

Statistics

Departmental Counselor: Michael Stein, E 135, 702.8326 or 702.8333,

stein@galton.uchicago.edu

Web: www.stat.uchicago.edu

Program of Study

The modern science of statistics involves the invention, study, and development of principles and methods for modeling uncertainty through mathematical probability; for designing experiments, surveys, and observational programs; and for analyzing and interpreting empirical data. Mathematics plays a major role in all statistical activity, whether of an abstract nature or dealing with specific techniques for analyzing data. Statistics is an excellent field for students with strong mathematical skills and an interest in applying these skills to problems in the natural and social sciences. A program leading to the BA degree in statistics offers coverage of the principles and methods of statistics in combination with a solid training in mathematics and some exposure to computing, which is essential to nearly all modern data analysis. In addition, there is considerable elective freedom enabling interested students to examine those areas of knowledge in the biological, physical, and social sciences that are often subjected to detailed statistical analysis. The major provides a base for graduate study in statistics or in other subjects with strong quantitative components. Students considering graduate study in statistics or related fields are encouraged to discuss their programs with the departmental counselor at an early stage, whether or not they plan to receive an undergraduate degree in statistics.

Students who are majoring in other fields of study may also complete a minor in statistics. Information follows the description of the major.

Statistics Courses for Students in Other Majors. Courses at the 20000 level are designed to provide instruction in statistics, probability, and statistical computation for students from all parts of the University. These courses differ in emphasis on theory or methods, on the mathematical level, and in the direction of applications. Most of the introductory courses make intensive use of computers to exemplify and explore statistical concepts and methods. The nature and extent of computer work varies according to the course and instructor. Statistics courses are not mathematics courses, but the mathematics prerequisites are a useful guide to the level of mathematical maturity assumed by a statistics course. Students with a background in calculus typically are advised to take STAT 22000 or higher.

Explanations and comparisons of the various courses, both entry level and more advanced, are provided in the following sections. Students will also find the course descriptions to be helpful in choosing appropriate courses.

Introductory Courses and Sequences. STAT 22000, which typically is the statistics course taken first, is a general introduction to statistical concepts,

techniques, and applications to data analysis and to problems in the design, analysis, and interpretation of experiments and observational programs. Computers are used throughout the course. A score of 4 or 5 on the AP test in statistics yields credit for STAT 22000, although this credit will not count toward the requirements for a major in statistics. STAT 23400 covers much of the same material as 22000, but at a somewhat higher mathematical level. STAT 23400 is a required course for students who are majoring in economics, but the class is a one-quarter introduction to statistics that is appropriate for any student with a good command of univariate calculus. For their introductory statistics course, students should choose *either* STAT 22000 *or* 23400 (*not both*). For students who do not intend to continue to more advanced statistics courses, STAT 20000 is an alternative with no calculus prerequisite that places less emphasis on statistical techniques. STAT 20000 may not be taken either by students who have already taken STAT 22000 or 23400 or by students who have received AP credit for statistics. STAT 25100 is an introductory course in probability.

STAT 24400-24500 is recommended for students who wish to have a thorough introduction to statistical theory and methodology. STAT 24400-24500 is more mathematically demanding than either STAT 22000 or 23400 and assumes some familiarity with multiple integration and with linear algebra. Students considering a major in statistics are encouraged to take STAT 24400 rather than STAT 23400. Although students with a strong mathematical background can and do take STAT 24400-24500 without prior course work in statistics or probability, many students find it helpful to take a more elementary course as preparation. Students who have already taken STAT 22000 or 23400 and wish to study statistics at a higher mathematical level are welcome to take STAT 24400-24500. STAT 24600 is a follow-up to 24400-24500 that covers more advanced statistical methods. STAT 24400-24500 and 25100 form the core of the statistics major; this is recommended as a cognate sequence to students in the quantitative sciences and mathematics. Taking STAT 24400-24500 before STAT 25100 is recommended but not required.

