Department of Statistics

Daniel L. Gillen, Department Chair
2038 Donald Bren Hall
949-824-9862
Fax: 949-824-9863
http://www.stat.uci.edu/

Overview
Statistics is the science concerned with developing and studying methods for collecting, analyzing, interpreting, and presenting empirical data. Statistical principles and methods are important for addressing questions in public policy, medicine, industry, and virtually every branch of science. Interest in statistical methods has increased dramatically with the abundance of large databases in fields like computer science (Internet and Web traffic), business and marketing (transaction records), and biology (the human genome and related data). It is the substantive questions in such areas of application that drive the development of new statistical methods and motivate the mathematical study of the properties of these methods.

Undergraduate Major in Data Science
The Data Science Major prepares students for a career in data analysis, combining foundational statistical concepts with computational principles from computer science. In the first two years of the program students will take core courses in both the Statistics and Computer Science Departments, providing a strong foundation in the principles of each field. In the 3rd and 4th years of the program, students will take more specialized courses, on topics such as design of algorithms, machine learning, information visualization, and Bayesian statistics. A major component of this degree is the final year capstone project course, a 2-quarter course that teaches students how to apply statistical and computational principles to solve large-scale real-world data analysis problems.

Admissions
Freshman Applicants: See the Undergraduate Admissions section.

Transfer Applicants: Junior-level applicants who satisfactorily complete course requirements will be given preference for admission. Applicants must satisfy the following requirements:

1. Completion of one year of college level mathematics (calculus or discrete math) and one semester of college level statistics.
2. Completion of one year of transferable Computer Science courses*; at least one of these should involve concepts such as those found in the Python and C++ programming languages, or another high-level programming language.

*NOTE: Additional Computer Science and Statistics courses beyond those above are strongly recommended, particularly those that align with the major(s) of interest. Python, C++ and R are used extensively in the curriculum; therefore, transfer students should plan to learn these by studying on their own or by completing Python, C++, and R-related programming courses prior to their first quarter at UCI. Additional courses beyond those required for admission must be taken to fulfill the lower-division degree requirements, as many are prerequisites for upper-division courses. For some transfer students, this may mean that it will take longer than two years to complete their degree.

Major and Minor Restrictions
Bren School of ICS majors (including shared majors, BIM and CSE) pursuing minors within the Bren School of ICS may not count more than five courses toward both the major and minor. Some ICS majors and minors outside of the School are not permitted due to significant overlap. Visit the ICS Student Affairs Office website for Majors and Minors restrictions. (http://www.ics.uci.edu/ugrad/degrees/MajorMinor_Restrictions_Chart.pdf) All students should check the Double Major Restrictions Chart (http://www.ics.uci.edu/ugrad/degrees/Dbl_Major_Restr_Chart.pdf) and view our information page (http://www.ics.uci.edu/ugrad/degrees/Double_Majors.php) on double majoring to see what degree programs are eligible for double majoring.

Requirements for the B.S. in Data Science
All students must meet the University Requirements.

Data Science Major Requirements
Lower-division:

A. Select one of the following series:

<table>
<thead>
<tr>
<th>Course Code</th>
<th>Course Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>I&amp;C SCI 31- 32- 33</td>
<td>Introduction to Programming and Programming with Software Libraries and Intermediate Programming</td>
</tr>
<tr>
<td>or</td>
<td></td>
</tr>
<tr>
<td>I&amp;C SCI 32A- 33</td>
<td>Python Programming and Libraries (Accelerated) and Intermediate Programming</td>
</tr>
</tbody>
</table>

B. Complete:
<table>
<thead>
<tr>
<th>Course Code</th>
<th>Course Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>I&amp;C SCI 45C</td>
<td>Programming in C/C++ as a Second Language</td>
</tr>
<tr>
<td>I&amp;C SCI 46</td>
<td>Data Structure Implementation and Analysis</td>
</tr>
<tr>
<td>I&amp;C SCI 51</td>
<td>Introductory Computer Organization</td>
</tr>
<tr>
<td>IN4MATX 43</td>
<td>Introduction to Software Engineering</td>
</tr>
</tbody>
</table>

**C. Complete:**

<table>
<thead>
<tr>
<th>Course Code</th>
<th>Course Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>MATH 2A</td>
<td>Single-Variable Calculus</td>
</tr>
<tr>
<td>MATH 2B</td>
<td>Single-Variable Calculus</td>
</tr>
<tr>
<td>MATH 2D</td>
<td>Multivariable Calculus</td>
</tr>
<tr>
<td>MATH 3A</td>
<td>Introduction to Linear Algebra</td>
</tr>
<tr>
<td>or I&amp;C SCI 6N</td>
<td>Computational Linear Algebra</td>
</tr>
<tr>
<td>I&amp;C SCI 6B</td>
<td>Boolean Logic and Discrete Structures</td>
</tr>
<tr>
<td>I&amp;C SCI 6D</td>
<td>Discrete Mathematics for Computer Science</td>
</tr>
<tr>
<td>STATS 5</td>
<td>Seminar in Data Science</td>
</tr>
<tr>
<td>STATS 7</td>
<td>Basic Statistics</td>
</tr>
<tr>
<td>STATS 68</td>
<td>Statistical Computing and Exploratory Data Analysis</td>
</tr>
</tbody>
</table>

