Statistics and Actuarial Science

Chair

- Kung-Sik Chan

Director of Undergraduate Studies, Actuarial Science

- Elias S. Shiu

Director of Undergraduate Studies, Data Science and Statistics

- Sanvesh Srivastava

Director of Graduate Studies

- Boxiang Wang

Undergraduate majors: actuarial science (BS); statistics (BS)

Undergraduate minor: statistics

Graduate degrees: MS in actuarial science; MS in statistics; PhD in statistics

Faculty: https://stat.uiowa.edu/people

Website: https://stat.uiowa.edu/

The Department of Statistics and Actuarial Science offers undergraduate majors, an undergraduate minor, and graduate degree programs. They partner with other departments to offer the BS in data science and the undergraduate Certificate in Social Science Analytics. The department offers courses that any undergraduate student may use to satisfy the GE CLAS Core Quantitative or Formal Reasoning requirement.

Probability and statistics are important scientific disciplines essential to all fields of study that rely on information obtained from data. In a world bombarded with numerical information, informed decisions rely on the ability to separate fact from fiction by applying valid statistical analyses and visualizations. Statisticians can provide crucial guidance in determining what information is reliable and which predictions may be trusted. They often help search for clues to the solution of a scientific mystery and sometimes keep investigators from being misled by false impressions.

The work of a statistician may range from the theoretical (developing new methodologies and statistical theory) to the applied (working with scientists and decision-makers to collect, analyze, and interpret data). Regardless of the areas in which they work, statisticians need strong mathematical, computational, and communication skills. Because uncertainty and data arise in many settings, statisticians have the opportunity to work on a variety of projects in industry, education, government, and research. Thousands of statisticians work in medicine, law, agriculture, business, finance, public policy, marketing, manufacturing, engineering, and other fields in the social and natural sciences. The diversity of applications is an exciting aspect of the field and is one reason why the demand for well-trained statisticians continues to be strong.

An actuary is a business executive, professionally trained in the mathematical sciences. Actuaries specialize in the evaluation of financial risk—most often in the context of life, health, and casualty insurance, where they design, analyze, and refine varied programs to meet the insurance needs of society. Many actuaries are employed by insurance companies, where they have responsibility for all phases of the development and maintenance of their company's products. They have considerable influence on the financial soundness of their company through work in pricing insurance policies and compiling data for financial statements.

Many actuaries are employed as consultants. Their actuarial services are used by smaller insurance companies and individual employers who need actuarial guidance in establishing insurance and retirement programs for their employees. A growing number of actuaries work in the areas of asset/liability management and risk management. Some of these actuaries are employed by investment and consulting firms; others are employed by insurance companies.

Actuaries have been called financial architects and social mathematicians because their combined analytical and business skills help solve a growing variety of financial and social problems. The actuarial profession is a demanding yet rewarding career choice.

Related Major and Certificate

Major: Data Science

The BS in data science produces graduates with the sophisticated analytical and computational skills required to thrive in a quantitative world where new problems are encountered at an ever-increasing rate. The major emphasizes the statistical/probabilistic and algorithmic methods that underlie the preparation, analysis, and communication of complex data. With a focus on technical foundations, the data science program promotes skills useful for creating and implementing new or special-purpose analysis and visualization tools. It also promotes a fundamental understanding of how to best handle uncertainty when making data-driven decisions.

Statistics majors may not earn a major in data science. The Department of Statistics and Actuarial Science and the Department of Computer Science collaborate to offer the major in data science. The BS in data science is administered by the Department of Statistics and Actuarial Science; see the BS in data science in the catalog.

Certificate: Social Science Analytics

The growth of big data and informatics calls for a new set of skills for social science students and an increased understanding of the logic of data collection and analysis. The certificate focuses on the application side of data analysis. It allows students to focus on the specific research methods and data-driven quantitative skills that are most effective for understanding an increasingly complicated social-political world. The certificate offers an opportunity for interdisciplinary training on how data can be used to address important questions in the social sciences. The Department of Statistics and Actuarial Science collaborates with the departments of Political Science and Sociology and Criminology and the School of Earth, Environment, and Sustainability to offer the undergraduate program in social science analytics; see the Certificate in Social Science Analytics in the catalog.

Programs

Undergraduate Programs of Study

Majors

- Major in Statistics (Bachelor of Science)
- Major in Actuarial Science (Bachelor of Science)

Minor

- Minor in Statistics

Graduate Programs of Study

- Master of Science in Statistics
- Master of Science in Actuarial Science
- Doctor of Philosophy in Statistics

Facilities

The Department of Statistics and Actuarial Science is housed in Schaeffer Hall, adjacent to the Old Capitol, a National Historic Landmark and the center of campus. The department operates two computer labs in Schaeffer Hall. One, which also is used as an electronic classroom, contains 28 Windows PCs. The second houses 18 high-end UNIX workstations. Students use these labs for classwork and research.