For students interested in exploring methods and their applications, STAT 22200, 22400, 22600, and 22700 are recommended. These complementary second courses emphasize some class of methods for the analysis of data. They may be taken in any order, although there is some overlap between STAT 22600 and 22700. Also, STAT 22400 is a prerequisite for STAT 22700. Each presumes a previous course in statistics (STAT 22000 or equivalent) and experience using computers in data analysis (as in STAT 22000). The emphasis is on linear models and experimental design in STAT 22200, multiple regression and least squares in STAT 22400, categorical data analysis in STAT 22600, and statistical methods for medical applications in STAT 22700. STAT 26100, which covers time dependent data, is appropriate for students with some knowledge of linear models (STAT 22400 or 24400-24500).

For students who have completed STAT 24500, many graduate courses in statistics offer opportunities for further study of statistical theory, methods, and applications. The introductory probability course (STAT 25100) may be

taken separately from any statistics courses and can be supplemented with more advanced probability courses, such as STAT 25300 (=31700). Students with a strong mathematical background may take STAT 31200, 31300, 38100, and 38300. College students may register for a number of other 30000-level courses in statistics. For details, consult the instructor or the departmental counselor, or visit galton.uchicago.edu.

Program Requirements

Students pursuing the BA in statistics should meet the general education requirements in the mathematical sciences with courses in calculus. The program includes four additional prescribed mathematics courses and four prescribed statistics courses; students should complete the four mathematics courses by the end of their third year. Additional requirements include one course in computer science and three approved elective courses in statistics. The four required statistics courses are STAT 24400-24500 and 25100; and either 22400 or 34300. STAT 24400 typically is suggested as a first course in statistics or, if a more elementary introduction is desired, students may take STAT 22000 or 23400 as preparation for 24400. *If either STAT 22000 or 23400 is included in the program (students may not include both), it will be counted as one of three approved electives and must be taken before STAT 24400.* Candidates must obtain approval of their course program from the departmental counselor; not all combinations of statistics electives are allowed. Specifically, at least two of the three electives must be courses in statistical methodology beyond the introductory level. NOTE: Students who are completing majors in both statistics and economics and who already have taken MATH 19520-19620 should discuss with the departmental counselor how to best meet their mathematical requirements.

Summary of Requirements for the BA in Statistics

General Education	MATH 13100-13200, 15100-15200, or 16100-16200*		
Major	1	MATH 13300, 15300, or 16300*	
	2	MATH 20000-20100, 20400-20500, or 20800-20900	
	1	STAT 24300 or MATH 25500 or MATH 25800	
	4	STAT 24400, 24500, and 25100; and either 22400 or 34300	
	1	CMSC 10500, 10600, 12100, 15100, or 16100 (10600 or higher preferred)	
	3	approved elective courses in statistics**	
	12		

* Credit may be granted by examination.
** For example, STAT 22200, 22600 or 22700 (but not both), 24600, 25300 or 31200 (but not both), 26100, or 26700. Upon petition, one intermediate/advanced course in mathematics or computer science may be approved for this purpose by the statistics departmental counselor as relevant for a coherent degree program. The petition must include a documented strong case for the relevance.

Grading. Subject to College and divisional regulations, and with the consent of the instructor, *all students except majors in statistics* may register for quality grades or for *P/F* grading in any 20000-level statistics course. A grade of *P* is given only for work of *C-* quality or higher.

In addition to submitting the official Incomplete Form required by the College, the following policy applies to students who wish to receive a mark of *I* for a statistics course: students must have completed at least half of the total required course work with a grade of *C-* or better and they must be unable to complete all of the course work by the end of the quarter due to an emergency.

Students who are majoring in statistics must receive a quality grade of at least *C-* in each of the twelve required courses in the major; a grade of *P* is not acceptable for any of these courses.

Honors. The BA with honors is awarded to students with statistics as their primary major who have a GPA of 3.0 or higher overall and 3.25 or higher in the twelve required courses in the major, and who, in addition, also complete an approved honors paper (STAT 29900). This paper typically is based upon a structured research program that students undertake with faculty supervision in the first quarter of their fourth year. Eligible students who wish to be considered for honors must consult the departmental counselor before the end of their third year. The research paper or project used to meet this requirement may not be used to meet the BA paper or project requirement in another major. NOTE: Credit

for STAT 29900 will not count towards the twelve courses required for a major in statistics.

Joint BA/MS Program. This program enables qualified undergraduate students to complete an MS in statistics along with a BA during their four years at the College. Although a student may receive a BA in any field, a program of study other than statistics is recommended.

