STATISTICS (STAT)

STAT 240 – DATA SCIENCE MODELING I
4 credits.

Introduces reproducible data management, modeling, analysis, and statistical inference through a practical, hands-on case studies approach. Topics include the use of an integrated statistical computing environment, data wrangling, the R programming language, data graphics and visualization, random variables and concepts of probability including the binomial and normal distributions, data modeling, statistical inference in one- and two- sample settings for proportions and means, simple linear regression, and report generation using R Markdown with applications to a wide variety of data to address open-ended questions.
Requisites: Satisfied Quantitative Reasoning (QR) A requirement
Course Designation: Gen Ed - Quantitative Reasoning Part B
Breadth - Natural Science
Level - Intermediate
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: No
Last Taught: Spring 2024
Learning Outcomes: 1. Transform data, possibly from multiple sources, into a form convenient for analysis
Audience: Undergraduate

2. Visualize and summarize data, generate questions/hypotheses, and address them
Audience: Undergraduate

3. Write R code using the R Studio integrated statistical computing environment to carry out data wrangling, graphical data exploration,and analysis that is reproducible
Audience: Undergraduate

4. Provide low-dimensional summaries of data that captures signal and quantify the noise; assess the adequacy of the model; andunderstand random variables and probability concepts associated with the models
Audience: Undergraduate

5. Explain what can be inferred from the data analysis and make predictions
Audience: Undergraduate

6. Use R Markdown to integrate prose, visualizations, code, interpretation, and results
Audience: Undergraduate

7. Work with other students to solve data challenges
Audience: Undergraduate

STAT 301 – INTRODUCTION TO STATISTICAL METHODS
3 credits.

Distributions, measures of central tendency, dispersion and shape, the normal distribution; experiments to compare means, standard errors, confidence intervals; effects of departure from assumption; method of least squares, regression, correlation, assumptions and limitations; basic ideas of experimental design.
Requisites: Satisfied Quantitative Reasoning (QR) A requirement. Not open to students with credit for STAT 302, 324, or 371
Course Designation: Gen Ed - Quantitative Reasoning Part B
Breadth - Natural Science
Level - Intermediate
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: No
Last Taught: Spring 2024

STAT 303 – R FOR STATISTICS I
1 credit.

An understanding of the commonly used statistical language R. Topics will include using R to manipulate data and perform exploratory data analysis.
Requisites: STAT 240, 301, 302, 312, 324, 371, MATH/STAT 310, ECON 310, GEN BUS 303, 304, 306, 307, PSYCH 210, B M E 325, I SY E 210, SOC/C&E SOC 360, graduate/professional standing, or declared in Statistics VISP
Course Designation: Breadth - Natural Science
Level - Intermediate
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: No
Last Taught: Spring 2024
Learning Outcomes: 1. Use basic R vocabulary
Audience: Undergraduate

2. Manipulate data in R
Audience: Undergraduate

3. Produce graphics and reports
Audience: Undergraduate

4. Apply statistical methods
Audience: Undergraduate

5. Run basic simulations
Audience: Undergraduate

STAT 304 – R FOR STATISTICS II
1 credit.

Provides an understanding of the commonly used statistical language R. Topics will include writing conditional expressions, loops, and functions; manipulating data matrices and arrays; extracting data from text; and making high level visualizations of data.
Requisites: STAT 303
Course Designation: Breadth - Natural Science
Level - Intermediate
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: No
Last Taught: Spring 2024

STAT 305 – R FOR STATISTICS III
1 credit.

Provides an understanding of the commonly used statistical language R. Learn to combine R with high performance computing tools to do scientific computing.
Requisites: STAT 304
Course Designation: Breadth - Natural Science
Level - Intermediate
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: No
Last Taught: Spring 2024

STAT/MATH 309 – INTRODUCTION TO PROBABILITY AND MATHEMATICAL STATISTICS I
3 credits.

Probability and combinatorial methods, discrete and continuous, univariate and multivariate distributions, expected values, moments, normal distribution and derived distributions, estimation.
Requisites: MATH 234, 376, or concurrent enrollment. Not open to students with credit for STAT/MATH 431 or STAT 311
Course Designation: Breadth - Natural Science
Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: No
Last Taught: Spring 2024
Learning Outcomes: 1. Recall the definitions of fundamental objects and concepts underlying probability theory (e.g. sample spaces and events, the axioms of probability, the notions of conditional probability and independence, random variables and their probability distributions, mathematical expectation, and the joint distribution of one or more random variables) and demonstrate understanding of their properties
Audience: Undergraduate

2. Perform important operations in probability (e.g. calculate the probabilities of events, derive the probability distributions of random variables, compute moments and the expectation of functions of random variables, calculate covariances and correlations, and obtain conditional distributions and conditional expectations) and interpret the results
Audience: Undergraduate

3. Explain the meaning of key results in probability theory that are especially important in mathematical statistics (e.g. Bayes' Theorem, probabilistic tail inequalities such as Markov's and Chebyshev's inequalities, the Law of Large Numbers, and the Central Limit Theorem)
Audience: Undergraduate

4. Identify, utilize, and understand the key properties of, probability distributions that are especially important in statistics, including discrete families of distributions (e.g. the binomial, Poisson, geometric, and negative binomial distributions) and continuous families of distributions (e.g. the uniform, exponential, gamma, and normal distributions)
Audience: Undergraduate

STAT/MATH 310 – INTRODUCTION TO PROBABILITY AND MATHEMATICAL STATISTICS II
3 credits.

Mathematical statistical inference aims at providing an understanding of likelihood's central role to statistical inference, using the language of mathematical statistics to analyze statistical procedures, and using the computer as a tool for understanding statistics. Specific topics include: samples and populations, estimation, hypothesis testing, and theoretical properties of statistical inference.
Requisites: (STAT/MATH 309, STAT 311, STAT/MATH 431, or MATH 531) and (STAT 240, STAT 301, STAT 302, STAT 324, STAT 371, or ECON 310), or graduate/professional standing
Course Designation: Breadth - Natural Science
Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: No
Last Taught: Spring 2024
Learning Outcomes: 1. Construct point estimators including maximum likelihood estimators, understand the theoretical properties of point estimation methods, and evaluate their performance
Audience: Undergraduate

2. Construct hypothesis tests including likelihood ratio tests, interpret their results, evaluate their performance, and understand the theoretical properties of hypothesis testing methods
Audience: Undergraduate

3. Construct interval estimators to quantify uncertainty, understand the theoretical properties of interval estimation methods, and interpret their results
Audience: Undergraduate

4. Mathematically derive key quantities required for statistical inference methods and be familiar with simulation-based techniques for obtaining those quantities
Audience: Undergraduate

5. Describe the Bayesian approach to inference and contrast it with the frequentist approach
Audience: Undergraduate

6. Identify and describe the assumptions underlying methods of statistical inference and explain their importance
Audience: Undergraduate

STAT 311 – INTRODUCTION TO THEORY AND METHODS OF MATHEMATICAL STATISTICS I
3 credits.

Elements of probability, important discrete distributions, acceptance sampling by attributes, sample characteristics, probability distributions and population characteristics, the normal distribution, acceptance sampling plans based on sample means and variances, sampling from the normal, the central limit theorem, point and interval estimation.
Requisites: MATH 234, 376, or concurrent enrollment or graduate/professional standing. Not open to students with credit for STAT/MATH 309 or STAT/MATH 431
Course Designation: Breadth - Natural Science
Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: No
Last Taught: Spring 2024

STAT 312 – INTRODUCTION TO THEORY AND METHODS OF MATHEMATICAL STATISTICS II
3 credits.

Unbiased estimation, maximum likelihood estimation, confidence intervals, tests of hypotheses, Neyman-Pearson lemma, likelihood ratio test, regression, analysis of variance with applications.
Requisites: STAT/MATH 309, STAT 311, STAT/MATH 431, MATH 531, or graduate/professional standing
Course Designation: Breadth - Natural Science
Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: No
Last Taught: Spring 2024
Learning Outcomes: 1. Be able to use probability theory to understand and utilize three principal tools of statistical inference: point estimators, confidence intervals, and hypothesis tests.
Audience: Undergraduate

2. Understand and be able to use standard statistical procedures for analyzing numerical data in certain contexts. These include inference for the mean based on a single random sample; comparing two means based on two random samples; Analysis of Variance; simple linear regression.
Audience: Undergraduate

3. Understand and be able to use standard statistical procedures for analyzing binary and categorical data in certain contexts. These include one- and two-sample inference for proportions, and the analysis of one- and two-way contingency tables for multi-category data.
Audience: Undergraduate

4. Identify the assumptions behind statistical procedures and understand their importance. Be able to recognize when techniques based on standard reference distributions (normal, chi-square, T- and F-distribution) are valid or not, and be able to utilize certain alternatives such as exact or nonparametric methods when they are required or preferable.
Audience: Undergraduate

STAT 324 – INTRODUCTORY APPLIED STATISTICS FOR ENGINEERS
3 credits.