**Upper-division:**

**A. Data Science core requirements:**

<table>
<thead>
<tr>
<th>Course Code</th>
<th>Course Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>STATS 110</td>
<td>Statistical Methods for Data Analysis I</td>
</tr>
<tr>
<td>STATS 111</td>
<td>Statistical Methods for Data Analysis II</td>
</tr>
<tr>
<td>STATS 112</td>
<td>Statistical Methods for Data Analysis III</td>
</tr>
<tr>
<td>STATS 115</td>
<td>Introduction to Bayesian Data Analysis</td>
</tr>
<tr>
<td>STATS 120A</td>
<td>Introduction to Probability and Statistics I</td>
</tr>
<tr>
<td>STATS 120B</td>
<td>Introduction to Probability and Statistics II</td>
</tr>
<tr>
<td>STATS 120C</td>
<td>Introduction to Probability and Statistics III</td>
</tr>
<tr>
<td>I&amp;C SCI 139W</td>
<td>Critical Writing on Information Technology</td>
</tr>
<tr>
<td>COMPSCI 122A</td>
<td>Introduction to Data Management</td>
</tr>
<tr>
<td>COMPSCI 161</td>
<td>Design and Analysis of Algorithms</td>
</tr>
<tr>
<td>COMPSCI 178</td>
<td>Machine Learning and Data-Mining</td>
</tr>
<tr>
<td>IN4MATX 143</td>
<td>Information Visualization</td>
</tr>
</tbody>
</table>

**B. Three elective courses from the list below:**

<table>
<thead>
<tr>
<th>Course Code</th>
<th>Course Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>MATH 130B</td>
<td>Probability and Stochastic Processes</td>
</tr>
<tr>
<td>MATH 130C</td>
<td>Probability and Stochastic Processes</td>
</tr>
<tr>
<td>STATS 140</td>
<td>Multivariate Statistical Methods</td>
</tr>
<tr>
<td>I&amp;C SCI 53</td>
<td>Principles in System Design</td>
</tr>
<tr>
<td>COMPSCI 111</td>
<td>Digital Image Processing</td>
</tr>
<tr>
<td>COMPSCI 115</td>
<td>Computer Simulation</td>
</tr>
<tr>
<td>COMPSCI 121</td>
<td>Information Retrieval</td>
</tr>
<tr>
<td>COMPSCI 122B</td>
<td>Project in Databases and Web Applications</td>
</tr>
<tr>
<td>COMPSCI 122C</td>
<td>Principles of Data Management</td>
</tr>
<tr>
<td>COMPSCI 125</td>
<td>Next Generation Search Systems</td>
</tr>
<tr>
<td>COMPSCI 131</td>
<td>Parallel and Distributed Computing</td>
</tr>
<tr>
<td>COMPSCI 134</td>
<td>Computer and Network Security</td>
</tr>
<tr>
<td>COMPSCI 163</td>
<td>Graph Algorithms</td>
</tr>
<tr>
<td>COMPSCI 165</td>
<td>Project In Algorithms And Data Structures</td>
</tr>
<tr>
<td>COMPSCI 169</td>
<td>Introduction to Optimization</td>
</tr>
<tr>
<td>COMPSCI 171</td>
<td>Introduction to Artificial Intelligence</td>
</tr>
<tr>
<td>COMPSCI 172B</td>
<td>Neural Networks and Deep Learning</td>
</tr>
<tr>
<td>IN4MATX 131</td>
<td>Human Computer Interaction</td>
</tr>
<tr>
<td>IN4MATX 141</td>
<td>Information Retrieval</td>
</tr>
<tr>
<td>IN4MATX 161</td>
<td>Social Analysis of Computing</td>
</tr>
</tbody>
</table>

**C. Data Science capstone team-based project courses:** STATS 170A and STATS 170B
Sample Program of Study — Data Science

Freshman

<table>
<thead>
<tr>
<th>Fall</th>
<th>Winter</th>
<th>Spring</th>
</tr>
</thead>
<tbody>
<tr>
<td>I&amp;C SCI 31</td>
<td>I&amp;C SCI 32</td>
<td>I&amp;C SCI 33</td>
</tr>
<tr>
<td>MATH 2A</td>
<td>MATH 2B</td>
<td>MATH 2D</td>
</tr>
<tr>
<td>WRITING 39A</td>
<td>STATS 5</td>
<td>STATS 7</td>
</tr>
<tr>
<td></td>
<td>WRITING 39B</td>
<td>WRITING 39C</td>
</tr>
</tbody>
</table>