Participants must be admitted to the MS program in statistics. Students must submit applications by June 1 of their third year for admission to candidacy for an MS in statistics during their fourth year. Interested students are strongly encouraged to consult the departmental counselor and Ron Gorny, the BA/MA adviser, early in their third year. (For a appointment with Mr. Gorny, call the College Adviser’s Reception Desk at 702-8615.)

Participants in the joint BA/MS program must meet the same requirements as students in the MS program in statistics. Of the nine courses that are required at the appropriate level, up to three may also meet the requirements of an undergraduate program. For example, STAT 24400-24500 and 24600, which are required for the MS in statistics, could also be used to meet part of the requirements of a BA or BS program in mathematics for courses outside of mathematics.

Other requirements include a master’s paper and participation in the consulting program of the Department of Statistics. For details, visit galton.uchicago.edu/admissions/master.html.

Minor Program in Statistics

The focus in the minor is on statistical methodology, whereas the statistics major has a substantial theoretical component. Students can begin the statistics minor with either STAT 22000 or STAT 23400 as their introductory course, requiring just two or three quarters of calculus as prerequisite. STAT 24500 (but not STAT 24400) may also be used to satisfy the introductory statistics requirement.

The minor in statistics requires an introductory course, two courses in statistical methods, and two approved electives on statistical topics chosen to complement a student’s major or personal interest. If the introductory course is a required component of a student’s major or if AP credit for STAT 22000 is used to satisfy the introductory course requirement, a third approved elective must be included to complete the statistics minor.

Summary of Requirements

- Introductory Statistics (choose one course):
STAT 22000. Statistical Methods and Applications
STAT 23400. Statistical Models and Methods

Statistical Methods (choose two courses; one must be STAT 22400):

STAT 22200. Linear Models and Experimental Design
 STAT 22400. Applied Regression Analysis
 STAT 22600. Analysis of Categorical Data
 STAT 22700. Biostatistical Methods

Additional Topics in Statistics (choose two courses):

STAT 22200. Linear Models and Experimental Design
 STAT 22600. Analysis of Categorical Data
 STAT 22700. Biostatistical Methods
 STAT 24300. Numerical Linear Algebra
 STAT 24600. Complex Statistical Problems
 STAT 26100. Time Dependent Data
 STAT 26700. History of Statistics
 STAT 33200. Causal Inference
 STAT 35000. Principles of Epidemiology
 STAT 35600. Applied Survival Analysis
 STAT 36900. Longitudinal Data Analysis
 HSTD 32800. Modern Data Analysis in Biostatistics
 HSTD 32901. Introduction to Clinical Trials

The topics courses on the list above are approved for the statistics minor. Students may petition the departmental counselor for approval of another course. Such courses must have a minimum statistics prerequisite of STAT 22000 or equivalent. The following statistics courses may not be included in a statistics minor: STAT 20000, 25100, or 25300; or any graduate courses in probability. Students may not include both STAT 22600 and 22700 in the minor. If either STAT 22000 or 23400 is required for another degree, then one additional statistical topics course must be chosen to complete the minimum five-course requirement for the minor in statistics. Any prerequisite mathematics courses needed are not a part of the statistics minor and may be counted toward a major or toward general education requirements.

Students who elect the minor program in statistics must meet with the departmental counselor before the end of Spring Quarter of their third year to declare their intention to complete the minor. The approval of the departmental counselor for the minor program should be submitted to a student's College adviser by the deadline above on a form obtained from the College adviser. (The deadline for students graduating in June or August of 2010 is Friday of first week of Spring Quarter 2010.) Courses for the minor are chosen in consultation with the departmental counselor.

Courses in the minor (1) may not be double counted with the student's major(s) or with other minors and (2) may not be counted toward general education requirements. Courses in the minor must be taken for quality grades, and students must receive a grade of C- or higher in each course taken for the minor. More than half of the requirements for the minor must be met by registering for courses bearing University of Chicago course numbers.

