected Class Size: 20

Grading: yes pass/fail option, yes fifth course option

Distributions: (D3) (QFR)

This course is cross-listed and the prefixes carry the following divisional credit:

STAT 341 (D3) MATH 341 (D3)

Quantative/Formal Reasoning Notes: This is a 300-level Math/Stat course.

Fall 2022

LEC Section: 01 TF 2:35 pm - 3:50 pm Mihai Stoiciu

Spring 2023

LEC Section: 01 TF 2:35 pm - 3:50 pm Mihai Stoiciu

STAT 342 (S) Introduction to Stochastic Processes (QFR)

Stochastic processes are mathematical models for random phenomena evolving in time or space. Examples include the number of people in a queue at time t or the accumulated claims paid by an insurance company in an interval of time t. This course introduces the basic concepts and techniques of stochastic processes used to construct models for a variety of problems of practical interest. The theory of Markov chains will guide our discussion as we cover topics such as martingales, random walks, Poisson process, birth and death processes, and Brownian motion.

Requirements/Evaluation: weekly homework/labs, classwork, and exams

Prerequisites: STAT 341

Enrollment Limit: 30

Enrollment Preferences: senior Statistics majors

Expected Class Size: 15

Grading: yes pass/fail option, yes fifth course option

Distributions: (D3) (QFR)

Quantative/Formal Reasoning Notes: This is a statistics class with a focus on mathematical skills and translating real world phenomena into mathematical descriptions.

Not offered current academic year

STAT 344 (F) Statistical Design of Experiments (QFR)

When you hear the word experiment you might be picturing white lab coats and pipettes, but businesses, especially e-commerce, are constantly experimenting as well. How do you get the most out of both scientific and business investigations? By doing the right experiment in the first place. We'll explore the techniques used to plan experiments that are both efficient and statistically sound. We'll learn how to analyze the data that come from these experiments and the conclusions we can draw from that analysis. We'll look at both classical tools like fractional factorial designs as well as optimal design, and see how these two frameworks differ in their philosophy and in what they can do. Throughout the course, we'll make extensive use of both R and JMP software to work with real-world data.

Requirements/Evaluation: Homework problems--both individual and in groups, midterm, final, and projects (on topics that interest you!).

Prerequisites: STAT 161 or 201 or 202, or equivalent, and Math 140 or equivalent, or permission of instructor

Enrollment Limit: 20

Enrollment Preferences: Statistics majors, seniors, juniors, sophomores, first years

Expected Class Size: 15

Grading: yes pass/fail option, yes fifth course option

Distributions: (D3) (QFR)

Quantative/Formal Reasoning Notes: This course uses mathematical tools and computing programs to design experiments, analyze their results, and assess their effectiveness. We'll also emphasize choosing appropriate mathematical tools and interpreting their results in a real-world context.

Fall 2022

LEC Section: 01 MWF 10:00 am - 10:50 am Richard D. De Veaux

STAT 346 (F)(S) Regression Theory and Applications (QFR)

This course focuses on the building of empirical models through data in order to predict, explain, and interpret scientific phenomena. Regression modeling is the most widely used method for analyzing and predicting a response data and for understand the relationship with explanatory variables. This course provides both theoretical and practical training in statistical modeling with particular emphasis on simple linear and multiple regression, using R to develop and diagnose models. The course covers the theory of multiple regression and diagnostics from a linear algebra perspective with emphasis on the practical application of the methods to real data sets. The data sets will be taken from a wide variety of disciplines.

Requirements/Evaluation: Weekly homework, theory and data analysis exams, final course project.

Prerequisites: MATH 250, and at least one of STAT 201 or 202. Or permission of the instructor.

Enrollment Limit: 30

Enrollment Preferences: Statistics Majors

Expected Class Size: 20

Grading: yes pass/fail option, no fifth course option

Distributions: (D3) (QFR)

Quantative/Formal Reasoning Notes: This course prepares students in the use of quantitative methods for the modeling, prediction and understanding of scientific phenomena.