Sophomore

<table>
<thead>
<tr>
<th>Fall</th>
<th>Winter</th>
<th>Spring</th>
</tr>
</thead>
<tbody>
<tr>
<td>I&amp;C SCI 6B</td>
<td>I&amp;C SCI 45C</td>
<td>I&amp;C SCI 46</td>
</tr>
<tr>
<td>STATS 120A</td>
<td>I&amp;C SCI 51</td>
<td>I&amp;C SCI 6D</td>
</tr>
<tr>
<td>MATH 3A</td>
<td>STATS 120B</td>
<td>STATS 68</td>
</tr>
<tr>
<td>General Education III</td>
<td></td>
<td>STATS 120C</td>
</tr>
</tbody>
</table>

Junior

<table>
<thead>
<tr>
<th>Fall</th>
<th>Winter</th>
<th>Spring</th>
</tr>
</thead>
<tbody>
<tr>
<td>STATS 110</td>
<td>STATS 111</td>
<td>STATS 112</td>
</tr>
<tr>
<td>IN4MATX 43</td>
<td>COMPSCI 178</td>
<td>COMPSCI 122A</td>
</tr>
<tr>
<td>COMPSCI 161</td>
<td>I&amp;C SCI 139W</td>
<td>IN4MATX 143</td>
</tr>
<tr>
<td>General Education IV/VIII</td>
<td>General Education III/VII</td>
<td>General Education VI</td>
</tr>
</tbody>
</table>

Senior

<table>
<thead>
<tr>
<th>Fall</th>
<th>Winter</th>
<th>Spring</th>
</tr>
</thead>
<tbody>
<tr>
<td>STATS 115</td>
<td>STATS 170A</td>
<td>STATS 170B</td>
</tr>
<tr>
<td>General Education III</td>
<td>General Education IV</td>
<td>Data Science Major Elective</td>
</tr>
<tr>
<td>General Education IV</td>
<td>Data Science Major Elective</td>
<td>Data Science Major Elective</td>
</tr>
</tbody>
</table>

Additional Information

Career Opportunities

A wide variety of careers and graduate programs are open to graduates of the Data Science major. Demand for graduates with skills in both statistics and computer science currently outpaces supply - thus, students with these skills typically find employment quickly, across a wide variety of sectors, including internet companies, finance, engineering, business, medicine, and more. Data Science graduates are well-qualified for job titles such as “data scientist,” “data analyst,” or “statistician,” both in the public and private sectors. Graduate school in area such as Computer Science or Statistics is also a possible career path.

Undergraduate Program in Statistics

The Department of Statistics offers lower-division undergraduate courses designed to introduce students to the field of statistics (STATS 7, STATS 8, STATS 67) and upper-division undergraduate courses on the theoretical foundations of probability and statistics (STATS 120A-STATS 120B-STATS 120C) and statistical methodology (STATS 110-STATS 111-STATS 112). The Department is in the process of planning an undergraduate degree program in Statistics. In the interim, students interested in focusing on statistics are encouraged to consider a minor in Statistics along with a major in a field of interest.

Minor in Statistics

The minor in Statistics is designed to provide students with exposure to both statistical theory and practice. The minor requires a total of seven courses. These include a mathematics course, five core statistics courses, and an elective that may be taken from among several departments. Some of the courses used to complete the minor may include prerequisites that may or may not be part of a student’s course requirements for their major. Because of this, the minor is somewhat intensive, but it is a useful complement to a variety of undergraduate fields for mathematically inclined students. The minor, supplemented with a few additional courses (mathematics and computing), would provide sufficient background for graduate study in statistics. Students considering a minor in Statistics should meet with the academic counselor of their major as early as possible to plan their course work and incorporate the required courses into their four-year academic plan.

NOTE: Students may not receive both a minor in Statistics and a specialization in Statistics within the Mathematics major.

Requirements for the Minor

Required Courses

<table>
<thead>
<tr>
<th>MATH 3A</th>
<th>Introduction to Linear Algebra</th>
</tr>
</thead>
<tbody>
<tr>
<td>or I&amp;C SCI 6N</td>
<td>Computational Linear Algebra</td>
</tr>
<tr>
<td>STATS 110-111</td>
<td>Statistical Methods for Data Analysis I and Statistical Methods for Data Analysis II</td>
</tr>
</tbody>
</table>
STATS 120A- 120B- 120C

Introduction to Probability and Statistics I and Introduction to Probability and Statistics II and Introduction to Probability and Statistics III

