

# Statistics

---

Department Website: <http://www.stat.uchicago.edu>

## Program of Study

The modern science of statistics involves the development of principles and methods for modeling uncertainty, for designing experiments, surveys, and observational programs, and for analyzing and interpreting empirical data. Mathematics plays a major role in all areas of statistics, from probability theory to data analysis. Statistics is an appropriate field for students with strong mathematical and computational skills and an interest in applying these skills to problems in the natural and social sciences. A program leading to the bachelor's degree in Statistics offers coverage of the principles and methods of statistics in combination with solid training in mathematics and some additional training in computation. The major can provide appropriate preparation 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 Adviser for Majors 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 and are encouraged to discuss their course choices with the Departmental Adviser for Minors. Information on the minor follows the description of the major.

## General Course Information

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, in mathematical level, and in the direction of applications.

## Introductory Courses and Sequences

To begin their studies in statistics, students can choose from several courses. Students and College advisers are encouraged to contact the Departmental Adviser for Introductory Courses for advice on choosing an appropriate first course.

For students with little or no math background who do not intend to continue on to more advanced statistics courses, STAT 20000 Elementary Statistics is an introductory course that emphasizes concepts rather than statistical techniques. STAT 20000 Elementary Statistics may not be taken by students with credit for STAT 22000 Statistical Methods and Applications, STAT 23400 Statistical Models and Methods, or more advanced courses in the Department of Statistics. STAT 20000 Elementary Statistics does not count toward the major or minor in Statistics.

Students with at least MATH 13100 Elem Functions and Calculus I or placement into MATH 15100 Calculus I are encouraged to take STAT 22000 Statistical Methods and Applications instead of STAT 20000 Elementary Statistics. Students with three quarters of calculus may choose either STAT 22000 Statistical Methods and Applications or STAT 23400 Statistical Models and Methods. Students may count either STAT 22000 Statistical Methods and Applications or STAT 23400 Statistical Models and Methods, but not both, toward the forty-two credits required for graduation.

STAT 22000 Statistical Methods and Applications 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. A score of 5 on the AP Statistics exam yields credit for STAT 22000 Statistical Methods and Applications, although this credit will not count toward the requirements for a major or minor in Statistics. STAT 22000 Statistical Methods and Applications can count toward the minor in Statistics, but for students matriculating in Autumn Quarter 2016 and after, cannot count toward the major in Statistics.

STAT 23400 Statistical Models and Methods covers much of the same material as STAT 22000 Statistical Methods and Applications, but at a somewhat higher mathematical level. The course is a one-quarter introduction to statistics that is appropriate for any student with a good command of univariate calculus including sequences and series. STAT 23400 Statistical Models and Methods can count toward the minor in Statistics, but for students matriculating in Autumn Quarter 2016 and after, cannot count toward the major in Statistics.

Students cannot hold credit for both STAT 22000 Statistical Methods and Applications and STAT 23400 Statistical Models and Methods. Students completing either STAT 22000 Statistical Methods and Applications or STAT 23400 Statistical Models and Methods forego their AP Statistics credit for STAT 22000 Statistical Methods and Applications.

STAT 24400 Statistical Theory and Methods I-STAT 24500 Statistical Theory and Methods II is recommended for students who wish to have a thorough introduction to statistical theory and methodology. STAT 24400 Statistical Theory and Methods I-STAT 24500 Statistical Theory and Methods II is more mathematically demanding than either STAT 22000 Statistical Methods and Applications or STAT 23400 Statistical Models and Methods. STAT 24400 Statistical Theory and Methods I assumes some familiarity with multivariate calculus, and STAT 24500 Statistical Theory and Methods II assumes some familiarity with linear algebra.

STAT 24410 Statistical Theory and Methods Ia-STAT 24510 Statistical Theory and Methods IIa is an alternative version of STAT 24400 Statistical Theory and Methods I-STAT 24500 Statistical Theory and Methods II that requires STAT 25100 Introduction to Mathematical Probability (or STAT 25150 Introduction to Mathematical Probability-A) as a prerequisite and that replaces some probability topics with additional statistical topics not normally covered in STAT

24400 Statistical Theory and Methods I-STAT 24500 Statistical Theory and Methods II. STAT 24410 Statistical Theory and Methods Ia-STAT 24510 Statistical Theory and Methods IIa is particularly well-suited for students with a strong mathematical background who are interested in more extensive coverage of probability and statistics. Students may count either STAT 24400 Statistical Theory and Methods I or STAT 24410 Statistical Theory and Methods Ia, but not both, toward the 4200 units of credit required for graduation. Similarly, students may count either STAT 24500 Statistical Theory and Methods II or STAT 24510 Statistical Theory and Methods IIa, but not both, and they may count STAT 25100 Introduction to Mathematical Probability or STAT 25150 Introduction to Mathematical Probability-A, but not both, toward the 4200 units of credits required for graduation.

Students considering a major in Statistics are encouraged to begin with either STAT 24400 Statistical Theory and Methods I-STAT 24500 Statistical Theory and Methods II or with the alternative sequence consisting of STAT 25100 Introduction to Mathematical Probability and STAT 24410 Statistical Theory and Methods Ia-STAT 24510 Statistical Theory and Methods IIa, rather than with STAT 23400 Statistical Models and Methods. Although students with a strong mathematical background can and do take either STAT 24400 Statistical Theory and Methods I-STAT 24500 Statistical Theory and Methods II or the alternative sequence (STAT 25100 Introduction to Mathematical Probability and STAT 24410 Statistical Theory and Methods Ia-STAT 24510 Statistical Theory and Methods IIa) without prior course work in statistics or probability, some students find it helpful to take either STAT 22000 Statistical Methods and Applications or STAT 23400 Statistical Models and Methods as preparation.

The core of the Statistics major consists of three courses: STAT 25100 Introduction to Mathematical Probability and either STAT 24400 Statistical Theory and Methods I-STAT 24500 Statistical Theory and Methods II or STAT 24410 Statistical Theory and Methods Ia-STAT 24510 Statistical Theory and Methods IIa. Either of these is recommended as a three-quarter cognate sequence for students in the quantitative sciences and mathematics. Note that STAT 25100 Introduction to Mathematical Probability may be taken before, after, or concurrently with STAT 24400 Statistical Theory and Methods I-STAT 24500 Statistical Theory and Methods II, though it is a prerequisite for STAT 24410 Statistical Theory and Methods Ia-STAT 24510 Statistical Theory and Methods IIa.

### Additional Courses in Statistical Theory, Methods, and Applications

For students interested in continuing their study of statistics beyond the introductory level, STAT 22200 Linear Models And Experimental Design, STAT 22400 Applied Regression Analysis, STAT 22600 Analysis of Categorical Data, STAT 22700 Biostatistical Methods, and STAT 26700 History of Statistics are recommended. Note that because there is some overlap between STAT 22600 Analysis of Categorical Data and STAT 22700 Biostatistical Methods, only one of these two courses, not both, may be counted toward a major or minor in Statistics. The courses STAT 22200 Linear Models And Experimental Design, STAT 22400 Applied Regression Analysis, STAT 22600 Analysis of Categorical Data, and STAT 26700 History of Statistics may be taken in any order. Each presumes two quarters of calculus (except STAT 26700 History of Statistics) and a previous course in statistics (STAT 22000 Statistical Methods and Applications or higher). STAT 22700 Biostatistical Methods has STAT 22400 Applied Regression Analysis as a prerequisite.