Courses

Undergraduate Duplication and Regression Policy

Data science majors may not earn a major or minor in computer science or statistics, a major in computer science and engineering, or the Certificate in Social Science Analytics. Likewise, statistics majors may not earn a major in data science.

Undergraduate students should be aware of the duplication and regression policies concerning the following courses.

Students may earn credit for only two of these: STAT:1010 Statistics and Society, STAT:1015/DATA:1015 Introduction to Data Science, STAT:1020/PSQF:1020 Elementary Statistics and Inference, STAT:1030 Statistics for Business, and STAT:2010 Statistical Methods and Computing.

Credit for STAT:1010 Statistics and Society may be earned only if the course is taken before any of these: STAT:1015/DATA:1015 Introduction to Data Science, STAT:1020/PSQF:1020 Elementary Statistics and Inference, STAT:1030 Statistics for Business, and STAT:2010 Statistical Methods and Computing.

Students may receive credit for only one course from each of these pairs: STAT:2010 Statistical Methods and Computing and STAT:4200/IGPI:4200 Statistical Methods and Computing; STAT:3100/IGPI:3100 Introduction to Mathematical Statistics I and STAT:3120/DATA:3120/IGPI:3120 Probability and Statistics; STAT:3510/IGPI:3510 Biostatistics and STAT:4143/PSQF:4143 Introduction to Statistical Methods; STAT:4100 Statistical Inference I and STAT:5100 Statistical Inference I; and STAT:4101 Statistical Inference II and STAT:5101 Statistical Inference II.

Students may not take STAT:3101/IGPI:3101 Introduction to Mathematical Statistics II and STAT:4101/IGPI:4101 Statistical Inference II at the same time and get credit for both (nor go back to STAT:3101/IGPI:3101 after taking STAT:4101/IGPI:4101).

Statistics Courses

STAT:1000 First-Year Seminar 1 s.h.
Small discussion class taught by a faculty member; topics chosen by instructor; may include outside activities (e.g., films, lectures, performances, readings, visits to research facilities). Requirements: first- or second-semester standing.

STAT:1010 Statistics and Society 3 s.h.
Statistical ideas and their relevance to public policy, business, humanities, and the social, health, and physical sciences; focus on critical approach to statistical evidence. Requirements: one year of high school algebra or MATH:0100. GE: Quantitative or Formal Reasoning.

STAT:1015 Introduction to Data Science 3 s.h.
In today's world, massive amounts of data are increasingly collected and leveraged for knowledge discovery, policy assessment, and decision-making across many fields, including business, natural sciences, social sciences, and humanities. Topics covered include data collection, visualization, and data wrangling; basics of probability and statistical inference; fundamentals of data learning, including regression, classification, prediction, and cross-validation; computing, learning, and reporting in the R environment; and literate programming and reproducible research. Requirements: one year of high school algebra or MATH:0100. GE: Quantitative or Formal Reasoning. Same as DATA:1015.

STAT:1020 Elementary Statistics and Inference 3 s.h.
Graphing techniques for presenting data, descriptive statistics, correlation, regression, prediction, logic of statistical inference, elementary probability models, estimation and tests of significance. Requirements: one year of high school algebra or MATH:0100. GE: Quantitative or Formal Reasoning. Same as PSQF:1020.

STAT:1030 Statistics for Business 4 s.h.
Descriptive statistics, graphical presentation, elementary probability, estimation and testing, regression, correlation; statistical computer packages. GE: Quantitative or Formal Reasoning.

STAT:2010 Statistical Methods and Computing 3 s.h.
Methods of data description and analysis using SAS; descriptive statistics, graphical presentation, estimation, hypothesis testing, sample size, power; emphasis on learning statistical methods and concepts through hands-on experience with real data. Recommendations: undergraduate standing. GE: Quantitative or Formal Reasoning.

STAT:2020 Probability and Statistics for the Engineering and Physical Sciences 3 s.h.
Probability, random variables, important discrete and continuous distributions, joint distributions, transformations of random variables, descriptive statistics, point and interval estimation, tests of hypotheses, regression. Prerequisites: MATH:1560 or MATH:1860.

STAT:3100 Introduction to Mathematical Statistics I 4 s.h.
Review of mathematical basics for mathematical statistics, descriptive statistics, probability, conditional probability, discrete and continuous univariate and multivariate distributions, and sampling distributions. Prerequisites: MATH:1860 or MATH:1560. Same as IGPI:3100.

STAT:3101 Introduction to Mathematical Statistics II 3 s.h.
Point and interval estimation, testing statistical hypotheses, simple regression, nonparametric methods. Prerequisites: STAT:3100. Same as IGPI:3101.

STAT:3120 Probability and Statistics 4 s.h.
Models, discrete and continuous random variables and their distributions, estimation of parameters, testing statistical hypotheses. Prerequisites: MATH:1560 or MATH:1860. Same as DATA:3120, IGPI:3120.