Faculty

Y. Amit, M. Anitescu, M. Drton, P. Fischer, L. Hansen, D. Heinz, N. Hinrichs, S. Lalley, G. Lawler, L. Lim, P. McCullagh, M. S. McPeck, D. Mondal, P. Mykland, D. Nicolae, P. Niyogi, M. Stein, M. Stephens, S. Stigler, R. Thisted, M. Wang, M. Wichura, K. Wolter, W. Wu

Courses: Statistics (STAT)

20000. Elementary Statistics. *This course is recommended for students who do not plan to take advanced statistics courses, and it may not be used in the statistics major. It is not open to students with credit for STAT 22000 or 23400 who matriculated in the College after August 2008. This course meets one of the general education requirements in the mathematical sciences.* This course introduces statistical concepts and methods for the collection, presentation, analysis, and interpretation of data. Elements of sampling, simple techniques for analysis of means, proportions, and linear association are used to illustrate both effective and fallacious uses of statistics. *Autumn, Winter, Spring.*

22000. Statistical Methods and Applications. (=HDCP 22050) *PQ: Two quarters of calculus. Students who matriculate in the College after August 2008 may count either STAT 22000 or 23400, but not both, toward the forty-two credits required for graduation.* This course introduces statistical techniques and methods of data analysis, including the use of computers. Examples are drawn from the biological, physical, and social sciences. Students are required to apply the techniques discussed to data drawn from actual research. Topics include data description, graphical techniques, exploratory data analyses, random variation and sampling, one- and two-sample problems, analysis of variance, linear regression, and analysis of discrete data. *Autumn, Winter, Spring.*

22200. Linear Models and Experimental Design. *PQ: STAT 22000 or 23400 or 24500.* This course covers principles and techniques for the analysis of experimental data and the planning of the statistical aspects of experiments. Topics include linear models; analysis of variance; randomization, blocking, and factorial designs; confounding; and incorporation of covariate information. *Spring.*

22400. Applied Regression Analysis. (=HSTD 32400) *PQ: STAT 22000 or 23400 or 24500 or HSTD 32100.* This course introduces the methods and applications of fitting and interpreting multiple regression models. The primary emphasis is on the method of least squares and its many varieties. Topics include the examination of residuals, the transformation of data, strategies and criteria for the selection of a regression equation, the use of dummy variables, tests of fit, nonlinear models, biases due to excluded variables and measurement error, and the use and interpretation of computer package regression programs. The techniques discussed are illustrated by many real examples involving data from both the natural and social sciences. Matrix notation is introduced as needed. *Autumn.*

22600. Analysis of Categorical Data. (=HSTD 32600) *PQ: STAT 22000 or 22400 or 24500.* This course covers statistical methods for the analysis of structured, counted data. Topics may include Poisson, multinomial, and product-multinomial sampling models; chi-square and likelihood ratio tests; log-linear models for cross-classified counted data, including models for data with ordinal categories and log-multiplicative models; logistic regression and logit linear models; and measures of association. Applications in the social and biological sciences are considered, and the interpretation of models and fits, rather than mathematical details of computational procedures, is emphasized. *Winter.*

22700. Biostatistical Methods. (=HSTD 32700) *PQ: HSTD 32400/STAT 22400 or 24500.* This course is designed to provide students with tools for analyzing categorical, count, and time-to-event data frequently encountered in medicine, public health, and related biological and social sciences. This course emphasizes application of the methodology rather than statistical theory (e.g., recognition of the appropriate methods; interpretation and presentation of results). Methods covered include contingency table analysis, Kaplan-Meier survival analysis, Cox proportional-hazards survival analysis, logistic regression, and Poisson regression. *Winter.*

23400. Statistical Models and Methods. *PQ: MATH 13300, 15300, or 16300. Students who matriculate in the College after August 2008 may count either STAT 22000 or 23400, but not both, toward the forty-two credits required for graduation.* This course is recommended for students throughout the natural and social sciences who want a broad background in statistical methodology and exposure to probability models and the statistical concepts underlying the methodology. Probability is developed for the purpose of modeling outcomes of random phenomena. Random variables and their expectations are studied; including means and variances of linear combinations and an introduction to conditional expectation. Binomial, Poisson, normal and other standard probability distributions are considered. Some probability models are studied mathematically, and others are studied via simulation on a computer. Sampling distributions and related statistical methods are explored mathematically, studied via simulation, and illustrated on data. Methods include, but are not limited to, inference for means and variances for one- and two-sample problems, correlation, and simple linear regression. Graphical description and numerical data description are used for exploration, communication of results, and comparing mathematical consequences of probability models and data. Mathematics employed is to the level of univariate calculus, but it is less demanding than that required by STAT 24400. *Autumn, Winter, Spring.*