For students who have completed STAT 24400 Statistical Theory and Methods I-STAT 24500 Statistical Theory and Methods II and are interested in more advanced statistical methodology courses, STAT 24620 Multivariate Statistical Analysis: Applications and Techniques, STAT 26100 Time Dependent Data, STAT 27400 Nonparametric Inference, STAT 27850 Multiple Testing, Modern Inference, and Replicability, and STAT 34300 Applied Linear Stat Methods are recommended. Many other graduate courses in Statistics offer opportunities for further study of statistical theory, methods, and applications. For details, consult the instructor or the Departmental Adviser for Majors, or visit the Graduate Announcements (<http://graduateannouncements.uchicago.edu/departmentofstatistics>).

### Courses in Probability

Students interested in probability can begin with STAT 25100 Introduction to Mathematical Probability, which can be taken separately from any Statistics courses and can be supplemented with more advanced probability courses, such as STAT 25300 Introduction to Probability Models or MATH 23500 Markov Chains, Martingales, and Brownian Motion. Students with a strong mathematical background can take STAT 31200 Introduction to Stochastic Processes I, STAT 38100 Measure-Theoretic Probability I, and STAT 38300 Measure-Theoretic Probability III. Note that because there is some overlap between MATH 23500 Markov Chains, Martingales, and Brownian Motion and STAT 31200 Introduction to Stochastic Processes I, only one of these two courses, not both, may be counted toward a major in Statistics.

### Courses in Machine Learning

A student with a strong computer science background and some knowledge of elementary statistics could take STAT 27725 Machine Learning. Other courses in the category of machine learning include the advanced statistical methodology courses STAT 24620 Multivariate Statistical Analysis: Applications and Techniques and STAT 27400 Nonparametric Inference. Graduate course offerings in machine learning include STAT 37601 Machine Learning and Large-Scale Data Analysis and STAT 37710 Machine Learning.

### Courses in Optimization

A student with a strong mathematical background could take STAT 28000 Optimization. Graduate course offerings in optimization include STAT 31015 Convex Optimization and STAT 31020 Mathematical Computation IIB: Nonlinear Optimization.

## Grading

Students who are majoring or minoring in Statistics must receive a quality grade of at least C in all of the courses counted toward their major or minor program in Statistics. In addition, students who are majoring in Statistics must receive quality grades of at least C+ in both STAT 24400 Statistical Theory and Methods I and STAT 24500 Statistical Theory and Methods II (or at least C in both STAT 24410 Statistical Theory and Methods Ia and STAT 24510 Statistical Theory and Methods IIa). Subject to College and divisional regulations, and with the consent of the instructor, students may register for either quality grades or for P/F grading in any 20000-level Statistics course that is not counted toward a major or minor in Statistics. A grade of P is given only for work of C– quality or higher.

The following policy applies to students who wish to receive a mark of I for a Statistics course. In addition to submitting the official Incomplete Form required by the College, 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 the remaining course work by the end of the quarter due to an emergency. Students requesting a mark of I for STAT 20000 Elementary Statistics, STAT 22000 Statistical Methods and Applications, or STAT 23400 Statistical Models and Methods must obtain approval from both the current instructor and the Departmental Adviser for Introductory Courses.

## Program Requirements for Majors

The requirements for the BA and BS in Statistics were updated in 2017. Students who matriculated prior to Autumn 2017 may choose to follow these updated requirements; otherwise, they should consult the archived catalog from their year of matriculation (or, at their option, any later year) for the degree requirements in Statistics. All students who matriculated in Autumn 2017 or later should follow the updated requirements described below.

Every candidate must obtain approval of his or her course program from the Departmental Adviser for Majors. Students majoring in Statistics should meet the general education requirement in mathematical sciences with courses in calculus. The major program includes four additional prescribed mathematics courses, four prescribed statistics courses, and two prescribed computer science courses. Students are advised to complete the four mathematics courses by the end of their third year. Additional requirements include four approved elective courses in Statistics. The BS also requires an additional prescribed mathematics course and an approved, coherent, three-quarter sequence at the 20000 level in a field to which statistics can be applied. Students who are majoring in Statistics must receive quality grades of at least C+ in both STAT 24400 Statistical Theory and Methods I and STAT 24500 Statistical Theory and Methods II (or at least C in both STAT 24410 Statistical Theory and Methods Ia and STAT 24510 Statistical Theory and Methods IIa), and at least C in all other courses counted toward the Statistics major. A grade of P is not acceptable for any of these courses.

## Prescribed Mathematics Courses

The prescribed mathematics courses include a Calculus III requirement (MATH 13300 Elementary Functions and Calculus III or MATH 15300 Calculus III or MATH 16300 Honors Calculus III) and a Linear Algebra requirement (STAT 24300 Numerical Linear Algebra or MATH 20250 Abstract Linear Algebra). Note that MATH 19620 Linear Algebra may not be used to meet the Linear Algebra requirement.

For the BA, one of the following pairs of courses is required: MATH 20000 Mathematical Methods for Physical Sciences I-MATH 20100 Mathematical Methods for Physical Sciences II or MATH 20400 Analysis in Rn II-MATH 20500 Analysis in Rn III or MATH 20800 Honors Analysis in Rn II-MATH 20900 Honors Analysis in Rn III or the pair consisting of MATH 20000 Mathematical Methods for Physical Sciences I and STAT 28200 Dynamical Systems with Applications.

For the BS, students must take one of the following three courses: MATH 20000 Mathematical Methods for Physical Sciences I or MATH 20500 Analysis in Rn III or MATH 20900 Honors Analysis in Rn III; and, in addition, one of the following three courses: MATH 20100 Mathematical Methods for Physical Sciences II, MATH 27300 Basic Theory of Ordinary Differential Equations, or STAT 28200 Dynamical Systems with Applications; and, in addition, one of the following two courses: STAT 28000 Optimization or MATH 21100 Basic Numerical Analysis.

Students who are completing majors in both Statistics and Economics should follow the same mathematics requirements as Statistics majors. Students who have already taken MATH 19520 Mathematical Methods for Social Sciences and MATH 19620 Linear Algebra should discuss with the Departmental Adviser for Majors how best to meet the mathematics requirements for the Statistics major. For example, such students can petition to meet the requirements for the BA in Statistics by taking all three of MATH 20100 Mathematical Methods for Physical Sciences II, STAT 24300 Numerical Linear Algebra, and STAT 28200 Dynamical Systems with Applications.

## Prescribed Statistics Courses

The four prescribed Statistics courses are STAT 25100 Introduction to Mathematical Probability, STAT 24400 Statistical Theory and Methods I-STAT 24500 Statistical Theory and Methods II (or STAT 24410 Statistical Theory and Methods Ia-STAT 24510 Statistical Theory and Methods IIa), and either STAT 22400 Applied Regression Analysis or STAT 34300 Applied Linear Stat Methods.

It is recommended that students who have had some multivariable calculus begin the major by taking either STAT 25100 Introduction to Mathematical Probability or STAT 24400 Statistical Theory and Methods I as their first course in probability and statistics. An alternative route to beginning the major would be to first take either STAT 22000 Statistical Methods and Applications or STAT 23400 Statistical Models and Methods, neither of which count toward the major, but which could serve as a prerequisite for courses such as STAT 22400 Applied Regression Analysis, STAT 22200 Linear Models And Experimental Design, and STAT 22600 Analysis of Categorical Data, which do count toward the major. This

second path is recommended for students who need additional time to complete multivariable calculus and linear algebra prerequisites and who want to get started on the major in the meantime.

