STATISTICS

Statistics is one of the natural, mathematical, and biomedical sciences programs in the Columbian College of Arts and Sciences. The curriculum emphasizes the important role of statistics as it provides methodologies for making advances in medicine, genetics, and other research arenas and supports decision making in business and public policy. Students learn reasoning skills and methods for analyzing and understanding data and how these skills can be applied to develop new initiatives.

Visit the Department of Statistics website (https://statistics.columbian.gwu.edu/) for additional information.

UNDERGRADUATE

Bachelor's program

- Bachelor of Science with a major in statistics (http://bulletin.gwu.edu/arts-sciences/statistics/bs/)

Minor

- Minor in statistics (http://bulletin.gwu.edu/arts-sciences/statistics/minor/)

GRADUATE

Master's programs

- Master of Science in the field of statistics (http://bulletin.gwu.edu/arts-sciences/statistics/ms/)

Combined program

- Dual Master of Science in the field of statistics and graduate certificate in the field of data science (http://bulletin.gwu.edu/arts-sciences/statistics/dual-ms-gc-data-science/)

DOCTORAL

Doctoral programs

- Doctor of Philosophy in the field of statistics (http://bulletin.gwu.edu/arts-sciences/statistics/phd/)

FACULTY

Professors: J.L. Gastwirth, F. Hu, S. Kundu, J.M. Lachin III, Y. Lai, Z. Li, H. Liang, H.M. Mahmoud, R. Modarres, T.K. Nayak, Q. Pan, R. Soyer, H.J. Wang (Chair)

Associate Professors: T. Apanasovich, S. Bose, X. Zhang

Assistant Professors: S. Balaji, F. Jin, J. Landon

Adjunct Professors: S. Amini

COURSES

Explanation of Course Numbers

- Courses in the 1000s are primarily introductory undergraduate courses
- Those in the 2000s to 4000s are upper-level undergraduate courses that can also be taken for graduate credit with permission and additional work assigned
- Those in the 6000s and 8000s are for master's, doctoral, and professional-level students
- The 6000s are open to advanced undergraduate students with approval of the instructor and the dean or advising office

Note: STAT 1051 Introduction to Business and Economic Statistics, STAT 1053 Introduction to Statistics in Social Science, STAT 1111 Business and Economic Statistics I, and STAT 1127 Statistics for the Biological Sciences are related in their subject matter, and credit for only one of these courses may be applied toward a degree. One entrance unit in algebra is prerequisite to all courses in statistics.

STAT 1000. Dean's Seminar. 3 Credits.
The Dean's Seminars provide Columbian College first-year students focused scholarship on specific intellectual challenges. Topics vary by semester; see department for more details.

STAT 1051. Introduction to Business and Economic Statistics. 3 Credits.
Lecture (3 hours), laboratory (1 hour). Frequency distributions, descriptive measures, probability, probability distributions, sampling, estimation, tests of hypotheses, regression and correlation, with applications to business. Credit cannot be earned for this course and STAT 1053, STAT 1111, STAT 1127.

STAT 1053. Introduction to Statistics in Social Science. 3 Credits.
Lecture (3 hours), laboratory (1 hour). Frequency distributions, descriptive measures, probability, sampling, estimation, tests of hypotheses, regression and correlation, with applications to social sciences. Credit cannot be earned for this course and STAT 1051, STAT 1111, STAT 1127.

STAT 1099. Variable Topics. 1-36 Credits.

STAT 1111. Business and Economic Statistics I. 3 Credits.
Descriptive statistics, graphical methods, probability, special distributions, random variables, sampling, estimation and confidence intervals, hypothesis testing, correlation and regression. Credit cannot be earned for this course and STAT 1051, STAT 1053, STAT 1127.

STAT 1127. Statistics for the Biological Sciences. 3 Credits.
Introduction to statistical techniques and reasoning applicable to the biomedical and related sciences. Properties of basic probability functions: binomial, Poisson, and normal. Data analysis, inference, and experimental design. Credit cannot be earned for this course and STAT 1051, STAT 1053, STAT 1111.