Select one elective from the following: ¹

<table>
<thead>
<tr>
<th>Course</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>I&amp;C SCI 31</td>
<td>Introduction to Programming</td>
</tr>
<tr>
<td>MATH 130B</td>
<td>Probability and Stochastic Processes</td>
</tr>
<tr>
<td>or MATH 130C</td>
<td>Probability and Stochastic Processes</td>
</tr>
<tr>
<td>STATS 7</td>
<td>Basic Statistics (or equivalent course) ²</td>
</tr>
<tr>
<td>STATS 112</td>
<td>Statistical Methods for Data Analysis III</td>
</tr>
<tr>
<td>STATS 140</td>
<td>Introduction to Bayesian Data Analysis</td>
</tr>
</tbody>
</table>

¹ Or can substitute another course with approval of the Director of Undergraduate Studies.
² Only if taken prior to STATS 110

NOTE: A maximum of two courses can be taken Pass/Not Pass toward a minor. Visit the ICS Student Affairs Office website for Majors and Minors restrictions. (http://www.ics.uci.edu/ugrad/degrees/MajorMinor_Restrictions_Chart.pdf)

Graduate Programs in Statistics

Research in statistics can range from mathematical studies of the theoretical underpinnings of a statistical model or method to the development of novel statistical models and methods and a thorough study of their properties. Frequently, statistics research is motivated and informed by collaborations with experts in a particular substantive field. Their scientific studies and data collection efforts may yield complex data that cannot be adequately handled using standard statistical methodology. Statisticians aim to develop methods that address the scientific or policy questions of the researcher. In doing so, statisticians must consider how efficiently and effectively the proposed methodology can be implemented and what guarantees can be provided as to the performance of the proposed methods. Such questions can often be answered using a combination of mathematical, analytical, and computational techniques.

Background: Individuals from a variety of backgrounds can make significant contributions to the field of statistics as long as they have sufficient background in statistics, mathematics, and computing. Undergraduate preparation in statistics, mathematics, and computing should include multivariate calculus (the equivalent of UCI courses MATH 2A-MATH 2B, MATH 2D-MATH 2E), linear algebra (MATH 121A), elementary analysis (MATH 140A-MATH 140B), introductory probability and statistics (STATS 120A-STATS 120B-STATS 120C), and basic computing (I&C SCI 31). For students with undergraduate majors outside of mathematics and statistics, it is possible to make up one or two missing courses during the first year in the program.

Students may be admitted to either the master’s program or the doctoral program. For additional information about the Bren School of ICS’s graduate programs and admissions information, click here.

Master of Science in Statistics

Course Requirements

A. Complete:

<table>
<thead>
<tr>
<th>Course</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>STATS 210</td>
<td>Statistical Methods I: Linear Models</td>
</tr>
<tr>
<td>STATS 210B</td>
<td>Statistical Methods II: Categorical Data ¹</td>
</tr>
<tr>
<td>STATS 210C</td>
<td>Statistical Methods III: Longitudinal Data ¹</td>
</tr>
<tr>
<td>STATS 205</td>
<td>Introduction to Bayesian Data Analysis</td>
</tr>
</tbody>
</table>

B. Complete three quarters of seminar in STATS 280.

C. Select five additional graduate courses in or related to statistics, at least two of which are offered by the Department of Statistics. ²

¹ STATS 211 and STATS 212 may be substituted for STATS 210B and STATS 210C.
² At most one of the five elective courses may be an Individual Study (STATS 299), and only with prior approval of the Department Graduate Committee. STATS 281A-STATS 281B-STATS 281C may not be taken as an elective.

The entire program of courses must be approved by the Statistics Department Graduate Committee. Students with previous graduate training in statistics may petition the Committee to substitute other courses for a subset of the required courses. Students are required to pass a written comprehensive examination ordinarily at the end of the first year, covering the material from STATS 200A-STATS 200B-STATS 200C, and either STATS 210, STATS 210B, and STATS 210C, or STATS 210, STATS 211, and STATS 212.
Doctor of Philosophy in Statistics

Statistics Course Requirements

A. Complete:

<table>
<thead>
<tr>
<th>Course Code</th>
<th>Course Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>STATS 200A</td>
<td>Intermediate Probability and Statistical Theory</td>
</tr>
<tr>
<td>STATS 200B</td>
<td>and Intermediate Probability and Statistical Theory</td>
</tr>
<tr>
<td>STATS 200C</td>
<td>and Intermediate Probability and Statistical Theory</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Course Code</th>
<th>Course Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>STATS 210</td>
<td>Statistical Methods I: Linear Models</td>
</tr>
<tr>
<td>STATS 211</td>
<td>Statistical Methods II: Generalized Linear Models</td>
</tr>
<tr>
<td>STATS 212</td>
<td>Statistical Methods III: Methods for Correlated Data</td>
</tr>
<tr>
<td>STATS 220A</td>
<td>Advanced Probability and Statistics Topics</td>
</tr>
<tr>
<td>STATS 220B</td>
<td>and Advanced Probability and Statistics Topics</td>
</tr>
<tr>
<td>STATS 225</td>
<td>Bayesian Statistical Analysis</td>
</tr>
<tr>
<td>STATS 230</td>
<td>Statistical Computing Methods</td>
</tr>
<tr>
<td>STATS 275</td>
<td>Statistical Consulting</td>
</tr>
</tbody>
</table>

B. Select four additional graduate courses in or related to statistics, at least two of which are offered by the Department of Statistics. \(^1\)

C. In addition, continual enrollment in STATS 280 is required in all quarters.

\(^1\) STATS 202, STATS 203, and STATS 281A-STATS 281B-STATS 281C may not be taken as electives. These courses must be completed prior to candidacy.