#### Electives

Candidates for the BA are required to take four electives, at least two of which must be on List B below. The remaining two electives may be from either List B or C. Students may count either STAT 22600 Analysis of Categorical Data or STAT 22700 Biostatistical Methods, but not both, toward the BA. Similarly, students may count either MATH 23500 Markov Chains, Martingales, and Brownian Motion or STAT 31200 Introduction to Stochastic Processes I, but not both, toward the BA.

Candidates for the BS are required to take four electives. A candidate for the BS who has *not* taken STAT 34300 Applied Linear Stat Methods as one of the four prescribed statistics courses must take at least one elective from List A below, a second elective from List B, and the remaining two electives may be from either List B or C. A candidate for the BS who *has* taken STAT 34300 Applied Linear Stat Methods as one of the four prescribed statistics courses must take at least two electives from List B and the remaining two electives may be from either List B or C. For the BS in Statistics, STAT 28000 Optimization counts as a List C elective only if MATH 21100 Basic Numerical Analysis is also included in the program. In other words, students cannot double-count STAT 28000 Optimization toward both the four-elective requirement and the requirement to take one of STAT 28000 Optimization and MATH 21100 Basic Numerical Analysis. Students may count either STAT 22600 Analysis of Categorical Data or STAT 22700 Biostatistical Methods, but not both, toward the BS. Similarly, students may count either MATH 23500 Markov Chains, Martingales, and Brownian Motion or STAT 31200 Introduction to Stochastic Processes I, but not both, toward the BS.

Note: The following lists may change from time to time as courses change and new courses are added. Please consult the Departmental Adviser for Majors for approval of your electives.

#### LIST A: Advanced Statistical Methodology

STAT 24620	Multivariate Statistical Analysis: Applications and Techniques
STAT 26100	Time Dependent Data
STAT 27400	Nonparametric Inference
STAT 27850	Multiple Testing, Modern Inference, and Replicability
Some additional graduate courses in Statistics (must be approved by Departmental Adviser for Majors)	

#### LIST B: Statistical Methodology

STAT 22200	Linear Models And Experimental Design
STAT 22600	Analysis of Categorical Data *
STAT 22700	Biostatistical Methods *
STAT 24620	Multivariate Statistical Analysis: Applications and Techniques
STAT 26100	Time Dependent Data
STAT 26700	History of Statistics
STAT 27400	Nonparametric Inference
STAT 27850	Multiple Testing, Modern Inference, and Replicability
STAT 35800	Statistical Applications
STAT 37601	Machine Learning and Large-Scale Data Analysis
Some additional graduate courses in Statistics (must be approved by Departmental Adviser for Majors)	

\* Students may count either STAT 22600 Analysis of Categorical Data or STAT 22700 Biostatistical Methods, but not both, toward the major.

#### LIST C: Other Upper Level/Graduate Courses

MATH 23500	Markov Chains, Martingales, and Brownian Motion *
STAT 25300	Introduction to Probability Models
STAT 27725	Machine Learning
STAT 28000	Optimization **
STAT 30900	Mathematical Computation I: Matrix Computation Course
STAT 31015	Convex Optimization
STAT 31020	Mathematical Computation IIB: Nonlinear Optimization
STAT 31060	Further Mathematical Computation: Matrix Computation and Optimization
STAT 31200	Introduction to Stochastic Processes I *
STAT 37710	Machine Learning
STAT 38100	Measure-Theoretic Probability I

STAT 38300 Measure-Theoretic Probability III

Some additional graduate courses in Statistics (must be approved by Departmental Adviser for Majors)

- \* Students may count either MATH 23500 Markov Chains, Martingales, and Brownian Motion or STAT 31200 Introduction to Stochastic Processes I, but not both, toward the major.
- \*\* For the BA in Statistics, STAT 28000 Optimization counts as a List C elective. For the BS in Statistics, STAT 28000 Optimization counts as a List C elective only if MATH 21100 Basic Numerical Analysis is also included in the program. In other words, for the BS, students cannot double-count STAT 28000 Optimization toward both the four-elective requirement and the requirement to take at least one of STAT 28000 Optimization and MATH 21100 Basic Numerical Analysis.

**Computer Science Requirement**

Candidates for either the BA or the BS are required to take one of the following sequences: CMSC 12100-CMSC 12200 or CMSC 15100-CMSC 15200 or CMSC 16100-CMSC 16200.

**BS Requirement of Three-Quarter Sequence in a Field to Which Statistics Can Be Applied**

Candidates for the BS (but not the BA) are required to take an approved, coherent, three-quarter sequence at the 20000 level in a field to which statistics can be applied. Generally this sequence should be in the natural or social sciences, but a sequence in another discipline may be acceptable. Courses in MATH or CMSC may not be used for this requirement. Sequences in which earlier courses are prerequisites for later ones are preferred. Example sequences include BIOS 20198 Biodiversity-BIOS 20196 Ecology and Conservation-BIOS 23406 Biogeography; CHEM 22000-CHEM 22100-CHEM 22200; CHEM 26100-CHEM 26200-CHEM 26300; ECON 20000-ECON 20100-ECON 20200; GEOS 21000 Mineralogy-GEOS 21100 Introduction to Petrology-GEOS 21200 Physics of the Earth; and PHYS 23400-PHYS 23500 and PHYS 23700 Nuclei and Elementary Particles. All sequences must be approved by the Departmental Adviser for Majors.

**Summary of Requirements for the BA in Statistics**

**GENERAL EDUCATION**

One of the following sequences: *	200
MATH 13100-13200 Elementary Functions and Calculus I-II	
MATH 15100-15200 Calculus I-II	
MATH 16100-16200 Honors Calculus I-II	

Total Units 200

**MAJOR**

One of the following: *	100
MATH 13300 Elementary Functions and Calculus III	
MATH 15300 Calculus III	
MATH 16300 Honors Calculus III	

One of the following course pairs:	200
MATH 20000 Mathematical Methods for Physical Sciences I	

&	
STAT 28200 Dynamical Systems with Applications	
MATH 20000-20100 Mathematical Methods for Physical Sciences I-II	
MATH 20400-20500 Analysis in Rn II-III	
MATH 20800-20900 Honors Analysis in Rn II-III	

One of the following:	100
STAT 24300 Numerical Linear Algebra	
MATH 20250 Abstract Linear Algebra	

One of the following sequences:	200
STAT 24400-24500 Statistical Theory and Methods I-II	
STAT 24410-24510 Statistical Theory and Methods Ia-IIa	

One of the following:	100
STAT 25100 Introduction to Mathematical Probability	
STAT 25150 Introduction to Mathematical Probability-A	

One of the following:	100
STAT 22400 Applied Regression Analysis	
STAT 34300 Applied Linear Stat Methods	

One of the following sequences:	200
CMSC 12100-12200 Computer Science with Applications I-II	

CMSC 15100-15200	Introduction to Computer Science I-II	
CMSC 16100-16200	Honors Introduction to Computer Science I-II	
Four approved elective courses in Statistics **		400
Total Units		1400

\* Credit may be granted by examination.