Descriptive statistics, probability concepts and distributions, random variables. Hypothesis tests and confidence intervals for one- and two-sample problems. Linear regression, model checking, and inference. Analysis of variance and basic ideas in experimental design. Utilizes the R programming language.
Requisites: MATH 211, 217, or 221. Not open to students with credit for STAT 302 or 371
Course Designation: Level - Intermediate
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: No
Last Taught: Spring 2024
Learning Outcomes: 1. Articulate the basics of probability and statistics.
Audience: Undergraduate

2. Make numeric and graphical summaries of simple data
Audience: Undergraduate

3. Produce appropriate statistical analyses of simple data sets
Audience: Undergraduate

4. Design simple experiments with data that will suit basic statistical analysis
Audience: Undergraduate

5. Use R for statistical computations and graphs
Audience: Undergraduate

6. Learn additional statistical methods
Audience: Undergraduate

STAT 333 – APPLIED REGRESSION ANALYSIS
3 credits.

An introduction to regression with emphasis on the practical aspects. Topics include: straight-line model, role of assumptions, residual analysis, transformations, multiple regression (with some use of matrix notation), multicollinearity, subset selection, and a brief introduction to mixed models.
Requisites: (STAT 240, 301, 302, 312, 324, 371, ECON 310, B M E 325, or I SY E 210) and (STAT 327 or 303, or concurrent enrollment)
Course Designation: Gen Ed - Quantitative Reasoning Part B
Breadth - Natural Science
Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: No
Last Taught: Spring 2024
Learning Outcomes: 1. Correctly choose and apply common regression methods that are used in practice to analyze data, including simple and multiple linear regressions, ANOVAs/ANCOVAs, generalized linear models (e.g. logistic and Poisson) and fixed/random/mixed effect models
Audience: Undergraduate

2. Identify the underlying assumptions behind common regression methods and utilize diagnostic tools to detect violations of said assumptions
Audience: Undergraduate

3. Correctly interpret and explain results from regression methods, including interpretation of the coefficients, the p-values, R-squared, and other statistical summaries from regression
Audience: Undergraduate

4. Apply these methods to real data using the free statistical software R
Audience: Undergraduate

STAT 340 – DATA SCIENCE MODELING II
4 credits.

Teaches how to explore, model, and analyze data using R. Topics include basic probability models; the central limit theorem; Monte Carlo simulation; one- and two-sample hypothesis testing; Bayesian inference; linear and logistic regression; ANOVA; the bootstrap; random forests and cross-validation. Features the analysis of real-world data sets and the communication of findings in a clear and reproducible manner within a project setting.
Requisites: (MATH 211, 217, or 221) and STAT 240
Course Designation: Breadth - Natural Science
Level - Intermediate
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: No
Last Taught: Spring 2024
Learning Outcomes: 1. Understand and apply basic concepts in probability; combine basic probability models to build more complicated ones; and critique models and their assumptions
Audience: Undergraduate

2. Formulate statistical hypotheses for different kinds of research questions and test those hypotheses using both classical and Monte Carlo methods.
Audience: Undergraduate

3. Understand and apply principles of statistical estimation and prediction, including fitting models and assessing model quality, in the context of both linear and logistic regression.
Audience: Undergraduate

4. Apply statistical tools to answer research questions using real-world data and present these findings clearly in both spoken and written form to non-experts.
Audience: Undergraduate

STAT 349 – INTRODUCTION TO TIME SERIES
3 credits.

Autocorrelation; stationarity and non-stationarity; heteroscedasticity; dynamic models; auto-regressive and moving average models; identification and fitting; forecasting; seasonal adjustment; applications for financial time series, social sciences and environmental studies.
Requisites: STAT 333, 340, graduate/professional standing, or declared in Statistics VISP
Course Designation: Breadth - Natural Science
Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: No
Last Taught: Fall 2023

STAT 351 – INTRODUCTORY NONPARAMETRIC STATISTICS
3 credits.

Distribution free statistical procedures or methods valid under nonrestrictive assumptions: basic tools; counting methods; order statistics, ranks, empirical distribution functions; distribution free tests and associated interval and point estimators; sign test; signed rank tests; rank tests; Mann Whitney Wilcoxon procedures; Kolmogorov Smirnov tests; permutation methods; kernel density estimation; kernel and spline regression estimation; computer techniques and programs; discussion and comparison with parametric methods.
Requisites: STAT 333, 340, graduate/professional standing, or declared in Statistics VISP
Course Designation: Breadth - Natural Science
Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: No
Last Taught: Spring 2024

STAT 360 – TOPICS IN STATISTICS STUDY ABROAD
1-3 credits.

Credit is awarded for students having completed an advanced statistics course in a study abroad program for which there is no direct equivalence to the statistics department course offerings. The study abroad course must be pre-approved by the statistics department. Enrollment in a UW-Madison resident study abroad program.
Requisites: None
Course Designation: Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: Yes, unlimited number of completions

STAT 371 – INTRODUCTORY APPLIED STATISTICS FOR THE LIFE SCIENCES
3 credits.

Introduction to modern statistical practice in the life sciences, using the R programming language. Topics include: exploratory data analysis, probability and random variables; one-sample testing and confidence intervals, role of assumptions, sample size determination, two-sample inference; basic ideas in experimental design, analysis of variance, linear regression, goodness-of fit; biological applications.
Requisites: (MATH 112 and placed out of MATH 113), (MATH 113 and placed out of MATH 112), (MATH 112 and 113), MATH 114, 171, 211 or 221 or placement in MATH 221. Not open to students with credit for STAT 302 or 324
Course Designation: Gen Ed - Quantitative Reasoning Part B
Breadth - Natural Science
Level - Intermediate
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: No
Last Taught: Spring 2024
Learning Outcomes: 1. Articulate the basics of probability and statistics
Audience: Undergraduate

2. Make numeric and graphical summaries of simple data
Audience: Undergraduate

3. Produce appropriate statistical analyses of simple data sets
Audience: Undergraduate

4. Design simple experiments whose data will suit basic statistical analysis
Audience: Undergraduate

5. Use RStudio, a free statistical software package, for statistical computations and graphs
Audience: Undergraduate

6. Study and learn additional statistical methods
Audience: Undergraduate

STAT/COMP SCI 403 – INTERNSHIP COURSE IN COMP SCI AND DATA SCIENCE
1 credit.

Enables students with outside internships to earn academic credit connected to their work experience related to the Computer Sciences or Data Science programs.
Requisites: Consent of instructor
Course Designation: Level - Intermediate
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: Yes, for 3 number of completions
Last Taught: Spring 2024
Learning Outcomes: 1. Understand the challenges and opportunities in Computer Sciences and Data Science professions
Audience: Undergraduate

2. Be prepared to find, apply and interview for a job and/or additional education
Audience: Undergraduate

3. Articulate your career goals and long-term trajectory
Audience: Undergraduate

STAT 405 – DATA SCIENCE COMPUTING PROJECT
3 credits.

The development of tools necessary for collecting, managing, and analyzing large data sets. Examples of techniques and programs used include Linux, R, distributed computing, text editor(s), git/github, and other related tools. Work in teams to research, develop, write, and make presentations related to a variety of data analysis projects.
Requisites: (STAT 240 or 303) and (COMP SCI 200, 220, 300, or placement into COMP SCI 300), or graduate/professional standing
Course Designation: Breadth - Natural Science
Level - Intermediate
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: No
Last Taught: Spring 2024
Learning Outcomes: 1. Collect and manage data and write programs and documentation viatools suited to large computations including an operating system, aneditor, and a version control system.
Audience: Undergraduate

2. Run analyses too large for a laptop on cluster, grid, and/or cloudcomputing environments.
Audience: Undergraduate

3. Work in teams to research, develop, write, and make presentations ona data analysis proposal, a draft data analysis, and a revised dataanalysis.
Audience: Undergraduate

STAT 411 – AN INTRODUCTION TO SAMPLE SURVEY THEORY AND METHODS
3 credits.

An introduction to the methods used to design sample surveys and analyze the results. Topics covered include: basic tools, simple random sampling, ratio and regression estimation, stratification, systematic sampling, cluster (area) sampling, two-stage sampling, unequal probability sampling, non-sampling errors, and missing data. For illustration and clarification, examples are drawn from diverse areas of application.
Requisites: STAT 333, 340, graduate/professional standing, or declared in Statistics VISP
Course Designation: Breadth - Natural Science
Level - Intermediate
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: No
Last Taught: Fall 2023

STAT 421 – APPLIED CATEGORICAL DATA ANALYSIS
3 credits.

Analysis of multidimensional contingency tables, Poisson regression, and logistic regression, with emphasis on practical applications. Use of computer programs for such analyses. Model selection, testing goodness of fit, estimation of parameters, measures of association and methods for detecting sources of significance.
Requisites: STAT 333, 340, graduate/professional standing, or declared in Statistics VISP
Course Designation: Breadth - Natural Science
Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: No
Last Taught: Fall 2022

STAT/M E 424 – STATISTICAL EXPERIMENTAL DESIGN
3 credits.

Introduction to statistical design and analysis of experiments. Topics include: principles of randomization, blocking and replication, randomized blocking designs, Latin square designs, full factorial and fractional factorial designs and response surface methodology. Substantial focus will be devoted to engineering applications.
Requisites: STAT 240, 301, 302, 312, 324, 371, or MATH/STAT 310
Course Designation: Breadth - Natural Science
Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: No
Last Taught: Spring 2024

STAT/MATH 431 – INTRODUCTION TO THE THEORY OF PROBABILITY
3 credits.