STAT:3200 Applied Linear Regression 3 s.h.
Regression analysis with focus on applications; model formulation, checking, and selection; interpretation and presentation of analysis results; simple and multiple linear regression; logistic regression; ANOVA; polynomial regression; tree models; bootstrapping; hands-on data analysis with computer software. Prerequisites: STAT:2020 or STAT:2010 or STAT:3120. Same as DATA:3200, IGPI:3200, ISE:3760.

STAT:3210 Experimental Design and Analysis 3 s.h.
Single- and multifactor experiments; analysis of variance; multiple comparisons; contrasts; diagnostics; fixed, random, and mixed effects models; designs with blocking and/or nesting; two-level factorials and fractions thereof; use of statistical computing packages. Prerequisites: STAT:3200.

STAT:3510 Biostatistics 3 s.h.
Statistical concepts and methods for the biological sciences; descriptive statistics, elementary probability, sampling distributions, confidence intervals, parametric and nonparametric methods, one-way ANOVA, correlation and regression, categorical data. Requirements: MATH:0100 or MATH:1005 or ALEKS score of 30 or higher. Same as IGPI:3510.

STAT:3620 Quality Control 3 s.h.
Basic techniques of statistical quality control; application of control charts for process control variables; design of inspection plans and industrial experimentation; modern management aspects of quality assurance systems. Offered fall semesters. Prerequisites: STAT:2020 or BAIS:9100 or (STAT:3100 and STAT:3101 and STAT:3200). Same as CEE:3142, ISE:3600.

STAT:4100 Statistical Inference I 3 s.h.
Review of probability, distribution theory (multiple random variables, moment-generating functions, transformations, conditional distributions), sampling distributions, order statistics, convergence concepts, generating random samples. Prerequisites: MATH:2850 and STAT:3101. Recommendations: undergraduate standing. Same as IGPI:4100.

STAT:4101 Statistical Inference II 3 s.h.
Continuation of STAT:4100; principles of data reduction, point estimation theory (MLE, Bayes, UMVU), hypothesis testing, interval estimation, decision theory, asymptotic evaluations. Prerequisites: STAT:4100. Recommendations: undergraduate standing. Same as IGPI:4101.

STAT:4143 Introduction to Statistical Methods 3 s.h.
Analysis, interpretation of research data; descriptive statistics; introduction to probability, sampling theory, statistical inference (binomial, normal distribution, t-distribution models); linear correlation, regression. Same as PSQF:4143.

STAT:4200 Statistical Methods and Computing 3 s.h.
Methods of data description and analysis using SAS; descriptive statistics, graphical presentation, estimation, hypothesis testing, sample size, power; emphasis on learning statistical methods and concepts through hands-on experience with real data. Recommendations: graduate standing in non-statistics or less quantitative major. Same as IGPI:4200.

STAT:4520 Bayesian Statistics 3 s.h.
Bayesian statistical analysis, with focus on applications; Bayesian and frequentist methods compared; Bayesian model specification, choice of priors, computational methods; hands-on Bayesian data analysis using appropriate software; interpretation and presentation of analysis results. Prerequisites: STAT:3200 and (STAT:3101 or STAT:4101 or STAT:3120). Same as IGPI:4522, PSQF:4520.

STAT:4540 Statistical Learning 3 s.h.
Introduction to supervised and unsupervised statistical learning, with a focus on regression, classification, and clustering; methods will be applied to real data using appropriate software; supervised learning topics include linear and nonlinear (e.g., logistic) regression, linear discriminant analysis, cross-validation, bootstrapping, model selection, and regularization methods (e.g., ridge and lasso); generalized additive and spline models, tree-based methods, random forests and boosting, and support-vector machines; unsupervised learning topics include principal components and clustering. Requirements: an introductory statistics course and a regression course. Recommendations: prior exposure to programming and/or software, such as R, SAS, and Matlab. Same as BAIS:4540, DATA:4540, IGPI:4540.

STAT:4560 Statistics for Risk Modeling I 3 s.h.
Simple linear regression, multiple linear regression, model diagnostics, linear models from a statistical learning perspective, generalized linear models, and implementations of these models on real data. Prerequisites: STAT:4101 with a minimum grade of C+ or STAT:5101 with a minimum grade of C+.

STAT:4561 Statistics for Risk Modeling II 3 s.h.
Regression-based time series models, decision trees, principal components analysis, cluster analysis, and implementations of these analytic techniques on real data. Prerequisites: STAT:4560 with a minimum grade of C+.

STAT:4580 Data Visualization and Data Technologies 3 s.h.
Introduction to common techniques for visualizing univariate and multivariate data, data summaries, and modeling results; how to create and interpret these visualizations and assess effectiveness of different visualizations based on an understanding of human perception and statistical thinking; data technologies for obtaining and preparing data for visualization and further analysis; students learn how to present results in written reports and use version control to manage their work. Requirements: an introductory statistics course and a regression course. Recommendations: prior exposure to basic use of statistical programming software (e.g., R or SAS) as obtained from a regression course strongly recommended. Same as DATA:4580, IGPI:4580.