\*\* At least two of the electives must be on List B. The remaining two electives may be from either List B or C. Students may count either STAT 22600 Analysis of Categorical Data or STAT 22700 Biostatistical Methods, but not both, toward the BA. Students may count either MATH 23500 Markov Chains, Martingales, and Brownian Motion or STAT 31200 Introduction to Stochastic Processes I, but not both, toward the BA.

### Summary of Requirements for the BS in Statistics

#### GENERAL EDUCATION

One of the following sequences: *		200
MATH 13100-13200	Elementary Functions and Calculus I-II	
MATH 15100-15200	Calculus I-II	
MATH 16100-16200	Honors Calculus I-II	
Total Units		200

#### MAJOR

One of the following: *		100
MATH 13300	Elementary Functions and Calculus III	
MATH 15300	Calculus III	
MATH 16300	Honors Calculus III	
One of the following:		100
MATH 20000	Mathematical Methods for Physical Sciences I	
MATH 20500	Analysis in $R^n$ III	
MATH 20900	Honors Analysis in $R^n$ III	
One of the following:		100
MATH 20100	Mathematical Methods for Physical Sciences II	
MATH 27300	Basic Theory of Ordinary Differential Equations	
STAT 28200	Dynamical Systems with Applications	
One of the following:		100
STAT 24300	Numerical Linear Algebra	
MATH 20250	Abstract Linear Algebra	
One of the following:		100
STAT 28000	Optimization	
MATH 21100	Basic Numerical Analysis	
One of the following sequences:		200
STAT 24400-24500	Statistical Theory and Methods I-II	
STAT 24410-24510	Statistical Theory and Methods Ia-IIa	
One of the following:		100
STAT 25100	Introduction to Mathematical Probability	
STAT 25150	Introduction to Mathematical Probability-A	
One of the following:		100
STAT 22400	Applied Regression Analysis	
STAT 34300	Applied Linear Stat Methods	
One of the following sequences:		200
CMSC 12100-12200	Computer Science with Applications I-II	
CMSC 15100-15200	Introduction to Computer Science I-II	
CMSC 16100-16200	Honors Introduction to Computer Science I-II	
Four approved elective courses in Statistics **		400
A coherent three-quarter sequence at the 20000 level in a field to which statistics can be applied ***		300
Total Units		1800

- \* Credit may be granted by examination.
- \*\* A candidate for the BS who has *not* taken STAT 34300 Applied Linear Stat Methods as one of the four prescribed statistics courses must take at least one elective from List A, a second elective from List B, and the remaining two electives may be from either List B or C. A candidate for the BS who *has* taken STAT 34300 Applied Linear Stat Methods as one of the four prescribed statistics courses must take at least two electives from List B and the remaining two electives may be from either List B or C. For the BS in Statistics, STAT 28000 Optimization counts as a List C elective only if MATH 21100 Basic Numerical Analysis is also included in the program. In other words, students cannot double-count STAT 28000 Optimization toward both the four-elective requirement and the requirement to take at least one of STAT 28000 Optimization and MATH 21100 Basic Numerical Analysis. Students may count either STAT 22600 Analysis of Categorical Data or STAT 22700 Biostatistical Methods, but not both, toward the BS. Students may count either MATH 23500 Markov Chains, Martingales, and Brownian Motion or STAT 31200 Introduction to Stochastic Processes I, but not both, toward the BS.
- \*\*\* Generally, this sequence should be in the natural or social sciences, but a sequence in another discipline may be acceptable. Courses in MATH or CMSC may not be used for this requirement. Sequences in which earlier courses are prerequisites for later ones are preferred. Example sequences include BIOS 20198 Biodiversity-BIOS 20196 Ecology and Conservation-BIOS 23406 Biogeography; CHEM 22000-22100-22200 Organic Chemistry I-II-III; CHEM 26100-26200-26300 Quantum Mechanics; Thermodynamics; Chemical Kinetics and Dynamics; ECON 20000-20100-20200 The Elements of Economic Analysis I-II-III; GEOS 21000 Mineralogy-GEOS 21100 Introduction to Petrology-GEOS 21200 Physics of the Earth; and PHYS 23400-23500 Quantum Mechanics I-II-PHYS 23700 Nuclei and Elementary Particles. All sequences must be approved by the Departmental Adviser for Majors.

### Honors

The BA or BS 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 courses in the major and also complete an approved honors paper (STAT 29900 Bachelor's Paper). This paper is typically based on a structured research program that the student undertakes, with faculty supervision, in the first quarter of his or her fourth year. Eligible students who wish to be considered for honors should consult the Departmental Adviser for Majors before the end of their third year. The research paper or project used to meet this requirement may not be used to meet the bachelor's paper or project requirement in another major or course. NOTE: Credit for STAT 29900 Bachelor's Paper will not count towards the courses required for a major in Statistics.

### Joint BA/MS or BS/MS in Statistics

This program enables unusually well-qualified undergraduate students to complete an MS in Statistics along with a BA or BS during their four years at the College. Although a student may receive a BA or BS in any field, a program of study other than Statistics is recommended.

Only a small number of students will be selected for the program through a competitive admissions process. Participants must apply to the MS program in Statistics by June 1 of their third year for admission to candidacy for an MS in Statistics during their fourth year. To be considered, students should have completed almost all of their undergraduate requirements, including all of their general education and language competence requirements, by the end of their third year. They should also have completed, at a minimum, STAT 24400-24500 Statistical Theory and Methods I-II (or STAT 24410-24510 Statistical Theory and Methods Ia-Ia) with A or A- grades and all the mathematics requirements for the Statistics major with very high grades. While these are the minimum criteria, admission is competitive, and additional qualifications may be needed. Interested students are strongly encouraged to consult both the Departmental Adviser for Majors and their College adviser early in their third year.

Participants in the joint BA/MS or BS/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 24410-24510 Statistical Theory and Methods Ia-Ia and STAT 34300 Applied Linear Stat Methods, which satisfy requirements for the MS in Statistics, could also be used to satisfy requirements of a BA or BS program in Statistics.

Other requirements include a master's paper and participation in the Consulting Program of the Department of Statistics. For details, visit the Department of Statistics Admissions page (<http://www.stat.uchicago.edu/admissions/> featured).

### Minor Program in Statistics

The Statistics minor focuses on statistical methodology, in contrast to the Statistics major, which has a substantial theoretical component. The minor in Statistics requires five courses, some prescribed and some elective, chosen in consultation with the Departmental Adviser for Minors. Not every combination of elective courses is allowed. Generally, no more than two electives may be satisfied by courses offered by departments other than the Department of Statistics. Students are encouraged to obtain course advising early from the Departmental Adviser for Minors. By the end of Spring Quarter of the student's third year, a student who wishes to complete the Statistics minor must complete the Consent to Complete a Minor Form (<https://college.uchicago.edu/advising/forms-and-petitions>) to obtain formal approval of their degree program from the Departmental Adviser for Minors.

The core of the Statistics minor consists of STAT 22400 Applied Regression Analysis and either STAT 22200 Linear Models And Experimental Design or STAT 22600 Analysis of Categorical Data (or both). These three courses may be



taken in any order after meeting the prerequisite of at least two quarters calculus and introductory statistics: STAT 22000 Statistical Methods and Applications, STAT 23400 Statistical Models and Methods, STAT 24500 Statistical Theory and Methods II, STAT 24510 Statistical Theory and Methods IIa, or AP credit for STAT 22000 Statistical Methods and Applications.