24300. Numerical Linear Algebra. *PQ: Multivariate calculus (MATH 19520 or 20000, or equivalent).* This course covers linear algebra, with special attention to topics useful in statistical applications. In addition to addressing theoretical and algorithmic aspects of solving systems of linear equations, topics may include least squares, orthogonal projections, positive-definite matrices, quadratic forms, matrix decompositions, and an introduction to vector spaces. Computers are used to study some computational issues and mathematical explorations. *Autumn.*

24400-24500. Statistical Theory and Methods I, II. *PQ: Multivariate calculus (MATH 19520 or 20000, or equivalent) and linear algebra (MATH 19620, 25500 or STAT 24300 or equivalent).* Some previous experience with statistics helpful but not required. This course is a systematic introduction to the principles and techniques of statistics, as well as to practical considerations in the analysis of data, with emphasis on the analysis of experimental data. The first quarter covers tools from probability and the elements of statistical theory. Topics include the definitions of probability and random variables, binomial and other discrete probability distributions, normal and other continuous probability distributions, joint probability distributions and the transformation of random variables, principles of inference (including Bayesian inference), maximum likelihood estimation, hypothesis testing and confidence intervals, likelihood ratio tests, multinomial distributions, and chi-square tests. Examples are drawn from the social, physical, and biological sciences. The coverage of topics in probability is limited and brief, so students who have taken a course in probability find reinforcement rather than redundancy. The second quarter covers statistical methodology, including the analysis of variance, regression, correlation, and some multivariate analysis. Some principles of data analysis are introduced, and an attempt is made to present the analysis of variance and regression in a unified framework. Computers are used in the second quarter. NOTE: The Autumn/Winter sequence typically is slightly more advanced than the Winter/Spring sequence. *Autumn, Winter; Winter, Spring.*

24600. Complex Statistical Problems. *PQ: STAT 24400-24500. Knowledge of probability distributions, random variables, and estimation techniques (e.g., maximum likelihood at the level of STAT 24400-24500).* Topics vary from year to year. Recently, this course has treated the impact of missing data on statistical analyses (e.g., probability models and methods of estimation and inference); algorithms for iterative maximum likelihood estimation (e.g., the Expectation-Maximization [EM] and Newton-Raphson algorithms); and Bayesian computation (e.g., Data Augmentation and Monte Carlo Markov Chain methods). *Spring.*

25100. Introduction to Mathematical Probability. *PQ: MATH 20000 or 20500, or consent of instructor.* This course covers fundamentals and axioms; combinatorial probability; conditional probability and independence; binomial, Poisson, and normal distributions; the law of large numbers and the central limit theorem; and random variables and generating functions. *Spring.*

25300/31700. Introduction to Probability Models. *PQ: STAT 24400 or 25100.* This course introduces stochastic processes as models for a variety of phenomena in the physical and biological sciences. Following a brief review of basic concepts in probability, we introduce stochastic processes that are popular in applications in sciences (e.g., discrete time Markov chain, the Poisson process, continuous time Markov process, renewal process and Brownian motion). *Winter.*

26100. Time Dependent Data. *PQ: STAT 24400-24500 or 22400, or consent of instructor.* This course considers the modeling and analysis of data that are

ordered in time. The main focus is on quantitative observations taken at evenly spaced intervals and includes both time-domain and spectral approaches. *Spring*.