STAT 1129. Introduction to Computing. 3 Credits.
Introduction to elements of computer programming and problem-solving using a computer programming language. Hands-on experience is acquired through computer programming projects, including some simple statistical applications.

STAT 2000. Sophomore Colloquium. 3 Credits.
Sophomore colloquia are small, seminar-type classes that deeply engage CCAS second-year students in a discipline, focus on a narrow issue of high interest and impact, and require independent research projects. May be repeated provided topic differs. Consult the Schedule of Classes for more details. Restricted to CCAS sophomores.

STAT 2112. Business and Economic Statistics II. 3 Credits.
Continuation of STAT 1111. Emphasis on techniques of regression, chi-square, nonparametric inference, index numbers, time series, decision analysis, and other topics relevant to economics and business. Prerequisites: One of the following - STAT 1051, STAT 1053, STAT 1111, or DNSC 1001. Credit cannot be earned for this course and DNSC 2001.

STAT 2118. Regression Analysis. 3 Credits.
Lecture (3 hours), laboratory (1 hour). Simple and multiple linear regression, partial correlation, residual analysis, stepwise model building, multicollinearity and diagnostic methods, indicator variables. Prerequisites: One of the following - STAT 1051, STAT 1053, STAT 1111, or DNSC 1001.

STAT 2123. Introduction to Econometrics. 3 Credits.
Analyzing causal relationships among economic variables through the lens of regression analysis. Students learn to critically assess empirical research in economics and beyond. Prerequisites: ECON 1011 or HONR 2043; and ECON 1012 or HONR 2044; and MATH 1221 or MATH 1231; and STAT 1051 or STAT 1053 or STAT 1111 or DNSC 1001. Same As: ECON 2123.

STAT 2183. Intermediate Statistics Lab/Packages. 3 Credits.
Application of program packages (e.g., SAS, SPSS) to the solution of one-, two- and k-sample parametric and nonparametric statistical problems. Basic concepts in data preparation, modification, analysis and interpretation of results. Prerequisites: An introductory statistics course.

STAT 2183W. Intermediate Statistical Laboratory: Statistical Computing Packages. 3 Credits.
Application of program packages (e.g., SAS, SPSS) to the solution of one-, two- and k-sample parametric and nonparametric statistical problems. Basic concepts in data preparation, modification, analysis and interpretation of results. Includes a significant engagement in writing as a form of critical inquiry and scholarly expression to satisfy the WID requirement. Prerequisites: An introductory statistics course.

STAT 3099. Variable Topics. 1-12 Credits.

STAT 3119. Design and Analysis of Experiments. 3 Credits.
Lecture (3 hours), laboratory (1 hour). Introduction to the design of experiments and analysis of variance; randomized block, factorial, Latin square designs, and analysis of covariance, random effects, response surface methodology. Prerequisites: STAT 2118. Recommended background: Familiarity with R, linear algebra, and matrix theory.

STAT 3157. Introduction to Mathematical Statistics I. 3 Credits.
Basic concepts of probability theory, including random variables, independence, distribution theory, and sampling theory. Prerequisites: MATH 1232.

STAT 3187. Introduction to Sampling. 3 Credits.
Problems of sampling and sample design. Simple random, stratified, systematic, cluster, and multistate designs; control of sampling and non-sampling errors. Prerequisites: STAT 1051.

STAT 3187W. Introduction to Sampling. 3 Credits.
Problems of sampling and sample design. Simple random, stratified, systematic, cluster, and multistate designs; control of sampling and non-sampling errors. Includes a significant engagement in writing as a form of critical inquiry and scholarly expression to satisfy the WID requirement. Prerequisite: STAT 1051. Same As: STAT 3187.

STAT 4157. Introduction to Mathematical Statistics I. 3 Credits.
Basic concepts of probability theory, including random variables, independence, distribution theory, and sampling theory. Prerequisites: MATH 1232.

STAT 4158. Introduction to Mathematical Statistics II. 3 Credits.
Continuation of STAT 4157. Inference procedures, including estimation, hypothesis testing, regression analysis, and experimental design. Prerequisites: MATH 1232.