An approved substitute for STAT 22600 Analysis of Categorical Data is PBHS 32700 Biostatistical Methods (also designated as STAT 22700 Biostatistical Methods), which requires STAT 22400 Applied Regression Analysis as prerequisite and is offered by the Department of Public Health Sciences. Students may count either STAT 22600 Analysis of Categorical Data or STAT 22700 Biostatistical Methods, but not both, toward the Statistics minor. STAT 22700 Biostatistical Methods does not count against the limit of no more than two electives from outside the Department of Statistics.

To complete the five-course minor, students must choose from among the approved electives listed below. No more than two electives may be satisfied by courses offered by departments other than the Department of Statistics. Students may petition the Department Adviser for Minors to include more than two electives from outside the Department of Statistics. Regardless, at most one elective can be satisfied by a course offered by the Booth School of Business. Further, due to the course grading policies of the Booth School of Business, their courses cannot be counted toward the Statistics minor if taken during the quarter in which the student will graduate from the College.

Either STAT 22000 Statistical Methods and Applications or STAT 23400 Statistical Models and Methods, but not both, may be used as an elective in the Statistics minor if taken prior to any other courses for which at least STAT 22000 Statistical Methods and Applications or STAT 23400 Statistical Models and Methods is prerequisite and before either of STAT 24500 Statistical Theory and Methods II or STAT 24510 Statistical Theory and Methods IIa.

If either STAT 22000 Statistical Methods and Applications or STAT 23400 Statistical Models and Methods are used to fulfill a requirement for any major(s), other minors, or general education requirements, then neither course may be used to fulfill a requirement in the Statistics minor. Students may not use AP credit for STAT 22000 Statistical Methods and Applications to fulfill a requirement for the Statistics minor.

The list of courses approved for the minor may change from time to time as courses change and new courses are added. Please consult the Departmental Adviser for Minors for approval of your minor program plan. Students may petition the Departmental Adviser for Minors for approval of another course. Such courses must have a minimum statistics prerequisite of introductory statistics (STAT 22000 Statistical Methods and Applications, STAT 23400 Statistical Models and Methods, STAT 24500 Statistical Theory and Methods II, or STAT 24510 Statistical Theory and Methods IIa), incorporate a considerable amount of data analysis, and cannot substantially overlap with the topics covered in departmental courses or other courses in the student's minor program.

No courses in the Statistics minor can be double counted with the student's major(s), other minors, or general education requirements. An approved elective must replace any course required for the Statistics minor that is used to meet the requirements for any major(s), other minors, or general education requirements.

The following courses offered by the Department of Statistics cannot be included in a Statistics minor: STAT 11800/STAT 11800 Introduction to Data Science - I, STAT 11900 Introduction to Data Science II, STAT 20000 Elementary Statistics, STAT 24300 Numerical Linear Algebra, STAT 24400 Statistical Theory and Methods I, STAT 24410 Statistical Theory and Methods Ia, STAT 25100 Introduction to Mathematical Probability, STAT 25150 Introduction to Mathematical Probability-A, STAT 25300 Introduction to Probability Models, STAT 27400 Nonparametric Inference, STAT 27850 Multiple Testing, Modern Inference, and Replicability, STAT 28000 Optimization, STAT 28200 Dynamical Systems with Applications, STAT 29700 Undergraduate Research, or any graduate courses in probability. In addition, CMSC 25400 Machine Learning (also designated as STAT 27725 Machine Learning) and PBHS 32901 Introduction to Clinical Trials (also designated as STAT 35201 Introduction to Clinical Trials).

Students who are minoring in Statistics must receive a quality grade of at least C+ in all of the courses counted toward the minor. A grade of P is not acceptable for any of these courses. More than half of the courses counted toward the Statistics minor must be met by registering for courses bearing University of Chicago course numbers.

#### Summary of Requirements for the Minor in Statistics

The following course: *		100
STAT 22400	Applied Regression Analysis *	
One of the following: **		100
STAT 22200	Linear Models And Experimental Design *	
STAT 22600	Analysis of Categorical Data * ****	
Three approved electives ****		300
Total Units		500



- \* STAT 22200 Linear Models And Experimental Design, STAT 22400 Applied Regression Analysis, and STAT 22600 Analysis of Categorical Data may be taken in any order after meeting the prerequisite of at least two quarters calculus and introductory statistics: STAT 22000 Statistical Methods and Applications, STAT 23400 Statistical Models and Methods, STAT 24500 Statistical Theory and Methods II, STAT 24510 Statistical Theory and Methods Iia, or AP credit for STAT 22000 Statistical Methods and Applications.
- \*\* If STAT 22200 Linear Models And Experimental Design is used to fulfill a requirement of the Statistics minor, then STAT 22600 Analysis of Categorical Data may be used as an elective in the minor. Similarly, If STAT 22600 Analysis of Categorical Data is used to fulfill a requirement of the Statistics minor, then STAT 22200 Linear Models And Experimental Design may be used as an elective in the minor.
- \*\*\* An approved substitute for STAT 22600 Analysis of Categorical Data is STAT 22700 Biostatistical Methods, which requires STAT 22400 Applied Regression Analysis as prerequisite and is offered by the Department of Public Health Sciences. Students may count either STAT 22600 Analysis of Categorical Data or STAT 22700 Biostatistical Methods, but not both, toward the Statistics minor. STAT 22700 Biostatistical Methods does not count against the limit of no more than two electives from outside the Department of Statistics.
- \*\*\*\* Not every combination of elective courses is allowed. No more than two electives may be satisfied by courses offered by departments other than the Department of Statistics. Students may petition the Departmental Adviser for Minors to include more than two electives from outside the Department of Statistics. Regardless, at most one elective can be satisfied by a course offered by the Booth School of Business. Further, due to the course grading policies of the Booth School of Business, their courses cannot be counted toward the Statistics minor if taken during the quarter in which the student will graduate from the College.

**Departmental Electives Approved for the Minor in Statistics**

STAT 22000	Statistical Methods and Applications <sup>1,2</sup>	100
STAT 22200	Linear Models And Experimental Design <sup>3</sup>	100
STAT 22600	Analysis of Categorical Data <sup>3, 4</sup>	100
STAT 23400	Statistical Models and Methods <sup>1</sup>	100
STAT 24500	Statistical Theory and Methods II <sup>5</sup>	100
STAT 24510	Statistical Theory and Methods Iia <sup>5</sup>	100
STAT 26100	Time Dependent Data	100
STAT 26700	History of Statistics	100

- <sup>1</sup> Either STAT 22000 Statistical Methods and Applications or STAT 23400 Statistical Models and Methods, but not both, may be used as an elective if taken prior to any other courses for which at least STAT 22000 Statistical Methods and Applications or STAT 23400 Statistical Models and Methods is prerequisite. If either STAT 22000 Statistical Methods and Applications or STAT 23400 Statistical Models and Methods is used to fulfill a requirement for any major(s), other minors, or general education requirements, then neither course may be used to fulfill a requirement in the Statistics minor.
- <sup>2</sup> Students may not use AP credit for STAT 22000 Statistical Methods and Applications to meet a requirement for the Statistics minor.
- <sup>3</sup> If STAT 22200 Linear Models And Experimental Design is used to fulfill a requirement of the Statistics minor, then STAT 22600 Analysis of Categorical Data may be used as an elective in the minor. Similarly, if STAT 22600 Analysis of Categorical Data is used to fulfill a requirement of the Statistics minor, then STAT 22200 Linear Models And Experimental Design may be used as an elective in the minor.
- <sup>4</sup> An approved substitute for STAT 22600 Analysis of Categorical Data is PBHS 32700 Biostatistical Methods (also designated as STAT 22700 Biostatistical Methods), which requires STAT 22400 Applied Regression Analysis as prerequisite and is offered by the Department of Public Health Sciences. Students may count either STAT 22600 Analysis of Categorical Data or STAT 22700 Biostatistical Methods, but not both, toward the Statistics minor. STAT 22700 Biostatistical Methods does not count against the limit of no more than two electives from outside the Department of Statistics.
- <sup>5</sup> If either STAT 24500 Statistical Theory and Methods II or STAT 24510 Statistical Theory and Methods Iia is used as an elective in the Statistics minor, then the prerequisite STAT 24400 Statistical Theory and Methods I or STAT 24410 Statistical Theory and Methods Ia may not be counted toward the minor, but may be counted toward any major(s) or other minors.