Topics covered include axioms of probability, random variables, the most important discrete and continuous probability distributions, expectation and variance, moment generating functions, conditional probability and conditional expectations, multivariate distributions, Markov's and Chebyshev's inequalities, laws of large numbers, and the central limit theorem.
Requisites: MATH 234 or 376 or graduate/professional standing or member of the Pre-Masters Mathematics (Visiting International) Program
Course Designation: Breadth - Natural Science
Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: No
Last Taught: Spring 2024
Learning Outcomes: 1. Recall and state the formal definitions of the mathematical objects and their properties used in probability theory (e.g., probability spaces, random variables and random vectors and their probability distributions, named distributions, conditional probability, independence, linearity of expectation, etc.).
Audience: Undergraduate

2. Use such definitions to argue that a mathematical object does or does not have the condition of being a particular type or having a particular property (e.g., whether certain events or random variables are independent or not, whether a random variable has one of the named distributions, whether or not a sequence of random variables is exchangeable, etc.).
Audience: Undergraduate

3. Recall and state the standard theorems of probability theory. (e.g., Bayes' theorem, the law of large numbers, the central limit theorem, etc.), and apply these theorems to solve problems in probability theory.
Audience: Undergraduate

4. Use multiple approaches to compute and estimate probabilities and expectations (e.g., using the indicator method, using conditioning, estimating probabilities using normal or Poisson approximation etc.).
Audience: Undergraduate

5. Construct mathematical arguments related to the above definitions, properties, and theorems, including the construction of examples and counterexamples.
Audience: Undergraduate

6. Convey his or her arguments in oral and written forms using English and appropriate mathematical terminology and notation (and grammar).
Audience: Undergraduate

7. Model simple real-life situations using techniques in probability theory and calculate probabilities and expectations associated with those models.

Audience: Undergraduate

STAT 433 – DATA SCIENCE WITH R
3 credits.

Perform Data Science as an iterative (back and forth) process of four different types of activities (data collection, data wrangling, data analysis, communication). Traverse through the five requisite stances (scientist, coder, mathematician, methodologist, skeptic). Develop and hone a broad set of computational tools in R (but not the broadest) and a broad set of statistical/machine learning tools (but not the broadest). Focus on doing these with agility to make the coding "transparent" and serve the large goals of the project.
Requisites: (STAT 333 or 340) and (MATH 320, 340, 341, or 375), graduate/professional standing, or declared in Statistics VISP
Course Designation: Breadth - Natural Science
Level - Intermediate
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: No
Last Taught: Fall 2022
Learning Outcomes: 1. Describe and apply the key steps taken in most data science projects and how these steps fit into a coherent whole
Audience: Undergraduate

2. Identify how a data set can be used for a specific purpose
Audience: Undergraduate

3. Clean and analyze data
Audience: Undergraduate

4. Communicate the results of data analysis
Audience: Undergraduate

5. Develop agile and reproducible code that enables iterative development of a data science project using the tools in the tidyverse
Audience: Undergraduate

STAT 436 – STATISTICAL DATA VISUALIZATION
3 credits.

Techniques for visualization within data science workflows. Topics include data preparation; exploratory data analysis; spatial, tabular, and graph structured data; dimensionality reduction; model visualization and interpretability; interactive queries and navigation.
Requisites: (STAT 240 or 303), graduate professional/standing, or declared in Statistics VISP
Course Designation: Breadth - Natural Science
Level - Intermediate
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: No
Last Taught: Spring 2024
Learning Outcomes: 1. Apply preprocessing strategies, including outlier removal, missing data imputation, and tidying, in a way that supports downstream visualization
Audience: Undergraduate

2. Develop a vocabulary of visual encoding that support exploration of geospatial, temporal, tree-structured, and network data, and demonstrate facility implementing them using packages in the R programming language

Audience: Undergraduate

3. Design dynamic queries that support interactive visualization of heterogenous data and demonstrate facility implementing them using the shiny package in the R programming language
Audience: Undergraduate

4. Design effective visualizations to summarize the results of dimensionality reduction and clustering algorithms
Audience: Undergraduate

5. Use visual artifacts derived from complex statistical machine learning models to discuss the patterns they learn and mistakes they make
Audience: Undergraduate

6. Recognize chart junk in real-word visualizations and propose improved alternatives
Audience: Undergraduate

STAT 443 – CLASSIFICATION AND REGRESSION TREES
3 credits.

Introduction to algorithms and applications of classification and regression trees. Recursive partitioning, pruning, and cross-validation estimation of prediction error. Class priors and misclassification costs. Univariate and linear splits. Linear and kernel discriminant analysis and nearest-neighbor classification. Unbiased variable selection and importance scoring of variables. Least-squares, quantile, Poisson, logistic, and proportional hazards regression tree models. Tree ensembles. Subgroup identification of differential treatment effects. Multiple and longitudinal response variables. Missing values and multiple missing value codes. Comparisons with neural networks, support vector machines, and other methods. Bootstrap calibration and post-selection inference. Applications to business, social science, engineering, biology, medicine, and other fields.
Requisites: STAT 333, 340, graduate/professional standing, or declared in Statistics VISP
Course Designation: Breadth - Either Social Science or Natural Science
Level - Intermediate
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: No
Last Taught: Fall 2022
Learning Outcomes: 1. Know the weaknesses and limitations of traditional statistical modeling methods
Audience: Undergraduate

2. Understand the strengths and capabilities of the classification and regression tree approach
Audience: Undergraduate

3. Learn to analyze data with missing values without missing value imputation
Audience: Undergraduate

4. Learn to analyze data containing circular or periodic variables, such as angle of impact, time of day, day of week, and month of year
Audience: Undergraduate

5. Learn how to build regression tree models for least squares regression, logistic regression, Poisson regression, quantile regression, and proportional hazards regression
Audience: Undergraduate

6. Learn how to build prediction models for univariate, multivariate, longitudinal, and censored dependent variables
Audience: Undergraduate

7. Learn to use GUIDE and R software for algorithms such as AMELIA and MICE (for missing value imputation), and RPART, MOB, random Forest (for tree and forest models)
Audience: Undergraduate

STAT 451 – INTRODUCTION TO MACHINE LEARNING AND STATISTICAL PATTERN CLASSIFICATION
3 credits.

Pattern classification, regression analysis, clustering, and dimensionality reduction. For each category, covers fundamental algorithms and selections of contemporary, current state-of-the-art algorithms. Focus on evaluation of machine learning models using statistical methods. Statistical pattern classification approaches, including maximum likelihood estimation and Bayesian decision theory, algorithmic and nonparametric approaches. Practical use of machine learning algorithms using open source libraries from the Python programming ecosystem.
Requisites: MATH 320, 321, 340, 341, 375, graduate/professional standing, or declared in Statistics VISP
Course Designation: Breadth - Natural Science
Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: No
Last Taught: Spring 2024
Learning Outcomes: 1. Understand the different subfields of machine learning, such as supervised and unsupervised learning and being familiar with essential algorithms from each subfield.
Audience: Undergraduate

2. Identify whether machine learning is appropriate for solving a given problem task and which class of algorithms is best suited for real-world problem solving.
Audience: Undergraduate

3. Use statistical learning theory to combine multiple machine learning models via ensemble methods.
Audience: Undergraduate

4. Apply best-practices for statistical model evaluation, model selection and algorithm comparisons including suitable statistical hypothesis tests.
Audience: Undergraduate

5. Use contemporary programming languages and machine learning libraries for implementing machine learning algorithms such that they can be readily applied for practical problem solving.
Audience: Undergraduate

6. Connect concepts from probability theory with supervised learning by implementing models based on Bayes' theorem.
Audience: Undergraduate

STAT 453 – INTRODUCTION TO DEEP LEARNING AND GENERATIVE MODELS
3 credits.

Deep learning is a field that specializes in discovering and extracting intricate structures in large, unstructured datasets for parameterizing artificial neural networks with many layers. Since deep learning has pushed the state-of-the-art in many research and application areas, it's become indispensable for modern technology. Focuses on a understanding deep, artificial neural networks by connecting it to related concepts in statistics. Beyond covering deep learning models for predictive modeling, focus on deep generative models. Besides explanations on a mathematical and conceptual level, emphasize the practical aspects of deep learning. Open-source computing provides hands-on experience for implementing deep neural nets, working on supervised learning tasks, and applying generative models for dataset synthesis.
Requisites: MATH 320, 321, 340, 341, 375, graduate/professional standing, or declared in Statistics VISP
Course Designation: Breadth - Natural Science
Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: No
Last Taught: Spring 2024
Learning Outcomes: 1. Develop an advanced understanding of deep learning and generative models, which represent state-of-the-art approaches for predictive modeling in today’s data-driven world.
Audience: Undergraduate

2. Identify scenarios where it makes sense to deep learning for real-world problem-solving.
Audience: Undergraduate

3. Build a repertoire of different algorithms and approaches to deep learning and understanding their various strengths and weaknesses.
Audience: Undergraduate

4. Employ the Python programming language and Python’s scientific computing stack for implementing deep learning algorithms to 1) enhance the learning experience, 2) conduct research and be able to develop novel algorithms, and 3) apply deep learning to problem-solving in various fields and application areas.
Audience: Undergraduate

5. Apply both the theoretical and practical concepts taught in this class to creative, real-world problem solving and communicating the outcome professionally in form of a scientific paper and a formal oral presentation.
Audience: Undergraduate

STAT 456 – APPLIED MULTIVARIATE ANALYSIS
3 credits.