STAT:4600 Causal Inference for Data Science 3 s.h.
Introduce methods for reasoning about causes, effects, and bias when analyzing experimental and observational data. Topics include the potential outcomes framework, counterfactuals, confounding, and missing data; the identification and estimation of causal effects via propensity score methods, marginal structural models, instrumental variables, and directed acyclic graphs; as well as applications of machine learning and Bayesian methods to causal inference. Prerequisites: (DATA:3120 or STAT:3120) and (DATA:3200 or STAT:3200). Requirements: familiarity with the R programming. Same as DATA:4600.

STAT:4740 Large Data Analysis 3 s.h.
Current areas that deal with problem of big data; techniques from computer science, mathematics, statistics; high performance and parallel computing, matrix techniques, cluster analysis, visualization; variety of applications including Google PageRank, seismology, Netflix-type problems, weather forecasting; fusion of data with simulation; projects. Prerequisites: (CS:1210 with a minimum grade of C- or ENGR:2730 with a minimum grade of C-) and (MATH:3800 or CS:3700) and (STAT:3200 or IE:3760 or IGPI:3200). Same as CS:4740, IGPI:4740, MATH:4740.

STAT:4750 Probabilistic Statistical Learning 3 s.h.
Essential machine learning and statistics ideas that are critical in analyzing modern complex and large data; supervised learning topics include linear models, deep neural networks, and nonparametric models; essential topics include nonlinear dimension reduction, clustering, and recommender systems. Prerequisites: (CS:1210 with a minimum grade of C- or ENGR:2730 with a minimum grade of C-) and (MATH:2700 or MATH:2550) and (STAT:2010 or STAT:2020 or STAT:4200) and STAT:4540. Same as DATA:4750.

STAT:5090 ALPHA Seminar 1 s.h.
Resources available to students, program requirements, tips for academic success, professional statistical organizations, library and career center resources, statistical computing, scientific document preparation, history of statistics. Requirements: graduate standing in statistics.

STAT:5091 Statistics Career Seminar 1 s.h.
Practical, hands-on approach to demystify the job search process and build essential career-readiness skills for roles in statistics, actuarial science, and data science across academia, industry, government, and non-profit sectors. Same as GRAD:5091.

STAT:5100 Statistical Inference I 3 s.h.
Review of probability, distribution theory (multiple random variables, moment-generating functions, transformations, conditional distributions), sampling distributions, order statistics, convergence concepts, generating random samples. Prerequisites: MATH:2850 and STAT:3101.

STAT:5101 Statistical Inference II 3 s.h.
Continuation of STAT:5100; principles of data reduction, point estimation theory (MLE, Bayes, UMVU), hypothesis testing, interval estimation, decision theory, asymptotic evaluations. Prerequisites: STAT:5100.

STAT:5120 Mathematical Methods for Statistics 3 s.h.
Real numbers, point set theory, limit points, limits, metric spaces, continuity, sequences and series, Taylor series (multivariate), uniform convergence, Riemann-Stieltjes integrals.

STAT:5200 Applied Statistics I 4 s.h.
Descriptive statistics, basic inferential methods (confidence intervals, chi-square tests); linear models (regression and ANOVA models—specification and assumptions, fitting, diagnostics, selection, testing, interpretation); nonlinear models, logistic regression. Prerequisites: STAT:3101. Corequisites: STAT:4100 or STAT:5100. Requirements: facility with matrix algebra. Same as IGPI:5199.

STAT:5201 Applied Statistics II 3 s.h.
Design of experiments and analysis of designed experiments; models for fixed and random effects; mixed models; design and analysis of complex plans; sample-size methods. Prerequisites: STAT:5200. Recommendations: prior exposure to SAS software.

STAT:5400 Computing in Statistics 3 s.h.
R; database management; graphical techniques; importing graphics into word-processing documents (e.g., LaTeX); creating reports in LaTeX; SAS; simulation methods (Monte Carlo studies, bootstrap, etc.). Prerequisites: CS:1210 and STAT:3200 and (STAT:3120 or STAT:3101 or STAT:4101). Corequisites: STAT:5100 and STAT:5200 if not already completed. Same as DATA:5400, IGPI:5400.

STAT:5610 Regression Modeling and ANOVA in the Health Sciences 3 s.h.
Continuation of BIOS:4120; correlation, simple and multiple linear regression, confounding, interactions, model selection, single and multiple factor ANOVA (analysis of variance) models, contrasts, multiple comparisons, nested and block designs; introduction to mixed models; for non-biostatistics majors. Offered spring semesters. Prerequisites: BIOS:4120. Same as BIOS:5120, IGPI:5120.

STAT:5810 Research Data Management 3 s.h.
Introduction to data management techniques and problems encountered in gathering and processing data from biomedical investigations; introduction to SAS, techniques taught in SAS; designed for non-biostatistics majors. Recommendations: prior programming experience with C, C++, Python, Java, or other. Same as BIOS:5310, IGPI:5310.