Additional Ph.D. requirements

Each Ph.D. student is required to take a written comprehensive examination, ordinarily at the end of the first year, covering the material from STATS 200A-STATS 200B-STATS 200C, STATS 210, STATS 211, and STATS 212. In addition, each student is required to take a written comprehensive examination after completion of the second year course work, covering material from STATS 220A-STATS 220B.

Ph.D. students who have passed the written comprehensive examinations are required to give a post-comprehensive research presentation each year.

Ph.D. students are required to serve as teaching assistants for at least two quarters.

Ph.D. students are required to demonstrate substantive knowledge of an application area outside of statistics (e.g., computer science, economics, cognitive sciences, biology, or medicine). Such knowledge can be demonstrated by course work in the application area (three quarter courses), co-authorship of publishable research in the application area, or other evidence of supervised collaborative work that is substantiated by an expert in the field. In the case of a theoretically oriented student, the outside application area may be mathematics.

The normative time for advancement to candidacy is three years. The normative time for completion of the Ph.D. is five years, and the maximum time permitted is seven years.

Faculty

Brigitte Baldi, Ph.D. Massachusetts Institute of Technology, Lecturer of Statistics

Scott Bartell, Ph.D. University of California, Davis, Associate Professor of Program in Public Health; Environmental Health Sciences; Social Ecology; Statistics

Carter Butts, Ph.D. Carnegie Mellon University, Professor of Sociology; Electrical Engineering and Computer Science; Statistics (mathematical sociology, social networks, quantitative methodology, human judgment and decision making, economic sociology)

Daniel L. Gillen, Ph.D. University of Washington, Department Chair and Professor of Statistics; Program in Public Health

Michele Guindani, Ph.D. Universita Luigi Bocconi, Associate Professor of Statistics

Matthew Harding, Ph.D. Massachusetts Institute of Technology, Associate Professor of Economics; Statistics

Ivan G. Jeliazkov, Ph.D. Washington University, Associate Professor of Economics; Statistics

Wesley O. Johnson, Ph.D. University of Minnesota, Professor Emeritus of Statistics

Volodymyr Minin, Ph.D. University of California, Los Angeles, Professor of Statistics

Bin Nan, Ph.D. University of Washington, Professor of Statistics

Hernando C. Ombao, Ph.D. University of Michigan, Professor Emeritus of Statistics; Cognitive Sciences
Dale J. Poirier, Ph.D. University of Wisconsin-Madison, *Professor Emeritus of Economics; Statistics*
Babak Shahbaba, Ph.D. University of Toronto, *Associate Professor of Statistics; Computer Science*
Weining Shen, Ph.D. North Carolina State University, *Assistant Professor of Statistics*
Patrick J. Smyth, Ph.D. California Institute of Technology, *Professor of Computer Science; Statistics*
Hal S. Stern, Ph.D. Stanford University, *UCI Chancellor's Professor of Statistics; Cognitive Sciences*
Jessica Utts, Ph.D. Pennsylvania State University, *Professor Emeritus of Statistics*
Joachim S. Vandekerckhove, Ph.D. University of Leuven, *Assistant Professor of Cognitive Sciences; Statistics* (response time modeling, model fitting, computational statistics, psychometrics, Bayesian statistics)
Yaming Yu, Ph.D. Harvard University, *Associate Professor of Statistics*
Zhaoxia Yu, Ph.D. William Marsh Rice University, *Associate Professor of Statistics*

**Courses**

**STATS 5. Seminar in Data Science. 1 Unit.**
An introduction to the field of Data Science; intended for entering freshman and transfers.
Grading Option: Pass/no pass only.

Restriction: Information Computer Science Majors only.

**STATS 7. Basic Statistics. 4 Units.**
Introduces basic inferential statistics including confidence intervals and hypothesis testing on means and proportions, t-distribution, Chi Square, regression and correlation. F-distribution and nonparametric statistics included if time permits.

Overlaps with STATS 8, MGMT 7, SOCECOL 13.

Restriction: STATS 7 may not be taken for credit concurrently with or after STATS 110, STATS 111, STATS 112.