Theory and applications of multivariate statistical methods. Basic concepts and statistical reasoning which underlie the techniques of multivariate analysis. Ideas rather than derivations stressed although basic models discussed to give the student some feeling for their adequacy in particular situations. Acquaintance with and use of existing computer programs in the multivariate analysis area.
Requisites: (STAT 333 or 340) and (MATH 320, 340, 341, or 375), graduate/professional standing, or declared in Statistics VISP
Course Designation: Breadth - Natural Science
Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: No
Last Taught: Spring 2024

STAT 461 – FINANCIAL STATISTICS
3 credits.

Stochastic models and statistical methodologies are widely employed in modern finance. The models and their inferences are very important for academic research and financial practices. Financial stochastic models and their statistical inferences with applications to volatility analysis and risk management, introduction to discrete models such as binomial trees and GARCH and stochastic volatility models as well as simple continuous models like the Black-Scholes model. The focus will be on statistical inference, data analysis and risk management regarding these models.
Requisites: (STAT 333, 340, or ECON 410) and (MATH/STAT 309, STAT 311, MATH/STAT 431, or MATH 531), graduate/professional standing, or declared in Statistics VISP
Course Designation: Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: No
Last Taught: Fall 2023
Learning Outcomes: 1. Understand basic stochastic models used in pricing financial instruments.
Audience: Undergraduate

2. Develop statistical inferences for financial applications to volatility analysis and risk management.
Audience: Undergraduate

3. Use computer packages to perform statistical analysis of financial data.
Audience: Undergraduate

4. Interpret statistical analysis in the context of financial applications
Audience: Undergraduate

STAT/COMP SCI 471 – INTRODUCTION TO COMPUTATIONAL STATISTICS
3 credits.

Classical statistical procedures arise where closed-form mathematical expressions are available for various inference summaries (e.g. linear regression; analysis of variance). A major emphasis of modern statistics is the development of inference principles in cases where both more complex data structures are involved and where more elaborate computations are required. Topics from numerical linear algebra, optimization, Monte Carlo (including Markov chain Monte Carlo), and graph theory are developed, especially as they relate to statistical inference (e.g., bootstrapping, permutation, Bayesian inference, EM algorithm, multivariate analysis).
Requisites: STAT/MATH 310 and (STAT 333 or 340), graduate/professional standing, or declared in Statistics VISP
Course Designation: Breadth - Natural Science
Level - Intermediate
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: No
Last Taught: Fall 2020
Learning Outcomes: 1. Use computational tools (alongside mathematical ones) to extract information from (a) the likelihood function, the central object of interest in frequentist statistics, and (b) the posterior distribution, the central object of interest in Bayesian statistics
Audience: Undergraduate

2. Describe, understand the theoretical properties of, and implement basic algorithms for optimizing likelihood functions, including least squares and the IRLS algorithm, and the EM algorithm
Audience: Undergraduate

3. Understand random numbers and pseudorandom numbers and how to distinguish them, and utilize a variety of techniques for generating random variates from a probability distribution
Audience: Undergraduate

4. Use Monte Carlo methodology for such purposes as (a) carrying out a simulation study to study the properties of a statistical method, or (b) performing statistical inference via the bootstrap, or MCMC
Audience: Undergraduate

5. Understand the use of graphical models for representing the structure of complex joint distributions, and be able to use computational tools to extract information from graphical models
Audience: Undergraduate

STAT/COMP SCI/MATH 475 – INTRODUCTION TO COMBINATORICS
3 credits.

Problems of enumeration, distribution, and arrangement. Inclusion-exclusion principle. Generating functions and linear recurrence relations. Combinatorial identities. Graph coloring problems. Finite designs. Systems of distinct representatives and matching problems in graphs. Potential applications in the social, biological, and physical sciences. Puzzles. Problem solving.
Requisites: (MATH 320, 340, 341, or 375) or graduate/professional standing or member of the Pre-Masters Mathematics (Visiting International) Program
Course Designation: Breadth - Natural Science
Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: No
Last Taught: Spring 2024
Learning Outcomes: 1. Understand basic counting strategies, such as staged thought-experiments, inclusion/exclusion, generating functions, and recurrence relations, and apply these strategies to solve a wide variety of counting problems.
Audience: Undergraduate

2. Recall basic objects that are used in combinatorics, such as permutations and combinations of sets and multisets, binomial and multinomial coefficients, the Catalan numbers, the Stirling numbers, and the partition numbers.
Audience: Undergraduate

3. Analyze a given combinatorial problem using the standard theorems of combinatorics, such as the pigeonhole principle, the Newton binomial theorem, the multinomial theorem, the Ramsey theorem, the Dilworth theorem, the Burnside theorem, and the Polya counting theorem.
Audience: Undergraduate

4. Construct mathematical arguments related to combinatorial problems using the above definitions, properties, theorems, and counting strategies; including the construction of examples and counterexamples.
Audience: Undergraduate

5. Convey his or her arguments in oral and written form in English, using appropriate mathematical terminology, notation, and grammar.
Audience: Undergraduate

STAT 479 – SPECIAL TOPICS IN STATISTICS
1-3 credits.

Special topics of interest in undergraduate students.
Requisites: None
Course Designation: Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Repeatable for Credit: Yes, unlimited number of completions
Last Taught: Fall 2023

STAT/COMP SCI/I SY E/MATH 525 – LINEAR OPTIMIZATION
3 credits.

Introduces optimization problems whose constraints are expressed by linear inequalities. Develops geometric and algebraic insights into the structure of the problem, with an emphasis on formal proofs. Presents the theory behind the simplex method, the main algorithm used to solve linear optimization problems. Explores duality theory and theorems of the alternatives.
Requisites: MATH 320, 340, 341, 375, or 443 or graduate/professional standing or member of the Pre-Masters Mathematics (Visiting International) Program
Course Designation: Breadth - Natural Science
Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Spring 2024
Learning Outcomes: 1. Use linear programming to formulate real world decision problems.
Audience: Both Grad & Undergrad

2. Apply algorithms to solve linear programming problems and demonstrate their correctness.
Audience: Both Grad & Undergrad

3. Combine different proving techniques explored in class in an original way to show new results.
Audience: Graduate

STAT/B M I 541 – INTRODUCTION TO BIOSTATISTICS
3 credits.

Course designed for the biomedical researcher. Topics include: descriptive statistics, hypothesis testing, estimation, confidence intervals, t-tests, chi-squared tests, analysis of variance, linear regression, correlation, nonparametric tests, survival analysis and odds ratio. Biomedical applications used for each topic.
Requisites: Graduate/professional standing. Not open to students with credit for STAT 511 or POP HLTH/B M I 551
Course Designation: Level - Intermediate
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Fall 2023
Learning Outcomes: 1. Understand building blocks and fundamentals that support core themes of Biostatistics in the application of biomedicine and public health
Audience: Both Grad & Undergrad

2. Conduct basic statistical analyses of biomedical data
Audience: Both Grad & Undergrad

3. Use R for statistical computing
Audience: Both Grad & Undergrad

4. Critique methods and evidence from others' studies
Audience: Graduate

5. Collaborate effectively with biostatisticians
Audience: Graduate

STAT/B M I 542 – INTRODUCTION TO CLINICAL TRIALS I
3 credits.

Intended for biomedical researchers interested in the design and analysis of clinical trials. Topics include definition of hypotheses, measures of effectiveness, sample size, randomization, data collection and monitoring, and issues in statistical analysis.
Requisites: B M I/STAT 541
Course Designation: Level - Intermediate
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Spring 2024
Learning Outcomes: 1. Develop skills to critically review clinical trials literature
Audience: Graduate

2. Formulate focused research questions, specific aims, and key outcomes

Audience: Graduate

3. Recognize the strengths and weaknesses of alternative clinical trials designs and design components
Audience: Graduate

4. Develop related technical skills, including basic sample size calculations and survival analysis
Audience: Graduate

5. Write a clinical trial protocol with all its core components
Audience: Graduate

STAT/F&W ECOL/HORT 571 – STATISTICAL METHODS FOR BIOSCIENCE I
4 credits.

Descriptive statistics, distributions, one- and two-sample normal inference, power, one-way ANOVA, simple linear regression, categorical data, non-parametric methods; underlying assumptions and diagnostic work.
Requisites: Graduate/professional standing
Course Designation: Level - Intermediate
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Fall 2023

STAT/F&W ECOL/HORT 572 – STATISTICAL METHODS FOR BIOSCIENCE II
4 credits.

Polynomial regression, multiple regression, two-way ANOVA with and without interaction, split-plot design, subsampling, analysis of covariance, elementary sampling, introduction to bioassay.
Requisites: STAT/F&W ECOL/HORT 571
Course Designation: Level - Intermediate
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Spring 2024

STAT 575 – STATISTICAL METHODS FOR SPATIAL DATA
3 credits.