STAT:6200 Predictive Analytics 3 s.h.
Linear mixed models; generalized linear mixed models; generalized additive models; applications of these models using associated R packages. Prerequisites: STAT:4560. Corequisites: STAT:4561. Requirements: comfort working with R software environment. Same as ACTS:6200, DATA:6200.

STAT:6220 Consulting and Communication With Data 3 s.h.
Realistic supervised data analysis experiences, including statistical packages, statistical graphics, writing statistical reports, dealing with complex or messy data. Offered spring semesters. Prerequisites: (STAT:3200 and STAT:3210) or (STAT:5201 and STAT:5200). Requirements: for undergraduate majors—major GPA of 3.00 or above, and grades of B or higher in STAT:3200 and STAT:3210. Same as DATA:6220.

STAT:6300 Probability and Stochastic Processes I 3 s.h.
Conditional expectations; Markov chains, including random walks and gambler's ruin; classification of states; stationary distributions; branching processes; Poisson processes; Brownian motion. Prerequisites: STAT:4100.

STAT:6301 Probability and Stochastic Processes II 3 s.h.
Markov chains with continuous state space, Martingales, random walks, Brownian motion and other continuous-time Markov chains, simulation methods. Prerequisites: STAT:6300.

STAT:6513 Intermediate Statistical Methods 3 s.h.
Statistical inference and uncertainty estimation using general linear models (i.e., linear regression, analysis of variance); interpreting and conveying statistical results. Requirements: for PSQF:6243—PSQF:4143; for STAT:6513—STAT:4143. Same as PSQF:6243.

STAT:6514 Correlation and Regression 4 s.h.
Correlation techniques; selected bivariate procedures, multiple, partial, curvilinear correlation; multiple linear regression; sampling theory applied to regression analysis and correlation coefficients; simple causal models. Requirements: for PSQF:6244—PSQF:6243; for STAT:6514—STAT:6513. Same as PSQF:6244.

STAT:6516 Design of Experiments 3 s.h.
Theory and methods in the planning and statistical analysis of experimental studies; testing of hypotheses about linear contrasts among means in single-factor and multifactor, completely randomized, and repeated measurement designs. Requirements: for PSQF:6246—PSQF:6243; for STAT:6516—STAT:6513. Same as PSQF:6246.

STAT:6530 Environmental and Spatial Statistics 3 s.h.
Geostatistics kriging, variogram estimation, trend estimation, sampling design, extensions to river networks and the globe, lattice data analysis, analysis of spatial point patterns. Prerequisites: STAT:4101 and STAT:3200. Same as IGPI:6530.

STAT:6547 Nonparametric Statistical Methods 3 s.h.
Selected nonparametric methods; one- and two-sample location tests and estimation methods, measures of association, analyses of variance; emphasis on relationships to classical parametric procedures. Same as PSQF:6247.

STAT:6550 Introductory Longitudinal Data Analysis 3 s.h.
Introduction to statistical models and estimation methods for outcome variables (normal and non-normal) clustered or measured repeatedly in time or space; focus on applications and computer software methods for ANOVA based methods, hierarchical linear models, linear mixed models, correlated regression models, generalized estimating equations, and generalized linear mixed models. Offered fall semesters. Prerequisites: BIOS:5120 or STAT:3200. Same as BIOS:6310, IGPI:6310.

STAT:6560 Applied Time Series Analysis 3 s.h.
General stationary, nonstationary models, autocovariance autocorrelation functions; stationary, nonstationary autoregressive integrated moving average models; identification, estimation, forecasting in linear models; use of statistical computer packages. Offered spring semesters. Prerequisites: (STAT:5200 or STAT:3200) and STAT:3101.

STAT:6970 Topics in Statistics 3 s.h.
Selected advanced topics in statistics.

STAT:6990 Readings in Statistics arr.
Supervised reading and research in statistics.

STAT:7100 Advanced Inference I 3 s.h.
Concepts of convergence, asymptotic methods including the delta method, sufficiency, asymptotic efficiency, Fisher information and information bounds for estimation, maximum likelihood estimation, the EM-algorithm, Bayes estimation, decision theory. Prerequisites: STAT:5101 and STAT:5120.

STAT:7101 Advanced Inference II 3 s.h.
Hypothesis testing, asymptotics of the likelihood ratio test, asymptotic efficiency, statistical functionals, robustness, bootstrap and jackknife, estimation with dependent data. Prerequisites: STAT:7100.

STAT:7190 Seminar: Mathematical Statistics arr.
Selected advanced topics in mathematical statistics.

STAT:7200 Linear Models 4 s.h.
Linear spaces and selected topics in matrix algebra, full rank and non-full rank linear models, estimability, least squares and best linear unbiased estimation, multivariate normal distribution and distributions of quadratic forms, interval estimation, hypothesis testing, random and mixed models, best linear unbiased prediction, variance component estimation. Prerequisites: STAT:5101 and STAT:5200 and STAT:5201.