Fall 2022

LEC Section: 01 TR 9:55 am - 11:10 am Xizhen Cai

Spring 2023

LEC Section: 01 TR 9:55 am - 11:10 am Elizabeth M. Upton

STAT 355 (F) Multivariate Statistical Analysis (QFR)

To better understand complex processes, we study how variables are related to one another, and how they work in combination. Therefore, we want to make inferences about more than one variable at time? Elementary statistical methods might not apply. In this course, we study the tools and the intuition that are necessary to analyze and describe such data sets. Topics covered will include data visualization techniques for high dimensional data sets, parametric and non-parametric techniques to estimate joint distributions, techniques for combining variables, as well as classification and clustering algorithms.

Class Format: This will be a hybrid course for students who are both remote and in-person, with a mix of synchronous and asynchronous elements

Requirements/Evaluation: homework, project/presentations, possibly one or two exams.

Prerequisites: MATH 250, and STAT 346 or permission of instructor

Enrollment Limit: 15

Enrollment Preferences: students interested in statistics which have solid background in math and stat

Expected Class Size: 10

Grading: yes pass/fail option, no fifth course option

Distributions: (D3) (QFR)

Quantative/Formal Reasoning Notes: It is an advanced statistics class with prerequisites that are QFR courses

Not offered current academic year

STAT 356 (F) Time Series Analysis (QFR)

Time series -- data collected over time -- crop up in applications from economics to engineering to transit. But because the observations are generally not independent, we need special methods to investigate them. This course will include exploratory methods and modeling for time series, including descriptive methods and checking for significance, and a foray into the frequency domain. We will emphasize applications to a variety of real data, explored using R.

Class Format: Introductory lectures will be available asynchronously as text and video; synchronous sessions will discuss questions from lecture, dive further into the material, and work on examples. You'll use chat and discussion boards to build community, study with classmates, and ask questions outside of class time. There will also be optional synchronous office hours/review sessions.

Requirements/Evaluation: Evaluation is primarily based on quizzes and projects (on topics that interest you!). You'll be given the opportunity to assess your own work and resubmit/reattempt assignments as you gain mastery of a topic. Participation matters! Engagement with your peers is an important part of learning, of being a statistician in the Real World...and of your evaluation in this course. While most assignments will be submitted (and graded) individually, you'll be responsible for giving and receiving peer feedback, contributing to live and online discussions, and working together with classmates on practice problems.

Prerequisites: STAT 346 (may be taken concurrently) or permission of instructor

Enrollment Limit: 15

Enrollment Preferences: Statistics majors, seniors

Expected Class Size: 15

Grading: yes pass/fail option, yes fifth course option

Distributions: (D3) (QFR)

Quantative/Formal Reasoning Notes: This course uses mathematical tools and computing programs to create models, make predictions, assess uncertainty, and describe data. We'll also emphasize choosing appropriate mathematical tools and interpreting their results in a real-world context.

Fall 2022

LEC Section: 01 Cancelled

STAT 358 (S) Introduction to Categorical Data Analysis (QFR)

This course focuses on methods for analyzing categorical response data. In contrast to continuous data, categorical data consist of observations classified into two or more categories. Traditional tools of statistical data analysis (such as linear regression) are not designed to handle such data and pose inappropriate assumptions. We will develop methods specifically designed for modeling categorical data, with applications in the social and biological sciences as well as in medical research, engineering and economics. This course has two parts. The first part will discuss statistical inference for parameters of categorical distributions (Bernoulli, Binomial, Multinomial, Poisson) and for measures of association arising in contingency tables (difference and ratio of proportions and odds ratios). Inferential methods covered include Wald, score and likelihood ratio tests and confidence intervals, as well as the bootstrap. The longer second part will focus on statistical modeling of categorical response data via generalized linear models, with a heavy focus on logistic regression models with both quantitative and categorical predictors and their interactions. Model fitting and inference will be based on maximum likelihood and carried out via R.