26700/36700. History of Statistics. (=CHSS 32900, HIPS 25600) *PQ: Prior statistics course.* This course covers topics in the history of statistics, from the eleventh century to the middle of the twentieth century. We focus on the period from 1650 to 1950, with an emphasis on the mathematical developments in the theory of probability and how they came to be used in the sciences. Our goals are both to quantify uncertainty in observational data and to develop a conceptual framework for scientific theories. This course includes broad views of the development of the subject and closer looks at specific people and investigations, including reanalyses of historical data. *S. Stigler. Spring.*

29700. Undergraduate Research. *PQ: Consent of faculty adviser and departmental counselor. Students are required to submit the College Reading and Research Course Form. Open to all students, including nonmajors. May be taken either for quality grades or for P/F grading; however, students who wish to count this course toward the requirements for a major in statistics must receive prior approval of the departmental counselor and must receive a quality grade.* This course consists of reading and research in an area of statistics or probability under the guidance of a faculty member. A written report must be submitted at the end of the quarter. *Autumn, Winter, Spring.*

29900. Bachelor's Paper. *PQ: Consent of faculty adviser and departmental counselor. Students are required to submit the College Reading and Research Course Form. Open only to students who are majoring in statistics. May be taken for P/F grading. Credit for STAT 29900 may not be counted toward the twelve courses required for a major in statistics.* This course consists of reading and research in an area of statistics or probability under the guidance of a faculty member, leading to a bachelor's paper. The paper must be submitted at the end of the quarter. *Autumn, Winter, Spring.*

The following list of 30000-level courses may be of interest to students with advanced standing who are majoring in statistics. For more information, consult the departmental counselor. For a complete listing and updates, visit www.stat.uchicago.edu.

30100-30200. Mathematical Statistics I, II. *PQ: STAT 30400 or consent of instructor.* This course surveys the mathematical structure of modern statistics. Topics include statistical models, methods for parameter estimation, comparison of estimators, large sample theory, efficiency, confidence sets, theory of hypothesis tests, elements of linear hypothesis theory, analysis of discrete data, and an introduction to Bayesian analysis. *Winter, Spring.*

30400. Distribution Theory. *PQ: STAT 24500 and MATH 20500, or consent of instructor.* This course is a systematic introduction to random variables and probability distributions. Topics include standard distributions (i.e., uniform, normal, beta, gamma, F , t , Cauchy, Poisson, binomial, and hypergeometric);

moments and cumulants; characteristic functions; exponential families; modes of convergence; central limit theorem; and Laplace's method. *Autumn.*

30900. Mathematical Computation I: Matrix Computation Course. (=CMSC 37810) *PQ: Linear algebra (STAT 24300 or equivalent) and some previous experience with statistics.* This course covers the theory and practice of matrix computation, starting with the LU and Cholesky decompositions, the QR decompositions with applications to least squares, iterative methods for solving eigenvalue problems, iterative methods for solving large systems of equations, and (time permitting) the basics of the fast Fourier and fast wavelet transforms. The mathematical theory underlying the algorithms is emphasized, as well as their implementation in code. *Autumn.*

31000. Mathematical Computation II: Optimization and Simulation. *PQ: Solid grounding in multivariate calculus, linear algebra, and probability theory.* This course covers the fundamentals of continuous optimization, including constrained optimization, and introduces the use of Monte Carlo methods in computer simulation and combinatorial optimization problems. Several substantial programming projects (using MATLAB) are completed during the course. *Winter.*

31100. Mathematical Computation III: Numerical Methods for PDE's. *PQ: Some prior exposure to differential equations and linear algebra.* The first part of this course introduces basic properties of PDE's; finite difference discretizations; and stability, consistency, convergence, and Lax's equivalence theorem. We also cover examples of finite difference schemes; simple stability analysis; convergence analysis and order of accuracy; consistency analysis and errors (i.e., dissipative and dispersive errors); and unconditional stability and implicit schemes. The second part of this course includes solution of stiff systems in 1, 2, and 3D; direct vs. iterative methods (i.e., banded and sparse LU factorizations); and Jacobi, Gauss-Seidel, multigrid, conjugate gradient, and GMRES iterations. *Spring.*

31200. Introduction to Stochastic Processes I. *PQ: STAT 25100 and MATH 20500; STAT 30400 or consent of instructor.* This course introduces stochastic processes not requiring measure theory. Topics include branching processes, recurrent events, renewal theory, random walks, Markov chains, Poisson, and birth-and-death processes. *Winter.*

31300. Introduction to Stochastic Processes II. *PQ: STAT 31200.* Topics include continuous-time Markov chains, Markov chain Monte Carlo, discrete-time martingales, and Brownian motion and diffusions. Our emphasis is on defining the processes and calculating or approximating various related probabilities. The measure theoretic aspects of these processes are not covered rigorously. *Spring.*