STAT:7290 Seminar: Applied Statistics arr.
Selected advanced topics in applied statistics.

STAT:7300 Advanced Probability 3 s.h.
Probability theory, with emphasis on constructing rigorous proofs; measure spaces, measurable functions, random variables and induced measures, distribution functions, Lebesque integral, product measure and independence, Borel Cantelli lemma, modes of convergence. Prerequisites: STAT:5120.

STAT:7390 Seminar: Probability arr.
Selected advanced topics in probability.

STAT:7400 Computer Intensive Statistics 3 s.h.
Computer arithmetic, random variate generation, numerical optimization, numerical linear algebra, smoothing techniques, bootstrap methods, cross-validation, MCMC, EM and related algorithms; other topics per student/instructor interests. Prerequisites: (BIOS:5710 or STAT:5200) and STAT:3101 and STAT:5400. Requirements: proficiency in Fortran or C or C++ or Java. Same as DATA:7400, IGPI:7400.

STAT:7500 Statistical Machine Learning 3 s.h.
Regularization methods for sparse models, computational algorithms for large scale problems, statistical inference in high-dimensional models, reproducing kernel Hilbert space, supervised learning, nonparametric density and conditional density estimation, neural networks and deep learning, optimal transport and generative learning, dimension reduction and representation learning. Prerequisites: STAT:5100 or STAT:5200. Same as BAIS:7500.

STAT:7510 Analysis of Categorical Data 3 s.h.
Distributions and inference for categorical data; descriptive and inferential methods for contingency tables; theory and application of generalized linear models including methods for parameter estimation and testing, model selection and assessment of model adequacy; models for binary, count, and multi-category outcomes; loglinear models for contingency tables; generalized additive models; generalized linear models for longitudinal and clustered data. Offered spring semesters of even years. Prerequisites: (BIOS:5720 or STAT:5200) and (STAT:5101 or STAT:4101). Same as BIOS:7410.

STAT:7520 Bayesian Analysis 3 s.h.
Decision theory, conjugate families, structure of Bayesian inference, hierarchical models, asymptotic approximations for posterior distributions, Markov chain Monte Carlo methods and convergence assessment, model adequacy and model choice. Prerequisites: STAT:5101 and STAT:5400 and STAT:5200.

STAT:7560 Time Series Analysis 3 s.h.
Stationary time series, ARIMA models, spectral representation, linear prediction inference for the spectrum, multivariate time series, state space models and processes, nonlinear time series. Prerequisites: STAT:4101 and (STAT:3200 or STAT:6560).

STAT:7570 Survival Data Analysis 3 s.h.
Types of censoring and truncation; survival function estimation; parametric inference using exponential, Weibull, and accelerated failure time models; nonparametric tests; sample size calculation; Cox regression with stratification and time-dependent covariates; regression diagnostics; competing risks; topics may include analysis of correlated survival data and/or recurrent events; designed for biostatistics and statistics majors. Offered fall semesters of odd years. Prerequisites: BIOS:5720 and ((STAT:4100 and STAT:4101) or (STAT:5100 and STAT:5101)). Same as BIOS:7210, IGPI:7210.

STAT:7990 Reading Research arr.
Supervised reading and research in statistics.

Actuarial Science Courses

ACTS:1000 First-Year Seminar 1 s.h.
Small discussion class taught by a faculty member; topics chosen by instructor; may include outside activities (e.g., films, lectures, performances, readings, visits to research facilities). Requirements: first- or second-semester standing.

ACTS:1001 Introductory Seminar on Actuarial Science 1 s.h.
The actuarial profession, exam process, academic success, internships, resources for actuarial students, important skills for actuaries, and career paths in actuarial science.

ACTS:3080 Mathematics of Finance I 3 s.h.
Mathematics of compound interest, annuities certain, loan amortization schedules, bonds, yield rates, and introduction to interest rate risk management. Prerequisites: MATH:1860 with a minimum grade of B-. Requirements: calculus II or graduate standing.

ACTS:3110 Actuarial Exam P Preparation 1 s.h.
Preparation for the Society of Actuaries exam P. Corequisites: STAT:3100 or STAT:4100 or STAT:5100.

ACTS:3210 Actuarial Exam FM Preparation 1 s.h.
Preparation for the Society of Actuaries exam FM. Corequisites: ACTS:3080, if not taken as a prerequisite.

ACTS:4130 Quantitative Methods for Actuaries 3 s.h.
Survival distributions, life tables, life insurance, life annuities, and premiums. Offered fall semesters. Prerequisites: STAT:3100 with a minimum grade of B- and ACTS:3080 with a minimum grade of C+. Corequisites: STAT:4100 or STAT:5100. Requirements: multivariate calculus and linear algebra.