(Va)

**STATS 8. Introduction to Biological Statistics . 4 Units.**
Introductory statistical techniques used to collect and analyze experimental and observational data from health sciences and biology. Includes exploration of data, probability and sampling distributions, basic statistical inference for means and proportions, linear regression, and analysis of variance.

Overlaps with SOCECOL 13, MGMT 7, STATS 7.

Restriction: STATS 8 may not be taken for credit concurrently with or after STATS 110, STATS 111, STATS 112.

(Va)

**STATS 67. Introduction to Probability and Statistics for Computer Science. 4 Units.**
Introduction to the basic concepts of probability and statistics with discussion of applications to computer science.
Prerequisite: MATH 2B

Restriction: School of Info & Computer Sci students have first consideration for enrollment. Computer Science Engineering Majors have first consideration for enrollment. STATS 67 may not be taken for credit concurrently with or after STATS 120B.

(Va)

**STATS 68. Statistical Computing and Exploratory Data Analysis. 4 Units.**
Introduces key concepts in statistical computing. Techniques such as exploratory data analysis, data visualization, simulation, and optimization methods, will be presented in the context of data analysis within a statistical computing environment.
Prerequisite: STATS 7 and I&C SCI 31
STATS 110. Statistical Methods for Data Analysis I. 4 Units.
Introduction to statistical methods for analyzing data from experiments and surveys. Methods covered include two-sample procedures, analysis of variance, simple and multiple linear regression.
Prerequisite: STATS 7 or STATS 8 or AP90 or (STATS 120A and STATS 120B and STATS 120C). AP90 with a minimum score of 3
Restriction: School of Info & Computer Sci students only.

STATS 111. Statistical Methods for Data Analysis II. 4 Units.
Introduction to statistical methods for analyzing data from surveys or experiments. Emphasizes application and understanding of methods for categorical data including contingency tables, logistic and Poisson regression, loglinear models.
Prerequisite: STATS 110
Concurrent with STATS 202.

STATS 112. Statistical Methods for Data Analysis III. 4 Units.
Introduction to statistical methods for analyzing longitudinal data from experiments and cohort studies. Topics covered include survival methods for censored time-to-event data, linear mixed models, non-linear mixed effects models, and generalized estimating equations.
Prerequisite: STATS 111
Concurrent with STATS 203.

STATS 115. Introduction to Bayesian Data Analysis. 4 Units.
Basic Bayesian concepts and methods with emphasis on data analysis. Special emphasis on specification of prior distributions. Development for one-two samples and on to binary, Poisson, and linear regression. Analyses performed using free OpenBugs software.
Prerequisite: STATS 120C. Recommended: STATS 110.
Concurrent with STATS 205.

STATS 120A. Introduction to Probability and Statistics I. 4 Units.
Introduction to basic principles of probability and statistical inference. Axiomatic definition of probability, random variables, probability distributions, expectation.
Prerequisite: MATH 2A and MATH 2B and (MATH 2D or MATH 4)
Overlaps with MATH 130A.
Restriction: Data Science Majors have first consideration for enrollment. Quantitative Economics majors have second consideration.
Concurrent with STATS 281A.

STATS 120B. Introduction to Probability and Statistics II. 4 Units.
Introduction to basic principles of probability and statistical inference. Point estimation, interval estimating, and testing hypotheses, Bayesian approaches to inference.
Prerequisite: STATS 120A
Restriction: Data Science Majors have first consideration for enrollment. Quantitative Economics majors have second consideration.
Concurrent with STATS 281B.

STATS 120C. Introduction to Probability and Statistics III. 4 Units.
Introduction to basic principles of probability and statistical inference. Linear regression, analysis or variance, model checking.
Prerequisite: STATS 120B and (MATH 3A or MATH 6G or I&C SCI 6N)
Restriction: Data Science Majors have first consideration for enrollment. Quantitative Economics majors have second consideration.
Concurrent with STATS 281C.
STATS 140. Multivariate Statistical Methods. 4 Units.
Theory and application of multivariate statistical methods. Topics include statistical inference for the multivariate normal model and its extensions to multiple samples and regression, use of statistical packages for data visualization and reduction, discriminant analysis, cluster analysis, and factor analysis.
Prerequisite: STATS 120C and (MATH 3A or I&C SCI 6N)
Concurrent with STATS 240.

STATS 170A. Project in Data Science I. 4 Units.
Problem definition and analysis, data representation, algorithm selection, solution validation, and results presentation. Students do team projects and lectures cover analysis alternatives, project planning, and data analysis issues. First quarter emphasizes approach selection, project planning, and experimental design.
Prerequisite: STATS 68 and STATS 112 and IN4MATX 43 and COMPSCI 122A and COMPSCI 161 and COMPSCI 178
Grading Option: In progress only.
Restriction: Seniors only. Data Science Majors have first consideration for enrollment.