Requirements/Evaluation: Weekly homework assignments consisting of exercises from the textbook as well as data analysis problems, carried out using R. Occasional short in-class quizzes at the beginning of class. One Midterm (with both in-class and take-home component). Final Project with presentation. Final exam. Homework accounts for roughly 15% of the grade, quizzes for another 15%, midterm (in-class and take-home combined) and final for about 30% each, and project for the remaining 10%.

Prerequisites: STAT 346: Regression and Forecasting

Enrollment Limit: 15

Enrollment Preferences: stats majors

Expected Class Size: 15

Grading: yes pass/fail option, yes fifth course option

Distributions: (D3) (QFR)

Quantative/Formal Reasoning Notes: Students learn how to analyze data and communicate results.

Not offered current academic year

STAT 360 (S) Statistical Inference (QFR)

How do we estimate unknown parameters and express the uncertainty we have in our estimate? Is there an estimator that works best? Many topics from introductory statistics such as random variables, the central limit theorem, point and interval estimation and hypotheses testing will be revisited and put on a more rigorous mathematical footing. The focus is on maximum likelihood estimators and their properties. Bayesian and computer intensive resampling techniques (e.g., the bootstrap) will also be considered.

Class Format: For the Spring 2021 semester, synchronous zoom lectures are planned, where the instructor uses Google's jamboard to interact with students

Requirements/Evaluation: Homework, Quizzes, Exams

Prerequisites: MATH 250, STAT 201 or 202, STAT 341

Enrollment Limit: 15

Enrollment Preferences: Statistics majors

Expected Class Size: 15

Grading: no pass/fail option, yes fifth course option

Distributions: (D3) (QFR)

Quantative/Formal Reasoning Notes: A rigourous mathematical course laying the foundation for reasoning with data

Spring 2023

LEC Section: 01 TR 11:20 am - 12:35 pm Daniel B. Turek

STAT 365 (F) Bayesian Statistics (QFR)

The Bayesian approach to statistical inference represents a reversal of traditional (or frequentist) inference, in which data are viewed as being fixed and model parameters as unknown quantities. Interest and application of Bayesian methods have exploded in recent decades, being facilitated by recent advances in computational power. We begin with an introduction to Bayes' Theorem, the theoretical underpinning of Bayesian statistics which dates back to the 1700's, and the concepts of prior and posterior distributions, conjugacy, and closed-form Bayesian inference. Building on this, we introduce modern computational approaches to Bayesian inference, including Markov chain Monte Carlo (MCMC), Metropolis-Hastings sampling, and the theory underlying these simple and powerful methods. Students will become comfortable with modern software tools for MCMC using a variety of applied hierarchical modeling examples, and will use R for all statistical computing.

Requirements/Evaluation: homework and exams

Prerequisites: STAT 201 and MATH 150 and 250, or permission of instructor

Enrollment Limit: 30

Enrollment Preferences: juniors and seniors, Statistics majors

Expected Class Size: 20

Grading: yes pass/fail option, yes fifth course option

Distributions: (D3) (QFR)

Quantative/Formal Reasoning Notes: This course utilizes mathematics and computer-based tools for the Bayesian approach for analyzing data and making statistical inferences.

Fall 2022

LEC Section: 01 TR 11:20 am - 12:35 pm Daniel B. Turek

STAT 368 (S) Modern Nonparametric Statistics (QFR)

Many statistical procedures and tools are based on a set of assumptions, such as normality or other parametric models. But, what if some or all of these assumptions are not valid and the adopted models are miss-specified? This question leads to an active and fascinating field in modern statistics called nonparametric statistics, where few assumptions are made on data's distribution or the model structure to ensure great model flexibility and robustness. In this course, we start with a brief overview of classic rank-based tests (Wilcoxon, K-S test), and focus primarily on modern nonparametric inferential techniques, such as nonparametric density estimation, nonparametric regression, selection of smoothing parameter (cross-validation), bootstrap, randomization-based inference, clustering, and nonparametric Bayes. Throughout the semester we will examine these new methodologies and apply them on simulated and real datasets using R.