33100. Sample Surveys. *PQ: Consent of instructor.* This course covers random sampling methods; stratification, cluster sampling, and ratio estimation; and methods for dealing with nonresponse and partial response. *Autumn.*

34300. Applied Linear Statistical Methods. *PQ: STAT 24500 or equivalent, and linear algebra (STAT 24300 or equivalent).* This course introduces the theory, methods, and applications of fitting and interpreting multiple regression models. Topics include the examination of residuals, the transformation of data, strategies and criteria for the selection of a regression equation, nonlinear models, biases due to excluded variables and measurement error, and the use and interpretation of computer package regression programs. The theoretical basis of the methods, the relation to linear algebra, and the effects of violations of assumptions are studied. Techniques discussed are illustrated by examples involving both physical and social sciences data. *Autumn.*

34500. Design and Analysis of Experiments. *PQ: STAT 34300.* This course introduces the methodology and application of linear models in experimental design. We emphasize the basic principles of experimental design (e.g., blocking, randomization, incomplete layouts). Many of the standard designs (e.g., fractional factorial, incomplete block, split unit designs) are studied within this context. The analysis of these experiments is developed as well, with particular emphasis on the role of fixed and random effects. Additional topics may include response surface analysis, the use of covariates in the analysis of designed experiments, and spatial analysis of field trials. *Winter.*

34700. Generalized Linear Models. *PQ: STAT 34300 or consent of instructor.* This applied course covers factors, variates, contrasts, and interactions; exponential-family models (i.e., variance function); definition of a generalized linear model (i.e., link functions); specific examples of GLMs; logistic and probit regression; cumulative logistic models; log-linear models and contingency tables; inverse linear models; Quasi-likelihood and least squares; estimating functions; and partially linear models. *Spring.*

35000. Principles of Epidemiology. (=BIOS 29318, ENST 27400, HSTD 30900, PPHA 36400) *Introductory statistics recommended.* Epidemiology is the study of the distribution and determinants of health and disease in human populations. This course introduces the basic principles of epidemiologic study design, analysis, and interpretation, through lectures, assignments, and critical appraisal of both classic and contemporary research articles. The course objectives include: (1) to be able to critically read and understand epidemiologic studies; (2) to be able to calculate and interpret measures of disease occurrence and measures of disease-exposure associations; and (3) to understand the contributions of epidemiology to clinical research, medicine, and public health. *Autumn.*

35600. Applied Survival Analysis. (=HSTD 33100) *PQ: HSTD 32100/STAT 22000, introductory statistics, or consent of instructor.* This course introduces principles and methods for the analysis of time-to-event data. This type of data occurs extensively in both observational and experimental biomedical and public health studies, as well as in industrial applications. While some theoretical statistical detail is given (at the level appropriate for a master's student in statistics), we primarily focus on data analysis. Problems are motivated from an

epidemiologic and clinical perspective, concentrating on the analysis of cohort data and time-to-event data from controlled clinical trials. *Autumn.*

36900. Longitudinal Data Analysis. (=HSTD 33300) *PQ: STAT 22600 or 22400 (or equivalent), and STAT 22700 or 34700 (or equivalent); or consent of instructor.* Longitudinal data consist of multiple measures over time on a sample of individuals. This type of data occurs extensively in both observational and experimental biomedical and public health studies, as well as in studies in sociology and applied economics. This course introduces principles and methods for the analysis of longitudinal data. Although some supporting statistical theory is given, we emphasize data analysis and interpretation of models for longitudinal data. Problems are motivated by applications in epidemiology, clinical medicine, health services research, and disease natural history studies. *Autumn.*

38100. Measure-Theoretic Probability I. *PQ: STAT 31300 or consent of instructor.* This course provides a detailed, rigorous treatment of probability from the point of view of measure theory, as well as existence theorems, integration and expected values, characteristic functions, moment problems, limit laws, Radon-Nikodym derivatives, and conditional probabilities. *Autumn.*

38300. Measure-Theoretic Probability III. *PQ: STAT 38100.* This course continues material covered in STAT 38100, with topics that include L_p spaces, Radon-Nikodym theorem, conditional expectation, and martingale theory. *Winter.*