Detecting, quantifying, and modeling spatial patterns and structure in data. Variograms and covariance functions, linear predictions with uncertainty qualification, and conditional simulations. Spectral domain models and spectral densities. Spatial point processes. Contemporary applications and Gaussian process model fitting at scale.
Requisites: (STAT 333 or 340) and (MATH 320, 340, 341, or 375), graduate/professional standing, or declared in Statistics VISP
Course Designation: Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Spring 2024
Learning Outcomes: 1. Undertake exploratory analysis of spatial and spatio-temporal data with nonparametric estimators like empirical variograms.
Audience: Both Grad & Undergrad

2. Fit parametric and semi-parametric mean and covariance models to spatial and spatio-temporal data for the purposes of prediction and uncertainty quantification.
Audience: Both Grad & Undergrad

3. Understand important differences between interpolation and extrapolation and how to choose and design models more effectively for each application.
Audience: Both Grad & Undergrad

4. Use the R programming language to load, manipulate, and effectively work with spatial data in popular file formats like CSV, NetCDF, HDF5.
Audience: Undergraduate

5. Use several popular software libraries and modeling paradigms for spatial and spatio-temporal problems like estimation and prediction.
Audience: Both Grad & Undergrad

6. Understand problems of identifiability and consistency in popular modeling paradigms.
Audience: Graduate

STAT 601 – STATISTICAL METHODS I
4 credits.

Provides a thorough grounding in modern statistical methods. The specific learning outcomes for the course are to understand data collection in context (how/why data were collected, key questions under study); explore data by effective graphical and numerical summaries; understand probability concepts and models as tools for studying random phenomena and for statistical inference; analyze data using appropriate, modern statistical models, methods, and software; understand the statistical concepts underlying methods; develop the ability to interpret results and critically evaluate the methods used; communicate data analysis and key findings in context.
Requisites: Graduate/professional standing or declared in Statistics VISP
Course Designation: Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Spring 2024

STAT 602 – STATISTICAL METHODS II
4 credits.

Provides a thorough grounding in modern statistical methods. The specific learning outcomes for the course are to understand data collection in context (how/why data were collected, key questions under study); explore data by effective graphical and numerical summaries; understand probability concepts and models as tools for studying random phenomena and for statistical inference; analyze data using appropriate, modern statistical models, methods, and software; understand the statistical concepts underlying methods; develop the ability to interpret results and critically evaluate the methods used; communicate data analysis and key findings in context.
Requisites: STAT 601
Repeatable for Credit: No
Last Taught: Spring 2022

STAT 605 – DATA SCIENCE COMPUTING PROJECT
3 credits.

The development of tools necessary for collecting, managing, and analyzing large data sets. Examples of techniques and programs utilized include Linux, R, distributed computing, powerful editor(s), git/github, and other related tools. Work in the class will be done in teams to research, develop, write, and make presentations related to a variety of data analysis projects.
Requisites: Declared in Statistics MS or Statistics VISP
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Spring 2024
Learning Outcomes: 1. Collect and manage data and write programs and documentation via tools suited to large computations. Use the Linux operating system and write shell scripts. Use an editor to write and manage local and remote files. Use the git/github version control system to track changes and manage collaboration.
Audience: Graduate

2. Be able to use Linux, R, and the Slurm job scheduler to run tens of parallel jobs on the Statistics High Performance Computing (HPC) Cluster. Use Linux, R, and distributed high-throughput computing via HTCondor to run thousands of parallel jobs at UW's Center for High-Throughput Computing (CHTC).
Audience: Graduate

3. Work in teams to research, develop, write, and make three presentations including one on a data analysis proposal consisting of data, a question, and a suggested analysis; one on a draft data analysis; and one on a revised data analysis.
Audience: Graduate

STAT 606 – COMPUTING IN DATA SCIENCE AND STATISTICS
3 credits.

A survey of some of the tools and frameworks that are currently popular among data scientists and statisticians working in both academia and industry. Begins with an accelerated introduction to the Python programming language and brief introductions to object-oriented and functional programming. Covers some of the scientific computing platforms available in Python, including tools for numerical and scientific computing; training basic machine learning models; and data visualization. Discusses collecting data from the web both by scraping and using APIs. Concludes with a brief survey of distributed computing platforms, focusing on the MapReduce framework.
Requisites: Declared in Statistics MS or Statistics VISP (undergraduate)
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Spring 2024
Learning Outcomes: 1. Understand and apply the basics of the Python programming language and basic programming patterns in both the object-oriented and functional programming frameworks
Audience: Graduate

2. Collect and clean data from a variety of data sources including markup languages from the web, databases, and by using APIs
Audience: Graduate

3. Understand the MapReduce framework and apply it to large-scale data sets in a distributed environment using modern cloud computing platforms

Audience: Graduate

4. Use numerical and scientific computing libraries to build and fit statistical models on large datasets
Audience: Graduate

STAT 609 – MATHEMATICAL STATISTICS I
3 credits.

Review of probability, random variables and vectors and their distributions, moments and inequalities, generating functions, transformations of random variables, sampling and distribution theory, convergence concepts for sequences of random variables, laws of large numbers, central limit and other limit theorems.
Requisites: Graduate/professional standing or declared in Statistics VISP
Course Designation: Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: Yes, unlimited number of completions
Last Taught: Fall 2023

STAT 610 – INTRODUCTION TO STATISTICAL INFERENCE
4 credits.

Conditioning, distribution theory, approximation to distributions, modes of convergence, limit theorems, statistical models, parameter estimation, comparision of estimators, confidence sets, theory of hypothesis tests, introduction to Bayesian inference and nonparametric estimation.
Requisites: Graduate/professional standing or declared in Statistics VISP
Course Designation: Breadth - Physical Sci. Counts toward the Natural Sci req
Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Spring 2024

STAT 611 – STATISTICAL MODELS FOR DATA SCIENCE
3 credits.

Probability, random variables and their distributions, joint and conditional distributions, moments and inequalities, generating functions, transformations of random variables, sampling and distribution theory, convergence concepts and limit theorems for sequences of random variables.
Requisites: Declared in Data Science MS or Data Engineering MS
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Fall 2023
Learning Outcomes: 1. Describe the foundations of probability theory, including: the axioms of probability, the concepts of sample space and probability measure, events and their probabilities, and the notions of conditional probabilityand independence
Audience: Graduate

2. Explain the concepts of random variables and probability distributions, and calculate or otherwise utilize key mathematical objects and results related to them, including: random number generation, expected values and moments, moment-based probabilistic inequalities, moment generating functions, joint probability distributions of multiple random variables
Audience: Graduate

3. Identify the most important probability distributions used in Statistics, cite their properties, create computer visualizations of them, and simulate random variates from these distributions. This includes the normal distribution, the binomial distribution, the exponential distribution, the Poisson distribution, and more
Audience: Graduate

4. Apply, and explain the significance of, key results related to the limits of sequences of random variables, including laws of large numbers, the Central Limit Theorem, and the delta method
Audience: Graduate

5. Grasp important elements of sampling theory, including: random samples and how to generate them; statistics, such as the sample mean and sample variance, and how to derive or simulate their sampling distributions; and the distributions that arise from a normal random sample, namely the chi-square,T, and F distributions.
Audience: Graduate

STAT 612 – STATISTICAL INFERENCE FOR DATA SCIENCE
3 credits.

Statistical models, methods and theory for parameter estimation, Bayesian approach to parameter estimation, methods and theory for hypothesis tests, confidence sets, two-sample testing and ANOVA, categorical data analysis, linear regression.
Requisites: STAT 611
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Spring 2024
Learning Outcomes: 1. Construct point estimators including maximum likelihood estimators, understand the theoretical properties of point estimation methods, evaluate their performance using mathematical derivations and simulation-based techniques, and identify optimal point estimators
Audience: Graduate

2. Describe the Bayesian approach to point estimation and contrast it with the frequentist approach
Audience: Graduate

3. Construct and evaluate hypothesis tests (such as likelihood ratio tests) using mathematical derivations or simulation-based techniques, interpret their results, understand theoretical properties of hypothesis testing methods, and identify optimal hypothesis tests
Audience: Graduate

4. Construct and evaluate interval estimators using mathematical and simulation-based techniques, understand the theoretical properties of interval estimation methods, and interpret their results
Audience: Graduate

5. Identify and describe the assumptions underlying methods of statistical inference and explain their importance
Audience: Graduate

6. Fit models and carry out statistical inference in classical situations: namely: the comparison of two or more samples (ANOVA), the analysis of categorical data, linear regression, generalized linear models, and random and mixed effects models
Audience: Graduate

STAT 613 – STATISTICAL METHODS FOR DATA SCIENCE
3 credits.

Provides a thorough grounding in modern statistical methods. Introduces statistical techniques and methods of data analysis, including data description, linear regression models, diagnostic tools, prediction and model selection, and experimental design.
Requisites: Declared in Data Science MS or Data Engineering MS
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Fall 2023
Learning Outcomes: 1. Strategize and select a regression equation, examine residuals, transform data, recognize biases due to excluded variables and measurement error
Audience: Graduate

2. Conduct general linear modeling for exponential family data and specifically models for binary, count, and categorical data, perform model fitting and inference
Audience: Graduate

3. Develop the concepts and relevant methodology and ability to design and analyze experiments
Audience: Graduate

4. Use and interpret computer package for regression programs
Audience: Graduate

5. Present clear data structure and analysis in the context of data drawn from real-world applications
Audience: Graduate

STAT 615 – STATISTICAL LEARNING
3 credits.