STATS 170B. Project in Data Science II. 4 Units.
Problem definition and analysis, data representation, algorithm selection, solution validation, and results presentation. Students do team projects and lectures cover analysis alternatives, project planning, and data analysis issues. Second quarter emphasizes project execution and analysis, and presentation of results.
Prerequisite: STATS 170A
Restriction: Seniors only. Data Science Majors have first consideration for enrollment.

STATS 199. Individual Study. 2-5 Units.
Individual research or investigations under the direction of an individual faculty member.
Repeatability: May be repeated for credit unlimited times.

STATS 200A. Intermediate Probability and Statistical Theory. 4 Units.
Basics of probability theory, random variables and basic transformations, univariate distributions—discrete and continuous, multivariate distributions.
Prerequisite: STATS 120C

STATS 200B. Intermediate Probability and Statistical Theory. 4 Units.
Random samples, transformations, limit laws, normal distribution theory, introduction to stochastic processes, data reduction, point estimation (maximum likelihood).
Prerequisite: STATS 200A

STATS 200C. Intermediate Probability and Statistical Theory. 4 Units.
Interval estimation, hypothesis testing, decision theory and Bayesian inference, basic linear model theory.
Prerequisite: STATS 200B

STATS 201. Statistical Methods for Data Analysis I. 4 Units.
Introduction to statistical methods for analyzing data from experiments and surveys. Methods covered include two-sample procedures, analysis of variance, simple and multiple linear regression.
Prerequisite: STATS 7 or STATS 8
Restriction: STATS 201 cannot be taken for credit after taking STATS 210.

STATS 202. Statistical Methods for Data Analysis II. 4 Units.
Introduction to statistical methods for analyzing data from surveys or experiments. Emphasizes application and understanding of methods for categorical data including contingency tables, logistic and Poisson regression, loglinear models.
Prerequisite: STATS 201 or STATS 210
Concurrent with STATS 111.
STATS 203. Statistical Methods for Data Analysis III. 4 Units.
Introduction to statistical methods for analyzing longitudinal data from experiments and cohort studies. Topics covered include survival methods for censored time-to-event data, linear mixed models, non-linear mixed effects models, and generalized estimating equations.

Prerequisite: STATS 202
Concurrent with STATS 112.

STATS 205. Introduction to Bayesian Data Analysis. 4 Units.
Basic Bayesian concepts and methods with emphasis on data analysis. Special emphasis on specification of prior distributions. Development for one-two samples and on to binary, Poisson and linear regression. Analyses performed using free OpenBugs software.

Prerequisite: STATS 120C. Recommended: STATS 201 or STATS 210.
Concurrent with STATS 115.

STATS 210. Statistical Methods I: Linear Models. 4 Units.
Statistical methods for analyzing data from surveys and experiments. Topics include randomization and model-based inference, two-sample methods, analysis of variance, linear regression and model diagnostics.

Prerequisite: Knowledge of basic statistics, calculus, linear algebra.

STATS 210A. Statistical Methods I: Linear Models. 4 Units.
Statistical methods for analyzing data from surveys and experiments. Topics include randomization and model-based inference, two-sample methods, analysis of variance, linear regression, and model diagnostics.

Prerequisite: Knowledge of basic statistics (at the level of STATS 7), calculus, and linear algebra.
Restriction: Graduate students only.

STATS 210B. Statistical Methods II: Categorical Data. 4 Units.
Introduction to statistical methods for analyzing discrete and non-normal outcomes. Emphasizes the development and application of methods for categorical data, including contingency tables, logistic and Poisson regression, loglinear models.

Prerequisite: STATS 210. May not be taken for graduate credit by Ph.D. students in Statistics.
Restriction: Graduate students only.

STATS 210C. Statistical Methods III: Longitudinal Data. 4 Units.
Introduction to statistical methods for analyzing longitudinal outcomes. Emphasizes the development and application of regression methods for correlated and censored outcomes. Methods for continuous and discrete correlated outcomes, as well as censored outcomes, are covered.

Prerequisite: STATS 210B. May not be taken for graduate credit by Ph.D. students in Statistics.
Restriction: Graduate students only.

STATS 211. Statistical Methods II: Generalized Linear Models. 4 Units.
Development of the theory and application of generalized linear models. Topics include likelihood estimation and asymptotic distributional theory for exponential families, quasi-likelihood and mixed model development. Emphasizes methodological development and application to real scientific problems.

Prerequisite or corequisite: STATS 210

STATS 212. Statistical Methods III: Methods for Correlated Data. 4 Units.
Development and application of statistical methods for analyzing correlated data. Topics covered include repeated measures ANOVA, linear mixed models, non-linear mixed effects models, and generalized estimating equations. Emphasizes both theoretical development and application of the presented methodology.