ACTS:4150 Fundamentals of Short-Term Actuarial Mathematics 3 s.h.
Severity, frequency, aggregate loss, estimation, credibility theory, pricing, and reserving for short-term insurance coverages; option pricing. Offered spring semesters. Prerequisites: STAT:4100 with a minimum grade of C+ or STAT:5100 with a minimum grade of C+. Corequisites: STAT:4101 or STAT:5101.

ACTS:4160 Topics in Actuarial Science arr.
Selected topics in actuarial science, financial mathematics, and quantitative risk management.

ACTS:4180 Life Contingencies I 3 s.h.
Reserves, multi-life models, multiple-decrement models, and Markov chains. Offered spring semesters. Prerequisites: ACTS:3080 with a minimum grade of C+ and ACTS:4130 with a minimum grade of C+ and (STAT:4100 with a minimum grade of C+ or STAT:5100 with a minimum grade of C+).

ACTS:4280 Life Contingencies II 3 s.h.
Multistate models, pension mathematics, emerging costs for traditional and equity-linked insurance, profit testing, profit measures, and embedded options. Offered fall semesters. Prerequisites: ACTS:4180 with a minimum grade of C+.

ACTS:4380 Mathematics of Finance II 3 s.h.
Derivatives markets, forwards, options, pricing models, and actuarial applications. Prerequisites: ACTS:3080 with a minimum grade of C+. Requirements: mathematical statistics, multivariate calculus, and linear algebra.

ACTS:4990 Readings in Actuarial Science arr.

ACTS:6160 Topics in Actuarial Science arr.
Selected topics in actuarial science, financial mathematics, and quantitative risk management.

ACTS:6200 Predictive Analytics 3 s.h.
Linear mixed models; generalized linear mixed models; generalized additive models; applications of these models using associated R packages. Prerequisites: STAT:4560. Corequisites: STAT:4561. Requirements: comfort working with R software environment. Same as DATA:6200, STAT:6200.

ACTS:6480 Loss Distributions 3 s.h.
Severity, frequency, and aggregate models and their modifications; risk measures; construction of empirical models. Offered spring semesters. Prerequisites: STAT:4101 or STAT:5101. Corequisites: ACTS:6580.

ACTS:6580 Credibility and Survival Analysis 3 s.h.
Construction and selection of parametric models, credibility, and simulation. Offered spring semesters. Prerequisites: STAT:4101 or STAT:5101. Corequisites: ACTS:6480.

ACTS:6990 Readings in Actuarial Science arr.
Supervised reading and research in actuarial science, financial mathematics, or quantitative risk management.

ACTS:7730 Advanced Topics in Actuarial Science/ Financial Mathematics arr.
Selected advanced topics in actuarial science, financial mathematics and quantitative risk management.

Data Science Courses

DATA:1015 Introduction to Data Science 3 s.h.
In today's world, massive amounts of data are increasingly collected and leveraged for knowledge discovery, policy assessment, and decision-making across many fields, including business, natural sciences, social sciences, and humanities. Topics covered include data collection, visualization, and data wrangling; basics of probability and statistical inference; fundamentals of data learning, including regression, classification, prediction, and cross-validation; computing, learning, and reporting in the R environment; and literate programming and reproducible research. Requirements: one year of high school algebra or MATH:0100. GE: Quantitative or Formal Reasoning. Same as STAT:1015.

DATA:3120 Probability and Statistics 4 s.h.
Models, discrete and continuous random variables and their distributions, estimation of parameters, testing statistical hypotheses. Prerequisites: MATH:1560 or MATH:1860. Same as IGPI:3120, STAT:3120.

DATA:3200 Applied Linear Regression 3 s.h.
Regression analysis with focus on applications; model formulation, checking, and selection; interpretation and presentation of analysis results; simple and multiple linear regression; logistic regression; ANOVA; polynomial regression; tree models; bootstrapping; hands-on data analysis with computer software. Prerequisites: STAT:2020 or STAT:2010 or STAT:3120. Same as IGPI:3200, ISE:3760, STAT:3200.

DATA:4540 Statistical Learning 3 s.h.
Introduction to supervised and unsupervised statistical learning, with a focus on regression, classification, and clustering; methods will be applied to real data using appropriate software; supervised learning topics include linear and nonlinear (e.g., logistic) regression, linear discriminant analysis, cross-validation, bootstrapping, model selection, and regularization methods (e.g., ridge and lasso); generalized additive and spline models, tree-based methods, random forests and boosting, and support-vector machines; unsupervised learning topics include principal components and clustering. Requirements: an introductory statistics course and a regression course. Recommendations: prior exposure to programming and/or software, such as R, SAS, and Matlab. Same as BAIS:4540, IGPI:4540, STAT:4540.

DATA:4580 Data Visualization and Data Technologies 3 s.h.
Introduction to common techniques for visualizing univariate and multivariate data, data summaries, and modeling results; how to create and interpret these visualizations and assess effectiveness of different visualizations based on an understanding of human perception and statistical thinking; data technologies for obtaining and preparing data for visualization and further analysis; students learn how to present results in written reports and use version control to manage their work. Requirements: an introductory statistics course and a regression course. Recommendations: prior exposure to basic use of statistical programming software (e.g., R or SAS) as obtained from a regression course strongly recommended. Same as IGPI:4580, STAT:4580.