The development of a variety of mathematical theories and statistical concepts (1) to understand the properties of those models and methods used for the purpose of prediction from data or decision making from data, and (2) to criticize such models, methods and their consequences. Specifically, the theories and tools that will be developed will include complexity theory, Hilbert spaces, Gaussian processes, Variational Analysis, and concentration inequalities.
Requisites: Declared in Statistics: Statistics and Data Science MS, Data Science MS, Data Engineering MS, or Statistics VISP
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Spring 2024
Learning Outcomes: 1. Demonstrate understanding of statistical theories, methodologies, and applications as tools in scientific inquiries.
Audience: Graduate

2. Be able to select and utilize the most appropriate statistical methodologies and practices.
Audience: Graduate

3. Be able to synthesize information pertaining to questions in empirical studies.
Audience: Graduate

4. Be able to communicate data concepts and analysis results clearly.
Audience: Graduate

5. Be able to recognize and apply principles of ethical and professional conduct.
Audience: Graduate

6. Demonstrate knowledge of theoretical properties of many procedures used in machine learning for the purposes of classification, regression, and beyond.
Audience: Graduate

7. Demonstrate knowledge of classical and modern notions in statisticallearning theory including concentration inequalities, measures of statistical complexity, kernel methods for learning, Gaussian processes, and basics of variational inference.
Audience: Graduate

STAT/B M I 620 – STATISTICS IN HUMAN GENETICS
3 credits.

Provides a comprehensive survey of statistical methods in human genetics research. Covered topics include linkage analysis, genome-wide association study, rare variant association analysis, meta-analysis, genome and variant annotation, heritability estimation, multi-trait modeling techniques, multi-omic data integration, and genetic risk prediction.
Requisites: STAT 333, 340, or graduate/professional standing
Course Designation: Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Learning Outcomes: 1. Recognize problems in human genetics that are appropriate for statistical modeling
Audience: Both Grad & Undergrad

2. Identify appropriate statistical procedures and computational algorithms for different tasks
Audience: Both Grad & Undergrad

3. Gain practical experience in applying a select set of statistical methods on real data and evaluate its outputs
Audience: Both Grad & Undergrad

4. Evaluate the strengths and weaknesses of different statistical and computational approaches designed for a specific biological problem
Audience: Graduate

STAT 627 – PROFESSIONAL SKILLS IN DATA SCIENCE
1-3 credits.

Covers important aspects of professional development in statistics, including skills with internet tools, sophisticated use of statistical languages (such as R) and other emerging topics.
Requisites: Graduate/professional standing
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: Yes, unlimited number of completions
Last Taught: Spring 2024

STAT 628 – DATA SCIENCE PRACTICUM
1-3 credits.

Provides an understanding of and experience with turning statistics concepts into practice through data science practicums inspired by realistic projects. Combine theory and methods expertise with communications skills to translate from a vaguely stated project description and complex data set into a concisely summarized analysis, including both written and graphical interpretation that can be used by decision makers in an organization.
Requisites: Graduate/professional standing
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: Yes, unlimited number of completions
Last Taught: Spring 2024
Learning Outcomes: 1. Frame real-world data problems into testable and solvable statistical problems.
Audience: Graduate

2. Develop analysis/solutions that are fast, scalable, and robust
Audience: Graduate

3. Communicate solutions concisely and clearly, creating understandable and accurate visual/tabular summaries of the data analysis, and responding to audience questions clearly
Audience: Graduate

STAT/I SY E/MATH/OTM 632 – INTRODUCTION TO STOCHASTIC PROCESSES
3 credits.

Topics include discrete-time Markov chains, Poisson point processes, continuous-time Markov chains, and renewal processes. Applications to queueing, branching, and other models in science, engineering and business.
Requisites: (STAT/MATH 431, 309, STAT 311 or MATH 531) and (MATH 320, 340, 341, 375, 421 or 531) or graduate/professional standing or member of the Pre-Masters Mathematics (Visiting International) Program
Course Designation: Breadth - Natural Science
Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Spring 2024
Learning Outcomes: 1. Recall and state the formal definitions of the mathematical objects and their properties for stochastic processes (e.g., discrete space Markov chains, Poisson processes, renewal processes, branching processes, etc.).
Audience: Both Grad & Undergrad

2. Use such definitions to argue that a mathematical object does or does not have the condition of being a particular type or having a particular property (e.g., irreducibility, aperiodicity, recurrence, transience, the Markov property, etc.).
Audience: Both Grad & Undergrad

3. Recall and state the standard theorems of stochastic processes. (e.g., laws of large numbers for Markov chains, existence of limiting/stationary distributions, law of large numbers for renewal processes, etc.) and recall the arguments for these theorems and the underlying logic of their proofs.

Audience: Both Grad & Undergrad

4. Construct mathematical arguments related to the above definitions, properties, and theorems, including the construction of examples and counterexamples.
Audience: Both Grad & Undergrad

5. Convey arguments in oral and written forms using English and appropriate mathematical terminology, notation and grammar.
Audience: Both Grad & Undergrad

6. Model simple real life situations by means of discrete-space stochastic processes and calculate probabilities associated with those processes.
Audience: Both Grad & Undergrad

7. Identify applications of course content in current areas of research.
Audience: Graduate

STAT/B M I 641 – STATISTICAL METHODS FOR CLINICAL TRIALS
3 credits.

Statistical issues in the design of clinical trials, basic survival analysis, data collection and sequential monitoring.
Requisites: STAT/MATH 310 or graduate/professional standing
Course Designation: Breadth - Natural Science
Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Fall 2023

STAT/B M I 642 – STATISTICAL METHODS FOR EPIDEMIOLOGY
3 credits.

Methods for analysis of case-control, cross sectional, and cohort studies. Covers epidemiologic study design, measures of association, rates, classical contingency table methods, and logistic and Poisson regression.
Requisites: STAT/MATH 310 or graduate/professional standing
Course Designation: Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Spring 2023
Learning Outcomes: 1. Summarize key concepts of statistical methods in epidemiology study: study design, association, causation
Audience: Both Grad & Undergrad

2. Build parametric or semiparametric model for analyzing categorical data and survival data
Audience: Both Grad & Undergrad

3. Utilize model design tools for model performance assessment
Audience: Both Grad & Undergrad

4. Build semiparametric model for analyzing categorical data and survival data
Audience: Graduate

STAT/B M I 643 – CLINICAL TRIAL DESIGN, IMPLEMENTATION, AND ANALYSIS
3 credits.

Gain an understanding of fundamental elements of clinical trials (such as objectives, endpoints, surrogate endpoints, and statistical decisions) and statistical design considerations (such as randomization and blinding). Designs of clinical trials for Phase I, II, and III studies including single-arm, two-arm, and drug combination trials. Introduction to adaptive designs for precision medicine and master protocol designs such as umbrella trials and basket trials.
Requisites: STAT 609, 610, B M I/STAT 641, or graduate/professional standing
Course Designation: Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Learning Outcomes: 1. Summarize the objectives of clinical trials and key statistical design components
Audience: Both Grad & Undergrad

2. Design the clinical trials and investigate the operating characteristics of the design to implement clinical trials
Audience: Both Grad & Undergrad

3. Write the protocol section of statistical considerations and communicate the design of clinical trials to both statisticians and clinicians
Audience: Both Grad & Undergrad

4. Build sequential and adaptive methods for clinical trials
Audience: Graduate

STAT 678 – INTRODUCTION TO STATISTICAL CONSULTING
3 credits.

Develop statistical consulting skills to be able to communicate design and analysis to non-technical research collaborators. Provides a supportive environment to experiment with statistical consulting in practice, which will sometimes be uncomfortable and strange. Consulting problems typically do not have a "right" answer, and mistakes are encouraged. Take risks in sharing developing ideas in class. Connections with external organizations, such as the private sector and government agencies, will be made through possible internship experiences.
Requisites: Declared in Statistics MS
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Spring 2024
Learning Outcomes: 1. summarize findings with graphs and non-technical writing
Audience: Graduate

2. ask questions to reveal problem design and study details
Audience: Graduate

3. present design, analysis approach and findings orally in plain language
Audience: Graduate

4. manage time and project workflow effectively
Audience: Graduate

5. contribute as an active member of a research team
Audience: Graduate

STAT 679 – SPECIAL TOPICS IN STATISTICS
1-3 credits.

Special topics in statistics at the master's level. Subject matter varies.
Requisites: Graduate/professional standing or declared in Statistics VISP
Course Designation: Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: Yes, unlimited number of completions
Last Taught: Spring 2024

STAT 681 – SENIOR HONORS THESIS
3 credits.

Mentored individual study for students writing honors thesis, as arranged with a faculty member.
Requisites: Consent of instructor
Course Designation: Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Honors - Honors Only Courses (H)
Repeatable for Credit: No
Last Taught: Fall 2023

STAT 682 – SENIOR HONORS THESIS
3 credits.

Mentored individual study for students writing honors thesis, as arranged with a faculty member.
Requisites: Consent of instructor
Course Designation: Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Honors - Honors Only Courses (H)
Repeatable for Credit: No
Last Taught: Spring 2024

STAT 698 – DIRECTED STUDY
1-6 credits.