Requirements/Evaluation: performance on exams, homework, and a project

Prerequisites: STAT 201 and STAT 346, or permission of instructor.

Enrollment Limit: 30

Enrollment Preferences: senior Statistics majors

Expected Class Size: 15

Grading: yes pass/fail option, yes fifth course option

Distributions: (D3) (QFR)

Quantative/Formal Reasoning Notes: This is a statistics class with a focus on mathematical, computational, and data analysis skills as well as appropriate practical application of analysis methods.

Not offered current academic year

STAT 372 (S) Longitudinal Data Analysis (QFR)

This course explores modern statistical methods for drawing scientific inferences from longitudinal data, i.e., data collected repeatedly on experimental units over time. The independence assumption made for most classical statistical methods does not hold with this data structure because we have multiple measurements on each individual. Topics will include linear and generalized linear models for correlated data, including marginal and random effect models, as well as computational issues and methods for fitting these models. As time permits, we will also investigate joint modeling of longitudinal and time-to-event data. We will consider many applications in the social and biological sciences.

Requirements/Evaluation: performance on exams, weekly homework, and a project

Prerequisites: STAT 201 and STAT 346

Enrollment Limit: 15

Enrollment Preferences: junior and senior Statistics majors

Expected Class Size: 15

Grading: yes pass/fail option, yes fifth course option

Distributions: (D3) (QFR)

Quantative/Formal Reasoning Notes: The course will cover a variety of statistical analysis methods for longitudinal data.

Attributes: PHLH Statistics Courses

Spring 2023

LEC Section: 01 MWF 11:00 am - 11:50 am Anna M. Plantinga

STAT 397 (F) Independent Study: Statistics

Directed independent study in Statistics.

Prerequisites: permission of department

Grading: yes pass/fail option, yes fifth course option

Distributions: (D3)

Fall 2022

IND Section: 01 TBA Richard D. De Veaux

STAT 398 (S) Independent Study: Statistics

Directed independent study in Statistics.

Prerequisites: permission of department

Grading: yes pass/fail option, yes fifth course option

Distributions: (D3)

Spring 2023

IND Section: 01 TBA Richard D. De Veaux

STAT 410 (F) Statistical Genetics (QFR)

Genetic studies explore patterns of genetic variation in populations and the effect of genes on diseases or traits. This course provides an introduction to statistical and computational methods for genetic studies. Topics will include Mendelian traits (such as single nucleotide polymorphisms), genome-wide association studies, pathway-based analysis, and methods for population genetics. Students will be introduced to some of the major computational tools for genetic analysis, including PLINK and R/Bioconductor. The necessary background in genetics and biology will be provided alongside the statistical and computational methods.

Requirements/Evaluation: project work, homework, exams, and contribution to discussion

Prerequisites: STAT 346 and STAT 360, or permission of instructor

Enrollment Limit: 14

Enrollment Preferences: Statistics majors, juniors and seniors

Expected Class Size: 10

Grading: no pass/fail option, no fifth course option

Distributions: (D3) (QFR)

Quantative/Formal Reasoning Notes: This is a statistics class with a focus on mathematical, computational, and data analysis skills as well as appropriate practical application of analysis methods.

Attributes: BIGP Courses PHLH Statistics Courses

Not offered current academic year

STAT 441 (F) Information Theory and Applications

Cross-listings: MATH 441 CSCI 441 STAT 441

Primary Cross-listing

What is information? And how do we communicate information effectively? This course will introduce students to the fundamental ideas of Information Theory including entropy, communication channels, mutual information, and Kolmogorov complexity. These ideas have surprising connections to a fields as diverse as physics (statistical mechanics, thermodynamics), mathematics (ergodic theory and number theory), statistics and machine learning (Fisher information, Occam's razor), and electrical engineering (communication theory).