STAT 4181. Applied Time Series Analysis. 3 Credits.
Autoregressive integrated moving average (ARIMA) modeling and forecasting of univariate time series. Estimation of spectral density functions, white noise tests, and tests for periodicities. Theory and applications using statistical software. Prerequisites: STAT 4157 and STAT 4158 or STAT 2118.

STAT 4188. Nonparametric Statistics Inference. 3 Credits.
Statistical inference when the form of the underlying distribution is not fully specified. Nonparametric procedures for estimation and testing hypotheses. An introduction to robust procedures. Prerequisites: STAT 1051.

STAT 4189. Mathematical Probability and Applications I. 3 Credits.
Probability theory, including combinatorial analysis, conditional probability, and stochastic independence. Random variables and their distributions; laws of large numbers and central limit theorem. Application of concepts to elementary stochastic processes (coin-tossing sequences, branching processes, Markov chains). Prerequisite: MATH 1232 .

STAT 4190. Mathematical Probability and Applications II. 3 Credits.
Continuation of STAT 4189. Probability theory, including combinatorial analysis, conditional probability, and stochastic independence. Random variables and their distributions; laws of large numbers and central limit theorem. Application of concepts to elementary stochastic processes (coin-tossing sequences, branching processes, Markov chains). Prerequisite: MATH 1232 .

STAT 4195. Reading and Research. 1-12 Credits.
May be repeated once for credit. Permission of the department chair required prior to enrollment.

STAT 4197. Fundamentals of SAS Programming for Data Management. 3 Credits.
Fundamentals of Statistical Analysis System (SAS) software for data management, statistical analysis, and report writing; data modification, programming, file handling, and macro writing. Students are expected to have knowledge of computer programming and to have completed an introductory statistics course. Credit cannot be earned for both STAT 4197 and STAT 6197.

STAT 4198. Special Topics. 3 Credits.
Topics vary by semester. May be repeated for credit provided the topic differs. See department for more details.

STAT 5099. Variable Topics. 1-99 Credits.

STAT 6197. Fundamentals of SAS Programming for Data Management. 3 Credits.
Fundamentals of Statistical Analysis System (SAS) software for data management, statistical analysis, and report writing. Data modification, programming, file handling, and macro writing. Recommended background: An introductory statistics course and knowledge of computer programming. Credit cannot be earned for this course and STAT 4197.

STAT 6201. Mathematical Statistics I. 3 Credits.
Basic Probability theory, Random variables and transformations, Common families of distribution, Conditional expectations and distributions, Bivariate and Multivariate distributions and transformations, Sampling distributions. Prerequisites: MATH 2233 and MATH 2184.

STAT 6202. Mathematical Statistics II. 3 Credits.
Continuation of STAT 6201. Order Statistics, Convergence concepts, Sufficient and Complete statistics, Likelihood Principle, Point and Interval Estimation, Hypothesis Testing, Bayesian Tests and Intervals. Prerequisites: MATH 2233, MATH 2184 and STAT 6201.

STAT 6207. Methods of Statistical Computing I. 3 Credits.
Error analysis, computational aspects of linear models, sweep operator, random number generation, simulation, resampling. Optimization, numerical integration (Gaussian quadrature, Simpson's rule); E–M algorithm. Prerequisites: STAT 2118, STAT 4157 and STAT 4158; and MATH 2184; and knowledge of a programming language.

STAT 6208. Methods of Statistical Computing II. 3 Credits.
Numerical linear algebra, matrix decomposition and eigenvalue problems. Smoothing and density estimation. Graphics, interactive and dynamic techniques for data display. Object-oriented programming. Prerequisites: STAT 2118, STAT 4157 and STAT 4158; and MATH 2184; and knowledge of a programming language.

STAT 6210. Data Analysis. 3 Credits.
Review of statistical principles of data analysis, using computerized statistical procedures. Multiple regression and the general linear model, analysis of contingency tables and categorical data, logistic regression for qualitative responses. Prerequisites: STAT 2118 or STAT 4157 or STAT 6201; and STAT 2183.