Directed study projects as arranged with a faculty member.
Requisites: Consent of instructor
Course Designation: Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: Yes, unlimited number of completions
Last Taught: Spring 2024

STAT 699 – DIRECTED STUDY
1-6 credits.

Directed study projects as arranged with a faculty member.
Requisites: Consent of instructor
Course Designation: Level - Advanced
L&S Credit - Counts as Liberal Arts and Science credit in L&S
Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: Yes, unlimited number of completions
Last Taught: Spring 2024

STAT 701 – APPLIED TIME SERIES ANALYSIS, FORECASTING AND CONTROL I
3 credits.

Theory and application of discrete time series models illustrated with forecasting problems. Principles of iterative model building. Representation of dynamic relations by difference equations. Autoregressive integrated Moving Average models. Identification, fitting, diagnostic checking of models. Seasonal model application to forecasting in business, economics, ecology, and engineering used at each stage, which the student analyzes using computer programs which have been specially written and extensively tested.
Requisites: Graduate/professional standing
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Fall 2023

STAT/MATH 709 – MATHEMATICAL STATISTICS
4 credits.

Introduction to measure theoretic probability; derivation and transformation of probability distributions; generating functions and characteristic functions; conditional expectation, sufficiency, and unbiased estimation; methods of large sample theory including laws of large numbers and central limit theorems; order statistics.
Requisites: Graduate/professional standing
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Fall 2023
Learning Outcomes: 1. Understand foundations of mathematical statistics, including key notations, important concepts, and basic definitions
Audience: Graduate

2. Develop proof-based theoretical skills for analyzing statistical problems
Audience: Graduate

3. Familiarize various theoretical tools relevant to statistical research, including modern probability theory, optimization, and information theory
Audience: Graduate

4. Prepare for statistical research by learning recent development in high-dimensional statistics and shrinkage estimation
Audience: Graduate

STAT/MATH 710 – MATHEMATICAL STATISTICS
4 credits.

Estimation, efficiency, Neyman-Pearson theory of hypothesis testing, confidence regions, decision theory, analysis of variance, and distribution of quadratic forms.
Requisites: STAT/MATH 709
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Spring 2024

STAT/COMP SCI/I SY E/MATH 726 – NONLINEAR OPTIMIZATION I
3 credits.

Theory and algorithms for nonlinear optimization, focusing on unconstrained optimization. Line-search and trust-region methods; quasi-Newton methods; conjugate-gradient and limited-memory methods for large-scale problems; derivative-free optimization; algorithms for least-squares problems and nonlinear equations; gradient projection algorithms for bound-constrained problems; and simple penalty methods for nonlinearly constrained optimization. Students are strongly encouraged to have knowledge of linear algebra and familiarity with basic mathematical analysis.
Requisites: Graduate/professional standing
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Spring 2024

STAT/B M I 727 – THEORY AND METHODS OF LONGITUDINAL DATA ANALYSIS
3 credits.

Theory and methods of fundamental statistical models for the analysis of longitudinal data, including repeated measures analysis of variance, linear mixed models, generalized linear mixed models, and generalized estimating equations. Introduction of how to implement these methods in statistical softwares such as in R and/or SAS, within the context of appropriate statistical models and carry out and interpret analyses.
Requisites: STAT 610
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Fall 2023
Learning Outcomes: 1. Analyze longitudinal data in a variety of settings and with a variety of outcome variables
Audience: Graduate

2. Apply statistical methods in fitting longitudinal data models for addressing scientific questions
Audience: Graduate

3. Perform longitudinal data analyses in statistical softwares such as R and/or SAS
Audience: Graduate

4. Interpret and communicate the scientific meanings of the results to both statisticians and non-statisticians (such as clinicians and scientists)
Audience: Graduate

STAT 732 – LARGE SAMPLE THEORY OF STATISTICAL INFERENCE
3 credits.

Stochastic modes of convergence. Asymptotic theory of normed sums of random variables with applications to asymptotic normality of estimators. Methods for deriving limit distributions of nonlinear statistics. Asymptotic relative efficiencies. Asymptotic confidence regions and tests of hypotheses. Models of non-identically distributed or dependent random variables.
Requisites: STAT 610 or MATH/STAT 709
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Spring 2020

STAT/MATH 733 – THEORY OF PROBABILITY I
3 credits.

An introduction to measure theoretic probability and stochastic processes.Topics include foundations, independence, zero-one laws, laws of large numbers, convergence in distribution, characteristic functions, central limit theorems, random walks, conditional expectations. Familiarity with basic measure theory (e.g. MATH 629 or 721) or concurrent registration in MATH 721 is strongly recommended.
Requisites: Graduate/professional standing or member of the Pre-Masters Mathematics (Visiting International) Program
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: Yes, unlimited number of completions
Last Taught: Fall 2023

STAT/MATH 734 – THEORY OF PROBABILITY II
3 credits.

Possible topics include martingales, weak convergence of measures, introduction to Brownian motion.
Requisites: Graduate/professional standing
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: Yes, unlimited number of completions
Last Taught: Spring 2024

STAT/B M I 741 – SURVIVAL ANALYSIS THEORY AND METHODS
3 credits.

Theory and practice of analytic methods for censored survival data, including nonparametric and parametric methods, the proportional hazards regression model, and a review of current topics in survival analysis.
Requisites: Graduate/professional standing
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Spring 2024
Learning Outcomes: 1. Summarize the features of censored data and their implications in drawing inference
Audience: Graduate

2. Implement proper non- and semi-parametric methods for analysis of various types of data
Audience: Graduate

3. Recognize and check the assumptions needed for estimation and inference
Audience: Graduate

4. Implement the inference procedures to solve real-world problems using statistical packages such as R (or SAS)
Audience: Graduate

5. Interpret and present the analytic results in a clear and coherent way to answer substantive questions
Audience: Graduate

STAT 760 – MULTIVARIATE ANALYSIS I
3 credits.

Multivariate normal distribution, estimation of mean and covariance matrix; Wishart distribution; distribution of partial and multiple correlation coefficients; Hotelling's T-squared, principal components.
Requisites: STAT 610 or MATH/STAT 710
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Fall 2023

STAT 761 – DECISION TREES FOR MULTIVARIATE ANALYSIS
3 credits.

Tree construction, including finding splits, tree-pruning and error estimation. Categorical predictor variables, missing or censored data, prior class-probabilities, and unequal misclassification costs. Selection bias. Comparison with other statistics and machine-learning methods. Extensions to piecewise linear and non-least squares regression models.
Requisites: Graduate/professional standing
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Spring 2024

STAT/B M I 768 – STATISTICAL METHODS FOR MEDICAL IMAGE ANALYSIS
3 credits.

Introduce key statistical methods and concepts for analyzing various medical images. Analyze publicly available and student/instructor supplied imaging data using the most up-to-date methods and tools.
Requisites: Graduate/professional standing
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Spring 2023
Learning Outcomes: 1. Implement the key concepts of biomedical image processing and analysis
Audience: Graduate

2. Describe the key concepts of statistical inference procedures for single and multiple images
Audience: Graduate

3. Apply scalable computation in breaking bigger imaging problems into smaller computable problems
Audience: Graduate

4. Describe functional data analysis (FA), geometric data analysis (GDA) and topological data analysis (TDA) methods in analyzing biomedical images
Audience: Graduate

STAT 771 – STATISTICAL COMPUTING
3 credits.

The design of statistical software including special techniques for probability distributions, methods of simulation of random processes, numerical methods for linear models and multivariate analysis, and methods for nonlinear models.
Requisites: Graduate/professional standing
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Fall 2023

STAT 772 – LINEAR RANDOMIZED ALGORITHMS FOR DATA SCIENCE
3 credits.

Introduce new algorithms that leverage randomization to address the scale, speed, and sensitivity needs of modern data science. Develop the mathematical foundations of such randomized algorithms. Criticize these algorithms through the lens of computational resource utilization. Implement these algorithms to address linear problems in data science.
Requisites: Consent of instructor
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Spring 2024
Learning Outcomes: 1. Apply common (constructive and inductive) thinking and reasoning patterns used in randomized algorithms for deriving solutions to linear problems arising in data science
Audience: Graduate

2. Transform constructive mathematical reasoning patterns into numerical algorithms, and criticize the computational resource utilization of these algorithms
Audience: Graduate

3. Implement numerical algorithms using the Julia programming language, and become familiar with assistive tools such as debuggers and profilers
Audience: Graduate

4. Create and implement rigorous numerical experiments to test and compare different algorithms and implementations, and draw appropriate conclusions from the results of these experiments
Audience: Graduate

STAT/ECON/GEN BUS 775 – INTRODUCTION TO BAYESIAN DECISION AND CONTROL I
3 credits.