DATA:4600 Causal Inference for Data Science 3 s.h.
Introduce methods for reasoning about causes, effects, and bias when analyzing experimental and observational data. Topics include the potential outcomes framework, counterfactuals, confounding, and missing data; the identification and estimation of causal effects via propensity score methods, marginal structural models, instrumental variables, and directed acyclic graphs; as well as applications of machine learning and Bayesian methods to causal inference. Prerequisites: (DATA:3120 or STAT:3120) and (DATA:3200 or STAT:3200). Requirements: familiarity with the R programming. Same as STAT:4600.

DATA:4610 Data Acquisition and Management 3 s.h.
Introduction to common techniques for manipulating relational databases for data analysis; SQL and PostgreSQL fundamentals: querying, data manipulation and transformation, joins and subqueries, aggregation and grouping, data types and management; advanced topics: window functions, subqueries, common table expressions, indexing strategies, performance optimization techniques, security considerations; database building. Prerequisites: DATA:3200 or STAT:3200. Recommendations: Familiarity with basic programming logic, e.g., variables, loops, conditional statements.

DATA:4620 Text Data Analysis 3 s.h.
Introduction to text analytics techniques for real-world applications; Python fundamentals for text data exploration and manipulation; text processing via NLP libraries (NLTK, spaCy, Gensim); feature engineering; sentiment analysis; topic modeling; text summarization, machine translation, and deep learning applications. Prerequisites: (CS:1210 or DATA:5400) and DATA:4540. Recommendations: Basic knowledge of Python programming.

DATA:4750 Probabilistic Statistical Learning 3 s.h.
Essential machine learning and statistics ideas that are critical in analyzing modern complex and large data; supervised learning topics include linear models, deep neural networks, and nonparametric models; essential topics include nonlinear dimension reduction, clustering, and recommender systems. Prerequisites: (CS:1210 with a minimum grade of C- or ENGR:2730 with a minimum grade of C-) and (MATH:2700 or MATH:2550) and (STAT:2010 or STAT:2020 or STAT:4200) and STAT:4540. Same as STAT:4750.

DATA:4880 Data Science Creative Component 1 s.h.
Readings, group discussions, and short-term projects in area of data science; emphasis on communication of ideas learned in student's data science coursework, data ethics, and potential bias in algorithms.

DATA:4890 Data Science Practicum 3 s.h.
On- or off-campus internship or group-based consulting project that provides experience in a real-world setting; application of knowledge and techniques learned in coursework; practice in communicating results to others.

DATA:5400 Computing in Statistics 3 s.h.
R; database management; graphical techniques; importing graphics into word-processing documents (e.g., LaTeX); creating reports in LaTeX; SAS; simulation methods (Monte Carlo studies, bootstrap, etc.). Prerequisites: CS:1210 and STAT:3200 and (STAT:3120 or STAT:3101 or STAT:4101). Corequisites: STAT:5100 and STAT:5200 if not already completed. Same as IGPI:5400, STAT:5400.

DATA:5890 MS Data Science Practicum 2 s.h.
On- or off-campus internship or group-based consulting project that provides experience in a real-world setting; application of knowledge and techniques learned in coursework and practice communicating results to others.

DATA:6200 Predictive Analytics 3 s.h.
Linear mixed models; generalized linear mixed models; generalized additive models; applications of these models using associated R packages. Prerequisites: STAT:4560. Corequisites: STAT:4561. Requirements: comfort working with R software environment. Same as ACTS:6200, STAT:6200.

DATA:6220 Consulting and Communication With Data 3 s.h.
Realistic supervised data analysis experiences, including statistical packages, statistical graphics, writing statistical reports, dealing with complex or messy data. Offered spring semesters. Prerequisites: (STAT:3200 and STAT:3210) or (STAT:5201 and STAT:5200). Requirements: for undergraduate majors—major GPA of 3.00 or above, and grades of B or higher in STAT:3200 and STAT:3210. Same as STAT:6220.

DATA:7350 High-Dimensional Probability for Data Science 3 s.h.
Nonasymptotic probability with a view towards applications in data science; concentration inequalities for functions of independent variables, martingale inequalities, entropy method, random matrices, matrix inequalities, suprema of random processes, and sparse recovery. Prerequisites: STAT:5101. Requirements: linear algebra course and familiarity with R or Python.

DATA:7400 Computer Intensive Statistics 3 s.h.
Computer arithmetic, random variate generation, numerical optimization, numerical linear algebra, smoothing techniques, bootstrap methods, cross-validation, MCMC, EM and related algorithms; other topics per student/instructor interests. Prerequisites: (BIOS:5710 or STAT:5200) and STAT:3101 and STAT:5400. Requirements: proficiency in Fortran or C or C++ or Java. Same as IGPI:7400, STAT:7400.