**Non-Departmental Electives Approved for the Minor in Statistics**

Because of the interdisciplinary nature of the field of statistics, other departments and committees offer courses approved for use as electives for the Statistics minor. No more than two electives may be satisfied by courses offered by departments other than the Department of Statistics. Students may petition the Departmental Adviser for Minors to include more than two electives from outside the Department of Statistics. Regardless, at most one elective can be satisfied by a course offered by the Booth School of Business. Further, due to the course grading policies of the Booth School of Business, their courses cannot be counted toward the Statistics minor if taken during the quarter in which the student will graduate from the College.

Offering departments include Public Health Sciences, Computer Science, Comparative Human Development, Human Genetics, Public Policy, Sociology, and the Booth School of Business.

BIOS 21216	Intro Statistical Genetics	100
BUSN 41201	Big Data <sup>1</sup>	100
BUSN 41204	Machine Learning <sup>1</sup>	100
CHDV 30102	Introduction to Causal Inference	100
CHDV 32411	Mediation, Moderation, and Spillover Effects	100
PBHS 30910	Epidemiology and Population Health	100
PBHS 31001	Epidemiologic Methods	100
PBHS 32700	Biostatistical Methods <sup>2</sup>	100
PBHS 33300	Applied Longitudinal Data Analysis	100
PBHS 33500	Statistical Applications	100
PBHS 35100	Health Services Research Methods	100
PBPL 28820	Machine Learning and Policy	100
SOCI 20112	Applications of Hierarchical Linear Models	100
SOCI 20253	Introduction to Spatial Data Science	100

<sup>1</sup> At most one elective can be satisfied by a course offered by the Booth School of Business. Due to the course grading policies of the Booth School of Business, their courses cannot be counted toward the Statistics minor if taken during the quarter in which the student will graduate from the College.

<sup>2</sup> An approved substitute for STAT 22600 Analysis of Categorical Data is PBHS 32700 Biostatistical Methods (also designated as STAT 22700 Biostatistical Methods), which requires STAT 22400 Applied Regression Analysis as prerequisite and is offered by the Department of Public Health Sciences. Students may count either STAT 22600 Analysis of Categorical Data or STAT 22700 Biostatistical Methods, but not both, toward the Statistics minor. STAT 22700 Biostatistical Methods does not count against the limit of no more than two electives from outside the Department of Statistics.

Some of the approved electives offered by other departments also bear a Statistics course number and some do not. Students should enroll in the relevant Department of Statistics course number when available. Examples include STAT 22700 Biostatistical Methods, STAT 22810 Epidemiology and Population Health, STAT 31900 Introduction to Causal Inference, STAT 33211 Mediation, Moderation, and Spillover Effects, STAT 35700 Epidemiologic Methods, STAT 35800 Statistical Applications, and STAT 36900 Applied Longitudinal Data Analysis.

Undergraduate registration in 30000-level courses is by instructor consent only. Undergraduates cannot pre-register for 30000-level courses. Instead, students should contact the instructor well in advance.

## Statistics Courses

### **STAT 11800. Introduction to Data Science - I. 100 Units.**

Data science provides tools for gaining insight into specific problems using data, through computation, statistics and visualization. This course introduces students to all aspects of a data analysis process, from posing questions, designing data collection strategies, management+storing and processing of data, exploratory tools and visualization, statistical inference, prediction, interpretation and communication of results. Simple techniques for data analysis are used to illustrate both effective and fallacious uses of data science tools. Although this course is designed to be at the level of mathematical sciences courses in the Core, with little background required, we expect the students to develop computational skills that will allow them to analyze data. Computation will be done using Python and Jupyter Notebook.

Instructor(s): Michael J. Franklin, Dan Nicolae Terms Offered: Autumn. Note: This course will be crosslisted with STAT 11800.

Prerequisite(s): None

Equivalent Course(s): CMSC 11800

### **STAT 11900. Introduction to Data Science II. 100 Units.**

This course is the second quarter of a two-quarter systematic introduction to the foundations of data science, as well as to practical considerations in data analysis. A broad background on probability and statistical methodology as well as a basic proficiency in RStudio will be provided. More advanced topics on data privacy and ethics, reproducibility in science, data encryption, and basic machine learning will be introduced. We will explore these concepts with real-world problems from different domains.

Terms Offered: Winter

Equivalent Course(s): CMSC 11900

**STAT 20000. Elementary Statistics. 100 Units.**

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.

Instructor(s): Staff Terms Offered: Autumn Spring Winter

Note(s): For students with little or no math background. Not recommended for students planning to take STAT 22000 or STAT 23400 or more advanced courses in Statistics. Students with credit for STAT 22000, STAT 23400 or more advanced courses in Statistics not admitted. This course may not be used in the Statistics major or minor. This course meets one of the general education requirements in the mathematical sciences. Only one of STAT 20000 and STAT 22000, but not both, can count toward the general education requirement in the mathematical sciences.

**STAT 22000. Statistical Methods and Applications. 100 Units.**

This course introduces statistical techniques and methods of data analysis, including the use of statistical software. 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, basic probability, random variables and expected values, confidence intervals and significance tests for one- and two-sample problems for means and proportions, chi-square tests, linear regression, and, if time permits, analysis of variance.

Terms Offered: Autumn Spring Winter

Prerequisite(s): MATH 13100 or placement into MATH 15100

Note(s): Students may count either STAT 22000 or 23400, but not both, toward the forty-two credits required for graduation. Students with credit for STAT 23400 not admitted. This course meets one of the general education requirements in the mathematical sciences. Only one of STAT 20000 and STAT 22000, but not both, can count toward the general education requirement in the mathematical sciences.

**STAT 22200. Linear Models And Experimental Design. 100 Units.**

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.

Instructor(s): Staff Terms Offered: Spring

Prerequisite(s): STAT 22000 or 23400 or 24500 or 24510 and two quarters of calculus.

**STAT 22400. Applied Regression Analysis. 100 Units.**

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. Prerequisite: PBHS 32100. Equivalent Course(s): PBHS 32400

Terms Offered: Autumn Spring

Prerequisite(s): STAT 22000 or 23400 or 24500 or PBHS 32100

Equivalent Course(s): PBHS 32400

**STAT 22600. Analysis of Categorical Data. 100 Units.**

This course covers statistical methods for the analysis of qualitative and counted data. Topics include description and inference for binomial and multinomial data using proportions and odds ratios; multi-way contingency tables; generalized linear models for discrete data; logistic regression for binary responses; multi-category logit models for nominal and ordinal responses; loglinear models for counted data; and inference for matched-pairs and correlated data. Applications and interpretations of statistical models are emphasized.