Requirements/Evaluation: Weekly homeworks, midterm(s), final exam.

Prerequisites: Math/Stat 341; Math 150 or 151; or permission of instructor.

Enrollment Limit: 30

Enrollment Preferences: Seniors; mathematics and statistics majors.

Expected Class Size: 25

Grading: yes pass/fail option, yes fifth course option

Distributions: (D3)

This course is cross-listed and the prefixes carry the following divisional credit:

MATH 441 (D3) CSCI 441 (D3) STAT 441 (D3)

Not offered current academic year

STAT 442 (S) Statistical Learning and Data Mining (QFR)

In both science and industry today, the ability to collect and store data can outpace our ability to analyze it. Traditional techniques in statistics are often unable to cope with the size and complexity of today's data bases and data warehouses. New methodologies in Statistics have recently been developed, designed to address these inadequacies, emphasizing visualization, exploration and empirical model building at the expense of traditional hypothesis testing. In this course we will examine these new techniques and apply them to a variety of real data sets.

Class Format: Students cannot take both STAT 315 and STAT 442. Only one of the two can be taken for credit.

Requirements/Evaluation: weekly homework, exams and an end-of-term project

Prerequisites: MATH/STAT 341 and STAT 346, or permission of instructor

Enrollment Limit: 20

Enrollment Preferences: Statistics majors, juniors and seniors. Students cannot take both STAT 315 and STAT 442. Only one of the two can be taken for credit.

Expected Class Size: 15

Grading: yes pass/fail option, no fifth course option

Distributions: (D3) (QFR)

Quantative/Formal Reasoning Notes: This is an advanced statistics class involving theory and application of statistical methods to data.

Spring 2023

LEC Section: 01 TR 9:55 am - 11:10 am Xizhen Cai

STAT 458 (F) Generalized Linear Models- Theory and Applications (QFR)

This course will explore generalized linear models (GLMs)--the extension of linear models, discussed in Stat346, to response variables that have specific non-normal distributions, such as counts and proportions. We will consider the general structure and theory of GLMs and see their use in a range of applications. As time permits, we will also examine extensions of these models for clustered data such as mixed effects models and generalized estimating equations.

Requirements/Evaluation: Weekly homework consisting of theoretical exercises and data analyses carried out in R. Short frequent quizzes and one midterm (with an in-class and take-home component). Final project and final exam.

Prerequisites: STAT 346, or permission of instructor

Enrollment Limit: 20

Enrollment Preferences: Seniors and Statistics majors

Expected Class Size: 10

Grading: yes pass/fail option, yes fifth course option

Distributions: (D3) (QFR)

Quantative/Formal Reasoning Notes: This is an intensive statistics course, involving theoretical and mathematical reasoning as well as the application of mathematical ideas to data using software.

Fall 2022

LEC Section: 01 TF 2:35 pm - 3:50 pm Elizabeth M. Upton

STAT 465 (F) Bayesian Statistics (QFR)

Interest and application of Bayesian methods have exploded in recent decades, being facilitated by recent advances in computational power. Indeed, the Bayesian approach is now recognized across scientific disciples as a modern and powerful tool. We begin with an introduction to Bayes' Theorem, the theoretical underpinning of Bayesian statistics which dates back to the 1700's, and the concepts of prior and posterior distributions, conjugacy, and closed-form Bayesian inference. Building on this, we introduce modern computational approaches to performing Bayesian inference, including Markov chain Monte Carlo (MCMC), Metropolis-Hastings sampling, and the theory underlying these simple and powerful methods, before moving on to multivariate sampling methods and methodology. Students will become comfortable with modern software tools for MCMC using a variety of applied hierarchical modeling examples, and will use R for all statistical computing. The course will culminate in an independent Bayesian research project.