STAT 6213. Intermediate Probability and Stochastic Processes. 3 Credits.
Discrete and continuous random variables and their distributions, conditional distributions and conditional expectation, generating functions and their applications, convergence of random variables; introduction to Brownian motion, homogeneous and nonhomogeneous Poisson processes and martingales. Prerequisites: STAT 6201 and STAT 6202.

STAT 6214. Applied Linear Models. 3 Credits.
Introduction to regression techniques for discrete and continuous response variables. The course includes a computing component using SAS and S> Prerequisite: MATH 2233 and MATH 2184.

STAT 6215. Applied Multivariate Analysis I. 3 Credits.
Statistical analysis of several variables, possibly dependent, following a joint normal distribution. matrix algebra and random vectors, multivariate sample geometry, multivariate normal distribution, inferences about a mean vector, and comparisons of several population means. Applications of multivariate techniques to the analysis of data from the behavioral, social, medical, and physical sciences. Prerequisites: STAT 4157 and STAT 4158; and MATH 2184.

STAT 6216. Applied Multivariate Analysis II. 3 Credits.
Continuation of STAT 6215. Statistical analysis of random vectors, following a multivariate normal distribution. multivariate linear regression models, principal components, factor analysis, inference for structured covariance matrices, canonical correlations, discrimination and classification, clustering and distance methods. Applications of multivariate techniques to the analysis of data from the behavioral, social, medical, and physical sciences. Prerequisites: STAT 3119, STAT 4157 and STAT 4158; and MATH 2184.

STAT 6217. Design of Experiments. 3 Credits.
Design and analysis of single- and multiple-factor experiments. Includes block designs, repeated measures, factorial and fractional factorial experiments, response surface experimentation. Prerequisites: STAT 4157 and STAT 4158; and MATH 2184.

STAT 6218. Linear Models. 3 Credits.
Theory of the general linear parametric model. Includes least squares estimation, multiple comparisons procedures, variance components estimation. Prerequisites: STAT 6201, STAT 6202, STAT 2118 and MATH 2184.

STAT 6223. Bayesian Statistics: Theory and Applications. 3 Credits.
An overview of Bayesian statistics, including its foundational issues, decision under uncertainty, linear models, expert opinion, and computational issues. Prerequisites: STAT 6201 and STAT 6202.

STAT 6225. Longitudinal Data Analysis. 3 Credits.
Introduction to the statistical models, estimation methods, and inferences for the analysis of longitudinal data; modern methods for the analysis of repeated measures as well as parametric and nonparametric regression models for longitudinal analysis. Restricted to master of science and doctoral program candidates. Prerequisites: Stat 2118, Stat 6201 and Stat 6202.

STAT 6227. Survival Analysis. 3 Credits.
Parametric and nonparametric methods for the analysis of events observed in time (survival data), including Kaplan–Meier estimate of survival functions, logrank and generalized Wilcoxon tests, the Cox proportional hazards model and an introduction to counting processes. Prerequisites: STAT 6201 and STAT 6202; or permission of the instructor.

STAT 6230. Causal Inference. 3 Credits.
Introduction to causal inference, potential outcomes, counterfactual reasoning, randomized experiments, observational study, propensity score, stratification, subclassification, matching, instrumental variables, applications of causal inference. Restricted to master of science and doctoral program candidates. Prerequisites: STAT 2118, STAT 6201, and MATH 2184. Recommended background: Familiarity with basic linear algebra, mathematical probability and statistics, and linear regression.

STAT 6231. Categorical Data Analysis. 3 Credits.
A study of the theoretical bases underlying the analysis of categorical data. Measures and tests of association; Mantel-Haenszel procedure; weighted least squares and maximum likelihood estimators in linear models; generalized linear models including Poisson and logistic regression models, generalized estimating equations; random effects models Prerequisites: Graduate level mathematical statistics courses - STAT 6201 and STAT 6202. Recommended background: Courses in Mathematical Statistics and Linear Models.

STAT 6236. Applied Sampling Techniques. 3 Credits.
Problems of sampling and sample design. Simple random, stratified, systematic, cluster, and multistate designs; control of sampling and non-sampling errors. Prerequisite: STAT 1051.