Terms Offered: Winter

Prerequisite(s): STAT 22000 or 23400 or 24500

Equivalent Course(s): PBHS 32600

**STAT 22700. Biostatistical Methods. 100 Units.**

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.

Instructor(s): F. Yang Terms Offered: Winter

Prerequisite(s): PBHS 32400, STAT 22400 or STAT 24500 or equivalent or consent of instructor.

Equivalent Course(s): PBHS 32700

**STAT 22810. Epidemiology and Population Health. 100 Units.**

This course does not meet requirements for the biological sciences major. 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.

Instructor(s): B. Lahey Terms Offered: Autumn

Prerequisite(s): Introductory statistics recommended or Consent of Instructor

Equivalent Course(s): ENST 27400, PBHS 30910, PPHA 36410

**STAT 23400. Statistical Models and Methods. 100 Units.**

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 computer simulation. 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 proportions for one- and two-sample problems, two-way tables, correlation, and simple linear regression. Graphical 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 single-variable differential and integral calculus and sequences and series.

Terms Offered: Autumn Spring Winter

Prerequisite(s): MATH 13300, 15300, or 16200

Note(s): Students may count either STAT 22000 or 23400, but not both, toward the forty-two credits required for graduation.

**STAT 24300. Numerical Linear Algebra. 100 Units.**

This course is devoted to the basic theory of linear algebra and its significant applications in scientific computing. The objective is to provide a working knowledge and hands-on experience of the subject suitable for graduate level work in statistics, econometrics, quantum mechanics, and numerical methods in scientific computing. Topics include Gaussian elimination, vector spaces, linear transformations and associated fundamental subspaces, orthogonality and projections, eigenvectors and eigenvalues, diagonalization of real symmetric and complex Hermitian matrices, the spectral theorem, and matrix decompositions (QR, Cholesky and Singular Value Decompositions). Systematic methods applicable in high dimensions and techniques commonly used in scientific computing are emphasized. Students enrolled in the graduate level STAT 30750 will have additional work in assignments, exams, and projects including applications of matrix algebra in statistics and numerical computations implemented in Matlab or R. Some programming exercises will appear as optional work for students enrolled in the undergraduate level STAT 24300.

Terms Offered: Autumn

Prerequisite(s): Multivariate calculus (MATH 19520 or 20000 or 20500 or equivalent). Previous exposure to linear algebra is helpful.

Equivalent Course(s): STAT 30750

**STAT 24400-24500. Statistical Theory and Methods I-II.**

This sequence 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.

**STAT 24400. Statistical Theory and Methods I. 100 Units.**

This course is the first quarter of a two-quarter 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. This course 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. Students who have already taken STAT 25100 have the option to take STAT 24410 (if offered) instead of STAT 24400.

Instructor(s): Staff Terms Offered: Autumn Winter

Prerequisite(s): MATH 19520 or 20000 with a grade of B or better, or MATH 16300 or 20250 or 20300 or 20700 or STAT 24300 or PHYS 22100. Concurrent or prior linear algebra (MATH 19620 or 20250 or STAT 24300 or equivalent) is recommended for students continuing to STAT 24500.

Note(s): Some previous experience with statistics and/or probability helpful but not required. Students may count either STAT 24400 or STAT 24410, but not both, toward the forty-two credits required for graduation.

**STAT 24500. Statistical Theory and Methods II. 100 Units.**

This course is the second quarter of a two-quarter 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. This course continues from either STAT 24400 or STAT 24410 and 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. Statistical software is used.

Instructor(s): Staff Terms Offered: Spring Winter

Prerequisite(s): Linear algebra (MATH 19620 or 20250 or STAT 24300 or equivalent) and STAT 24400 or STAT 24410.

Note(s): Students may count either STAT 24500 or STAT 24510, but not both, toward the forty-two credits required for graduation.

**STAT 24410-24510. Statistical Theory and Methods Ia-IIa.**

This sequence provides a sophisticated introduction to statistical principles and methods and their application to the analysis of data. In addition to serving undergraduates with a strong interest in Statistics, it is the recommended sequence in theoretical statistics for MS students in Statistics.

**STAT 24410. Statistical Theory and Methods Ia. 100 Units.**

This course is the first quarter of a two-quarter sequence providing a principled development of statistical methods, including practical considerations in applying these methods to the analysis of data. The course begins with a brief review of probability and some elementary stochastic processes, such as Poisson processes, that are relevant to statistical applications. The bulk of the quarter covers principles of statistical inference from both frequentist and Bayesian points of view. Specific topics include maximum likelihood estimation, posterior distributions, confidence and credible intervals, principles of hypothesis testing, likelihood ratio tests, multinomial distributions, and chi-square tests. Additional topics may include diagnostic plots, bootstrapping, a critical comparison of Bayesian and frequentist inference, and the role of conditioning in statistical inference. Examples are drawn from the social, physical, and biological sciences. The statistical software package R will be used to analyze datasets from these fields and instruction in the use of R is part of the course.

Instructor(s): Staff Terms Offered: Autumn

Prerequisite(s): STAT 25100 or STAT 25150 or MATH 23500. Concurrent or prior linear algebra (MATH 19620 or 20250 or STAT 24300 or equivalent) is recommended for students continuing to STAT 24510.

Note(s): Some previous experience with statistics helpful but not required. Students may count either STAT 24400 or STAT 24410, but not both, toward the forty-two credits required for graduation.

Equivalent Course(s): STAT 30030

**STAT 24510. Statistical Theory and Methods IIa. 100 Units.**

This course is a continuation of STAT 24410. The focus is on theory and practice of linear models, including the analysis of variance, regression, correlation, and some multivariate analysis. Additional topics may include bootstrapping for regression models, nonparametric regression, and regression models with correlated errors.

Terms Offered: May be offered in Winter.

Prerequisite(s): STAT 24410. Linear algebra (MATH 19620 or 20250 or STAT 24300 or equivalent).

Note(s): Students may count either STAT 24500 or STAT 24510, but not both, toward the forty-two credits required for graduation.

Equivalent Course(s): STAT 30040

**STAT 24500. Statistical Theory and Methods II. 100 Units.**

This course is the second quarter of a two-quarter 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. This course continues from either STAT 24400 or STAT 24410 and 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. Statistical software is used.

Instructor(s): Staff Terms Offered: Spring Winter

Prerequisite(s): Linear algebra (MATH 19620 or 20250 or STAT 24300 or equivalent) and STAT 24400 or STAT 24410.

Note(s): Students may count either STAT 24500 or STAT 24510, but not both, toward the forty-two credits required for graduation.

**STAT 24510. Statistical Theory and Methods IIa. 100 Units.**

This course is a continuation of STAT 24410. The focus is on theory and practice of linear models, including the analysis of variance, regression, correlation, and some multivariate analysis. Additional topics may include bootstrapping for regression models, nonparametric regression, and regression models with correlated errors.

Terms Offered: May be offered in Winter.

Prerequisite(s): STAT 24410. Linear algebra (MATH 19620 or 20250 or STAT 24300 or equivalent).

Note(s): Students may count either STAT 24500 or STAT 24510, but not both, toward the forty-two credits required for graduation.