Requirements/Evaluation: Homework, exams, and project

Prerequisites: STAT 346, or permission of instructor

Enrollment Limit: 30

Enrollment Preferences: Juniors and Seniors, and Statistics majors

Expected Class Size: 15

Grading: yes pass/fail option, yes fifth course option

Distributions: (D3) (QFR)

Quantative/Formal Reasoning Notes: This course mandates significant mathematical and statistical prowess.

Not offered current academic year

STAT 493 (F) Senior Thesis: Statistics

Each student carries out an individual research project under the direction of a faculty member that culminates in a thesis. See description under *The Degree with Honors in Statistics.*

Grading: yes pass/fail option, yes fifth course option

Distributions: (D3)

Fall 2022

HON Section: 01 TBA Richard D. De Veaux

STAT 494 (S) Senior Thesis: Statistics

Each student carries out an individual research project under the direction of a faculty member that culminates in a thesis. See description under *The Degree with Honors in Statistics.*

Grading: yes pass/fail option, yes fifth course option

Distributions: (D3)

Spring 2023

HON Section: 01 TBA Richard D. De Veaux

STAT 497 (F) Independent Study: Statistics

Directed independent study in Statistics.

Prerequisites: permission of department

Grading: yes pass/fail option, yes fifth course option

Distributions: (D3)

Fall 2022

IND Section: 01 TBA Richard D. De Veaux

STAT 498 (S) Independent Study: Statistics

Directed independent study in Statistics.

Prerequisites: permission of department

Grading: yes pass/fail option, yes fifth course option

Distributions: (D3)

Spring 2023

IND Section: 01 TBA Richard D. De Veaux

STAT 499 (F)(S) Statistics Colloquium

Statistics senior colloquium. Meets every week for an hour both fall and spring. Senior statistics majors must participate. This colloquium is in addition to the regular four semester-courses taken by all students.

Requirements/Evaluation: delivering a passing talk and participation throughout the year

Prerequisites: Statistics majors must take the colloquium in their senior year

Enrollment Limit: none

Enrollment Preferences: none

Expected Class Size: 25

Grading: non-graded

Distributions: (D3)

Fall 2022

SEM Section: 01 W 1:10 pm - 2:00 pm Richard D. De Veaux

Spring 2023

SEM Section: 01 W 1:10 pm - 2:00 pm Richard D. De Veaux

Winter Study ----------

STAT 19 (W) Chess, Speed Chess, Bughouse

This course will present a fast and fun introduction to chess, speed chess, and multi-player variants of classical chess. We'll begin with the rules of chess, and a study of classical openings, theory, checkmates, and endgames. These concepts will be practiced through in-class games. We will always make use of chess clocks, limiting a player's total thinking time. Chess clocks are an important part of tournament chess and speed chess, and are critically important in several chess variants we'll explore. This will open up your eyes to the high-paced, social, and extremely fun nature of recreational chess. Students will immensely enjoy learning and playing these variants, and will be surprised at how much fun chess can be. The course will culminate in a series of informal tournaments among the class.

Class Format: mornings

Requirements/Evaluation: understanding of the rules and basic principles of chess, and participation and performance in class

Prerequisites: none

Enrollment Limit: 20

Enrollment Preferences: a brief statement of your present chess knowledge and experience

Grading: pass/fail only

Materials/Lab Fee: $0

Not offered current academic year

STAT 30 (W) Senior Project: Statistics

To be taken by candidates for honors in Statistics other than by thesis route.

Class Format: senior project

Grading: pass/fail only

Not offered current academic year

STAT 31 (W) Senior Honors Thesis

Statistics senior honors thesis.

Class Format: thesis

Grading: pass/fail only

Not offered current academic year

STAT 99 (W) Indep Study: Statistics

Open to upperclass students. Students interested in doing an independent project (99) during Winter Study must make prior arrangements with a faculty sponsor. The student and professor then complete the independent study proposal form available online. The deadline is typically in late September. Proposals are reviewed by the pertinent department and the Winter Study Committee. Students will be notified if their proposal is approved prior to the Winter Study registration period.

Class Format: independent study

Grading: pass/fail only

Not offered current academic year