STAT 6240. Statistical Data Mining. 3 Credits.
Introduction to basic data mining concepts and techniques for discovering interesting patterns hidden in large-scale data sets, focusing on issues relating to effectiveness and efficiency. Students are expected to be familiar with R programming. Restricted to statistics majors or with the permission of the instructor. Prerequisites: STAT 6201, STAT 6202, and STAT 6214 or equivalents. Recommended background: coursework in mathematical statistics, applied linear models, and multivariate statistics.

STAT 6242. Modern Regression Analysis. 3 Credits.
Methodology, major software tools and applications of modern nonparametric methods. Regularized regression: shrinkage, ridge and lasso; nonparametric regression: kernels and splines; nonparametric classification: K-Nearest Neighbors and Decision Trees; resampling methods: bootstrap, boosting and bagging. Prerequisites: STAT 6201 or STAT 6202 or STAT 6214 or STAT 6218.

STAT 6245. Statistical Consulting. 3 Credits.
This course focuses on the following themes: (i) understanding the statistical consulting process; (ii) developing effective verbal and written communication skills; (iii) comprehending consulting environments in different industries; and (iv) obtaining consulting experience through case studies. Prerequisites: STAT 6201, STAT 6202, STAT 6214 and STAT 6215. Recommended background: second-year status in the graduate statistics or biostatistics program.

STAT 6250. A/B Testing. 3 Credits.
Concepts and basic designs of A/B testing, adaptive and covariate-adaptive designs, testing in medical studies, network data and adaptive designs, and statistical inferences for A/B testing. Restricted to MS and PhD students. Prerequisites: STAT 6201, STAT 6202, STAT 6214, and STAT 6215. Recommended background: Coursework in mathematical statistics, linear models, and multivariate analysis.

STAT 6252. Statistical Methods in Bioinformatics and Computational Biology. 3 Credits.

STAT 6253. Legal Statistics. 3 Credits.

STAT 6254. Statistical Genetics. 3 Credits.
Theories of population genetics and Mendelian genetics, Hardy-Weinberg equilibrium and linkage disequilibrium, statistical software (R or SAS) for linkage analysis and association analysis, research in statistical genetics. Prerequisites: STAT 6201 and STAT 6202.

STAT 6255. Clinical Trials. 3 Credits.
Introduction to the design and analysis of clinical trials. Clinical trials as a tool for medical research; phases and endpoints; the role of randomization, power, and sample size; statistical analysis of data; interim analysis and data monitoring. Recommended background: Knowledge of basic design of experiments, mathematical statistics (probability and inference), and familiarity with R and SAS.

STAT 6260. Statistical Deep Learning. 3 Credits.
Introduction to deep neural network models, architecture, and algorithms, with applications to image classification, natural language processing, and sequence of data prediction. Students learn to build deep learning models with Python. Restricted to statistics majors or with the permission of the instructor. Prerequisites: STAT 6201 or equivalent. Recommended background: familiarity with Python.

STAT 6287. Sample Surveys. 3 Credits.
Application of statistical theory to the sampling of finite populations. Simple, stratified, cluster, double and subsampling. Special topics, including super-populations and randomized response. Prerequisites: STAT 4157 and STAT 4158.

STAT 6289. Topics in Statistics. 3 Credits.
Topics vary by semester. May be repeated for credit provided the topic differs. See department for more details.

STAT 6291. Methods of Demographic Analysis. 3 Credits.
Basic methods for analysis of mortality, natality, and migration; population estimates and projections; estimation of demographic measures from incomplete data. Departmental prerequisite waived.

STAT 6295. Reading and Research. 3 Credits.
May be repeated once for credit.

STAT 6298. Seminar: Special Topics. 3 Credits.

STAT 6999. Thesis Research. 3,6 Credits.
Development of a thesis project and accompanying research.

STAT 8226. Advanced Biostatistical Methods. 3 Credits.
Statistical methods for the analysis of longitudinal data: nonparametric, fixed effects, mixed effects, generalized estimating equations. Methods for the analysis of emerging data: group sequential analysis, Brownian motion, Bayesian methods, and stochastic curtailment. Other advanced topics of current research in biostatistics. Prerequisites: STAT 6201 and STAT 6202; or permission of the instructor.