Equivalent Course(s): STAT 30040

**STAT 24620. Multivariate Statistical Analysis: Applications and Techniques. 100 Units.**

This course focuses on applications and techniques for analysis of multivariate and high dimensional data. Beginning subjects cover principal component analysis, factor model, canonical correlation, multi-dimensional scaling, discriminant analysis, clustering, and common techniques of dimension reduction. Further topics on statistical learning for high dimensional data and complex structures include penalized regression models (LASSO, ridge, elastic net), sparse PCA, independent component analysis, Gaussian mixture model, and Expectation-Maximization methods. Theoretical derivations will be presented with emphasis on motivations, applications, and hands-on data analysis.

Terms Offered: Spring

Prerequisite(s): STAT 24400-24500 or STAT 24410-24510 or consent of instructor

Equivalent Course(s): STAT 32950

**STAT 25100. Introduction to Mathematical Probability. 100 Units.**

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.

Instructor(s): Staff Terms Offered: Autumn Spring

Prerequisite(s): MATH 19520, 20000, 20500 or 20900. MATH 20000 or higher recommended.

Note(s): Students may count either STAT 25100 or STAT 25150, but not both, toward the forty-two credits required for graduation.

**STAT 25150. Introduction to Mathematical Probability-A. 100 Units.**

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.

Instructor(s): Robert Fefferman Terms Offered: Not offered 2018-2019

Prerequisite(s): MATH 20500 or consent of instructor

Note(s): Students may count either STAT 25100 or STAT 25150, but not both, toward the forty-two credits required for graduation.

**STAT 25300. Introduction to Probability Models. 100 Units.**

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).

Instructor(s): Staff Terms Offered: May be offered in Winter

Prerequisite(s): STAT 24400 or STAT 25100 or STAT 25150

Equivalent Course(s): STAT 31700

**STAT 26100. Time Dependent Data. 100 Units.**

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.

Instructor(s): Staff Terms Offered: Winter or Spring

Prerequisite(s): STAT 24500 or STAT 24510 is required; alternatively STAT 22400 and exposure to multivariate calculus. Some previous exposure to Fourier series is helpful but not required.

Equivalent Course(s): STAT 33600

**STAT 26300. Introduction to Statistical Genetics. 100 Units.**

As a result of technological advances over the past few decades, there is a tremendous wealth of genetic data currently being collected. These data have the potential to shed light on the genetic factors influencing traits and diseases, as well as on questions of ancestry and population history. The aim of this course is to develop a thorough understanding of probabilistic models and statistical theory and methods underlying analysis of genetic data, focusing on problems in complex trait mapping, with some coverage of population genetics. Although the case studies are all in the area of statistical genetics, the statistical inference topics, which will include likelihood-based inference, linear mixed models, and restricted maximum likelihood, among others, are widely applicable to other areas. No biological background is needed, but a strong foundation in statistical theory and methods is assumed.

Terms Offered: Spring

Prerequisite(s): STAT 24500 or STAT 24510

**STAT 26700. History of Statistics. 100 Units.**

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.

Instructor(s): S. Stigler Terms Offered: Spring

Prerequisite(s): Prior statistics course

Equivalent Course(s): STAT 36700, HIPS 25600, CHSS 32900

**STAT 27400. Nonparametric Inference. 100 Units.**

Nonparametric inference is about developing statistical methods and models that make weak assumptions. A typical nonparametric approach estimates a nonlinear function from an infinite dimensional space rather than a linear model from a finite dimensional space. This course gives an introduction to nonparametric inference, with a focus on density estimation, regression, confidence sets, orthogonal functions, random processes, and kernels. The course treats nonparametric methodology and its use, together with theory that explains the statistical properties of the methods.

Instructor(s): Staff Terms Offered: Autumn

Prerequisite(s): STAT 24400 or STAT 24410 is required; alternatively STAT 22400 and exposure to multivariate calculus and linear algebra.

Equivalent Course(s): STAT 37400

**STAT 27725. Machine Learning. 100 Units.**

This course offers a practical, problem-centered introduction to machine learning. Topics covered include the Perceptron and other online algorithms; boosting; graphical models and message passing; dimensionality reduction and manifold learning; SVMs and other kernel methods; artificial neural networks; and a short introduction to statistical learning theory. Weekly programming assignments give students the opportunity to try out each learning algorithm on real world datasets.

Instructor(s): R. Kondor Terms Offered: Winter

Prerequisite(s): CMSC 15400 or CMSC 12300. STAT 22000 or STAT 23400 strongly recommended.

Equivalent Course(s): CMSC 25400

**STAT 27850. Multiple Testing, Modern Inference, and Replicability. 100 Units.**

This course examines the problems of multiple testing and statistical inference from a modern point of view. High-dimensional data is now common in many applications across the biological, physical, and social sciences. With this increased capacity to generate and analyze data, classical statistical methods may no longer ensure the reliability or replicability of scientific discoveries. We will examine a range of modern methods that provide statistical inference tools in the context of modern large-scale data analysis. The course will have weekly assignments as well as a final project, both of which will include both theoretical and computational components.

Terms Offered: Winter

Equivalent Course(s): STAT 30850

**STAT 28000. Optimization. 100 Units.**

This is an introductory course on optimization that will cover the rudiments of unconstrained and constrained optimization of a real-valued multivariate function. The focus is on the settings where this function is, respectively, linear, quadratic, convex, or differentiable. Time permitting, topics such as nonsmooth, integer, vector, and dynamic optimization may be briefly addressed. Materials will include basic duality theory, optimality conditions, and intractability results, as well as algorithms and applications.

Instructor(s): Staff Terms Offered: Spring

Prerequisite(s): MATH 20500 or 20800; STAT 24300 or MATH 20250

Equivalent Course(s): CAAM 28000

**STAT 28200. Dynamical Systems with Applications. 100 Units.**

This course is concerned with the analysis of nonlinear dynamical systems arising in the context of mathematical modeling. The focus is on qualitative analysis of solutions as trajectories in phase space, including the role of invariant manifolds as organizers of behavior. Local and global bifurcations, which occur as system parameters change, will be highlighted, along with other dimension reduction methods that arise when there is a natural time-scale separation. Concepts of bi-stability, spontaneous oscillations, and chaotic dynamics will be explored through investigation of conceptual mathematical models arising in the physical and biological sciences.

Instructor(s): Mary Silber Terms Offered: Winter

Prerequisite(s): Multivariable calculus (MATH 19520, 20000 or 20400, or PHYS 22100, or equivalent). Linear algebra, including eigenvalues and eigenvectors (MATH 19620 or STAT 24300, or equivalent). Previous knowledge of elementary differential equations is helpful but not required.

Equivalent Course(s): CAAM 28200

**STAT 29700. Undergraduate Research. 100 Units.**

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.

Instructor(s): Staff Terms Offered: Autumn Spring Winter

Prerequisite(s): Consent of faculty adviser and Departmental Adviser for Majors

Note(s): 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.



**STAT 29900. Bachelor's Paper. 100 Units.**

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.

Terms Offered: Autumn Spring Winter

Prerequisite(s): Consent of faculty adviser and Departmental Adviser for Majors

Note(s): 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 major in Statistics.





### Font Notice

This document should contain certain fonts with restrictive licenses. For this draft, substitutions were made using less legally restrictive fonts. Specifically:

Times was used instead of Trajan.

Times was used instead of Palatino.

The editor may contact Leepfrog for a draft with the correct fonts in place.