Common sampling models in business and economic problems, information from data, likelihood function of parameters, choices of models, Bayes' Theorem, subjective basis for probability, sequential nature of Bayesian inference, prior and posterior distributions of parameters in binomial, poisson, exponential and normal populations, comparison of two normal distributions, predictive distributions, decision theory, utility, risk aversion, extensive form of analysis, two-action problems, point estimation, best population problems, economics of sampling.
Requisites: STAT 609 or STAT/MATH 709
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Spring 2024
Learning Outcomes: 1. Justify the use of probability for coherent uncertainty quantification
Audience: Graduate

2. Explain how Bayesian updating occurs in conjugate models and hierarchical models
Audience: Graduate

3. Compare and contrast the conceptual and practical benefits and challenges of different posterior approximation strategies like MCMC and variational inference
Audience: Graduate

4. Implement posterior approximation algorithms in modern statistical and probabilistic programming languages such as R or Stan
Audience: Graduate

5. Specify, fit, criticize, and revise Bayesian models in practice
Audience: Graduate

STAT 780 – INTRODUCTION TO QUANTUM DATA SCIENCE
3 credits.

Quantum computation issues, including probability, statistics, sensing, information, machine learning, and applying data science to quantum information science.
Requisites: STAT 601, (STAT 609 and 610), or (STAT 611 and 612)
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Learning Outcomes: 1. Demonstrate understanding of quantum probability, including axioms, observable, outcome, expectation, and distribution.
Audience: Graduate

2. Master quantum statistics methods such as likelihood, information inequality, quantum hypothesis test, and quantum tomography (quantum sensing).
Audience: Graduate

3. Comprehend concept of quantum computation and quantum information, including qubit and its properties, quantum entropy, and quantum cryptography.
Audience: Graduate

4. Understand essential elements of quantum algorithms and quantum machine learning.
Audience: Graduate

5. Understand statistics in quantum computational advantage (supremacy) studies.
Audience: Graduate

STAT 801 – ADVANCED FINANCIAL STATISTICS
3 credits.

Statistical theory and methodology for modern financial data. Topics include financial stochastic models based on time series and stochastic calculus, modern statistical inference, and statistical learning for financial data as well as their applications to financial problems.
Requisites: STAT 601 or 701
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Fall 2020

STAT/MATH 803 – EXPERIMENTAL DESIGN I
3 credits.

Summary of matrix algebra required, theory of estimable functions, incomplete blocks, balanced incomplete block designs, partially balanced incomplete block designs.
Requisites: Graduate/professional standing or member of the Pre-Masters Mathematics (Visiting International) Program
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Fall 2020

STAT 809 – NON PARAMETRIC STATISTICS
3 credits.

Statistical procedures valid under unrestrictive assumptions; sign test; confidence intervals; efficiency comparisons; signed rank procedures; Walsh sums; point estimators; two sample rank tests; zeros, ties, and other problems of discrete data; order statistics; Winsorized and truncated point estimators and connection with gross error models; permutation procedures; combinatorial problems, and computer applications.
Requisites: STAT 610 or MATH/STAT 710
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Spring 2019

STAT 811 – SAMPLE SURVEY THEORY AND METHOD
3 credits.

Simple random sampling; systematic sampling; probability sampling; stratified sampling; subsampling with units of equal and unequal size; double sampling; multi-stage and multi-phase sampling; ratio and regression estimates; model-based and model-assisted approaches; variance estimation; non-response.
Requisites: STAT 610 or MATH/STAT 710
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Fall 2017

STAT/B M I 828 – SEMIPARAMETRIC METHODS IN DATA SCIENCE
3 credits.

Review of statistical convergence modes, M-estimation, and basics of Hilbert space. Introduction of how to derive the nuisance tangent space, its complement, and the corresponding efficient influence function, from the geometric perspective of semiparametric models. Introduction of how to estimate nuisance functions using machine learning methods, and their implementations in R and/or Python. Introduction of a variety of semiparametric models in missing data analysis, causal inference, dimension reduction, precision medicine, semi-supervised learning, transfer learning and domain adaptation.
Requisites: Graduate/professional standing
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Learning Outcomes: 1. Derive the nuisance tangent space, its complement, and the corresponding efficient influence function in semiparametric models
Audience: Graduate

2. Apply a variety of semiparametric methods and models in applications ranging from biomedical studies to social sciences
Audience: Graduate

3. Perform machine learning algorithms for estimating nuisance functions in software such as R and/or Python
Audience: Graduate

STAT/MATH 833 – TOPICS IN THE THEORY OF PROBABILITY
3 credits.

Advanced topics in probability and stochastic processes.
Requisites: Graduate/professional standing or member of the Pre-Masters Mathematics (Visiting International) Program
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: Yes, unlimited number of completions
Last Taught: Fall 2023

STAT 834 – EMPIRICAL PROCESSES AND SEMIPARAMETRIC INFERENCE
1-3 credits.

Empirical process methods in statistics; semiparametric models; stochastic convergence in metric spaces; Glivenko-Cantelli and Donsker theorems; entropy calculations; bootstrapped empirical processes; functional delta method; Z-estimators; M-estimators; rates of convergence; semiparametric efficiency; semiparametric estimating equations; nonparametric maximum likelihood.
Requisites: STAT 610 or MATH/STAT 710
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Spring 2019

STAT 840 – STATISTICAL MODEL BUILDING AND LEARNING
3 credits.

Theory of reproducing kernel Hilbert spaces in statistical model building; bounded linear functionals and representer theory; smoothing splines; ANOVA spines; degees of freedom for signal and the bias-variance tradeoff; Bayesian confidence intervals; model selection.
Requisites: STAT 610 or MATH/STAT 710
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Fall 2015

STAT 841 – NONPARAMETRIC STATISTICS AND MACHINE LEARNING METHODS
3 credits.

Statistical function estimation and classification; reproducing kernel machines, support vector machines; high dimensional model selection and estimation; Bayesian, empirical Bayesian interpretation of nonparametric learning methods; log density ANOVA and graphical models; tree ensemble methods including bagging, boosting, and random forest.
Requisites: STAT 610 or MATH/STAT 710
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Fall 2019

STAT 849 – THEORY AND APPLICATION OF REGRESSION AND ANALYSIS OF VARIANCE I
3 credits.

Theory and applications of the general linear model; graphical methods; simultaneous inference; regression diagnostics; analysis of variance of fixed, random and mixed effects models; ANCOVA: violations of assumptions.
Requisites: Graduate/professional standing
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Fall 2023

STAT 850 – THEORY AND APPLICATION OF REGRESSION AND ANALYSIS OF VARIANCE II
3 credits.

Theory and applications of the general linear model; graphical methods; simultaneous inference; regression diagnostics; analysis of variance of fixed, random and mixed effects models; ANCOVA: violations of assumptions.
Requisites: STAT 849
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Spring 2024

STAT 860 – ESTIMATION OF FUNCTIONS FROM DATA
3 credits.

Statistical and appoximation theoretic methods of estimating functions and values of functionals from experimental data; experimental design and data analysis problems that arise as problems in approximation theory; convergence theorems; ill-posed inverse problems; Banach and Hilbert space penalty functionals.
Requisites: STAT 610 or MATH/STAT 710
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Fall 2016

STAT/COMP SCI/E C E 861 – THEORETICAL FOUNDATIONS OF MACHINE LEARNING
3 credits.

Advanced mathematical theory and methods of machine learning. Statistical learning theory, Vapnik-Chevronenkis Theory, model selection, high-dimensional models, nonparametric methods, probabilistic analysis, optimization, learning paradigms.
Requisites: E C E/COMP SCI 761 or E C E 830
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Fall 2023

STAT/B M I 877 – STATISTICAL METHODS FOR MOLECULAR BIOLOGY
3 credits.

Statistical and computational methods in statistical genomics for human and experimental populations. Review methods for quality control, experimental design, clustering, network analysis, and other downstream analysis of next-generation sequencing studies along with methods for genome wide association studies.
Requisites: Graduate/professional standing
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Spring 2024
Learning Outcomes: 1. Understand the statistical and computational background underlying many state-of-the-art techniques for the pre-processing and analysis of high-throughput genomics datasets
Audience: Graduate

2. Identify the appropriateness and limitations of such methods in a variety of settings.
Audience: Graduate

3. Discuss scientific problems and identify the statistical and computational aspects embedded in the processing and analysis of genomics datasets.
Audience: Graduate

4. Become proficient in select software packages commonly used in analysis of next-generation sequencing data.
Audience: Graduate

STAT/E C E/MATH 888 – TOPICS IN MATHEMATICAL DATA SCIENCE
1-3 credits.

Advanced topics in the mathematical foundations of data science
Requisites: Graduate/professional standing or member of the Pre-Masters Mathematics (Visiting International) Program
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: Yes, unlimited number of completions
Last Taught: Fall 2023
Learning Outcomes: 1. Apply advanced mathematical concepts to solve a variety of data science problems
Audience: Graduate

2. Analyze rigorously the mathematical properties of methods used in data science
Audience: Graduate

STAT 990 – RESEARCH
1-12 credits.

Independent research and writing for graduate students under the supervision of a faculty member.
Requisites: Consent of instructor
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: Yes, unlimited number of completions
Last Taught: Spring 2024

STAT 992 – SEMINAR
1-3 credits.

Special topics in statistics at the graduate level. Subject matter varies.
Requisites: Graduate/professional standing
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: Yes, unlimited number of completions
Last Taught: Spring 2024

STAT 998 – STATISTICAL CONSULTING
3 credits.

Consulting apprenticeship.
Requisites: Graduate/professional standing
Course Designation: Grad 50% - Counts toward 50% graduate coursework requirement
Repeatable for Credit: No
Last Taught: Spring 2024