STAT 8257. Probability. 3 Credits.
Probabilistic foundations of statistics, probability distributions, random variables, moments, characteristic functions, modes of convergence, limit theorems, probability bounds. Prerequisites: STAT 6201 and STAT 6202; and knowledge of calculus through functions of several variables and series.

STAT 8258. Distribution Theory. 3 Credits.
Special distributions of statistics, small and large sample theory, order statistics, and spacings. Prerequisite: STAT 8257.

STAT 8259. Advanced Probability. 3 Credits.
Conditional expectation and martingales; weak convergence in general metric spaces and functional central limit theorems for i.i.d. random variables and martingales; applications to biostatistics. Prerequisite: STAT 8257 or an equivalent measure-theoretic introduction to probability.

STAT 8262. Nonparametric Inference. 3 Credits.
Inference when the form of the underlying distribution is unspecified. Prerequisites: STAT 6201 and STAT 6202.

STAT 8263. Advanced Statistical Theory I. 3 Credits.
Decision theoretic estimation, classical point estimation, hypothesis testing. Prerequisites: STAT 6201 and STAT 6202.

STAT 8264. Advanced Statistical Theory II. 3 Credits.
Asymptotic theory, hypothesis testing, confidence regions. Prerequisites: STAT 8257 and STAT 8263.

STAT 8265. Multivariate Analysis. 3 Credits.
Characterization and properties of the multivariate normal distribution, conditional distributions, multiple correlation, partial correlation, estimation of the mean vector and the covariance matrix, Wishart and Hotelling distributions and applications to hypothesis testing, discrimination, classification, and principle component analysis. Prerequisites: STAT 6201 and STAT 6202.

STAT 8266. Topics-Multivariate Analysis. 3 Credits.
Multivariate analysis of variance, principal component analysis, canonical correlation, factor analysis. Prerequisites: STAT 6201, STAT 6202 and STAT 8265.

STAT 8271. Foundational and Philosophical Issues in Statistics. 3 Credits.
Axiomatic underpinnings of Bayesian statistics, including subjective probability, belief, utility, decision and games, likelihood principle, and stopping rules. Examples from legal, forensic, biological, and engineering sciences. Students are expected to have a background in computer science, economics, mathematics, or operations research. Prerequisites: STAT 6201 and STAT 6202.

STAT 8273. Stochastic Processes I. 3 Credits.
Fundamental notions of Markov chains and processes, generating functions, recurrence, limit theorems, random walks, Poisson processes, birth and death processes, applications. Prerequisites: STAT 6201 and STAT 6202.

STAT 8274. Stochastic Processes II. 3 Credits.
Continuation of STAT 8273. Fundamental notions of Markov chains and processes, generating functions, recurrence, limit theorems, random walks, Poisson processes, birth and death processes, applications. Prerequisites: STAT 6201 and STAT 6202.

STAT 8281. Advanced Time Series Analysis. 3 Credits.
Autoregressive integrated moving average (ARIMA) modeling and forecasting of univariate and multivariate time series. Statespace or Kalman filter models, spectral analysis of multiple time series. Theory and applications using the University computer. Prerequisites: MATH 2233, STAT 6201 and STAT 6202.

STAT 8288. Topics in Sample Surveys. 3 Credits.
Advanced topics and research in sample surveys. Prerequisite: STAT 6287.

STAT 8289. Seminar. 3 Credits.
Admission by permission of instructor. May be repeated for credit provided the content differs.

STAT 8375. Econometrics I. 3 Credits.
Statistical foundations for econometrics; standard methods of estimation and inference for classical and generalized regression models. Same as ECON 8375.

STAT 8376. Econometrics II. 3 Credits.
Topics may include asymptotic theory, statistical endogeneity, instrumental variables estimation, discrete and limited dependent variable models, and time–series models. Same as ECON 8376. Prerequisite: STAT 8375.

STAT 8998. Advanced Reading and Research. 1-12 Credits.
May be repeated for credit. Restricted to doctoral candidates preparing for the general examination.

STAT 8999. Dissertation Research. 3-12 Credits.
May be repeated for credit. Restricted to doctoral candidates.