Prerequisite: STATS 211

STATS 220A. Advanced Probability and Statistics Topics. 4 Units.
Advanced topics in probability and statistical inference including measure theoretic probability, large sample theory, decision theory, resampling and Monte Carlo methods, nonparametric methods.

Prerequisite: STATS 200C
STATS 220B. Advanced Probability and Statistics Topics. 4 Units.
Advanced topics in probability and statistical inference, including measure theoretic probability, large sample theory, decision theory, resampling and Monte Carlo methods, nonparametric methods.
Prerequisite: STATS 220A and MATH 140B

STATS 225. Bayesian Statistical Analysis. 4 Units.
Introduction to the Bayesian approach to statistical inference. Topics include univariate and multivariate models, choice of prior distributions, hierarchical models, computation including Markov chain Monte Carlo, model checking, and model selection.
Prerequisite: STATS 205 and STATS 230

STATS 226. Advanced Topics in Modern Bayesian Statistical Inference. 4 Units.
Modern Bayesian Statistics: selected topics from theory of Markov chains, application of theory to modern methods of Markov chain Monte Carlo sampling; Bayesian non-parametric and semiparametric modeling, including Dirichlet Process Mixtures; Mixtures of Polya Trees.
Prerequisite: STATS 200C and STATS 225

STATS 230. Statistical Computing Methods. 4 Units.
Numerical computations and algorithms with applications in statistics. Topics include optimization methods including the EM algorithm, random number generation and simulation, Markov chain simulation tools, and numerical integration.
Prerequisite: Two quarters of upper-division or graduate training in probability and statistics.
Overlaps with COMPSCI 206.

STATS 235. Modern Data Analysis Methods . 4 Units.
Introduces selected modern tools for data analysis. Emphasizes use of computational and resampling techniques for data analyses when the data do not conform to standard toolbox of regression models and/or complexity of modeling problem threatens validity of standard methods.
Prerequisite: STATS 120C and STATS 205 and (STATS 201 or STATS 210)
Restriction: Graduate students only.

STATS 240. Multivariate Statistical Methods. 4 Units.
Theory and application of multivariate statistical methods. Topics include statistical inference for the multivariate normal model and its extensions to multiple samples and regression, use of statistical packages for data visualization and reduction, discriminant analysis, cluster analysis, and factor analysis.
Prerequisite: STATS 120C and (MATH 3A or I&C SCI 6N)
Concurrent with STATS 140.

STATS 245. Time Series Analysis. 4 Units.
Statistical models for analysis of time series from time and frequency domain perspectives. Emphasizes theory and application of time series data analysis methods. Topics include ARMA/ARIMA models, model identification/estimation, linear operators, Fourier analysis, spectral estimation, state space models, Kalman filter.
Corequisite: STATS 200C
Prerequisite or corequisite: STATS 201 or STATS 210

STATS 246. Spectral Analysis . 4 Units.
Spectral methods that are most commonly utilized for analyzing univariate and multivariate time series and signals. These methods include spectral and coherence estimation, transfer function modeling, classification and discrimination of time series, non-stationary time series, time-frequency analysis, and wavelets analysis.
Prerequisite: STATS 200B and (STATS 201 or STATS 210)

STATS 250. Biostatistics. 4 Units.
Statistical methods commonly used to analyze data arising from clinical studies. Topics include analysis of observational studies and randomized clinical trials, techniques in the analysis of survival and longitudinal data, approaches to handling missing data, meta-analysis, nonparametric methods.
Prerequisite: STATS 210
STATS 255. Statistical Methods for Survival Data. 4 Units.
Statistical methods for analyzing survival data from cohort studies. Topics include parametric and nonparametric methods, the Kaplan-Meier estimator, log-rank tests, regression models, the Cox proportional hazards model and accelerated failure time models, efficient sampling designs, discrete survival models.

Corequisite: STATS 202 or STATS 211.
Prerequisite: STATS 210

STATS 257. Introduction to Statistical Genetics. 4 Units.
Provides students with knowledge of the basic principles, concepts, and methods used in statistical genetic research. Topics include principles of population genetics, and statistical methods for family- and population-based studies.

Prerequisite: Two quarters of upper-division or graduate training in statistical methods.

Same as EPIDEM 215.

STATS 260. Inference with Missing Data. 4 Units.
Statistical methods and theory useful for analysis of multivariate data with partially observed variables. Bayesian and likelihood-based methods developed. Topics include EM-type algorithms, MCMC samplers, multiple imputation, and general location model. Applications from economics, education, and medicine are discussed.

Prerequisite or corequisite: STATS 210 or STATS 200C. STATS 230.

STATS 262. Theory and Practice of Sample Surveys. 4 Units.
Basic techniques and statistical methods used in designing surveys and analyzing collected survey data. Topics include simple random sampling, ratio and regression estimates, stratified sampling, cluster sampling, sampling with unequal probabilities, multistage sampling, and methods to handle nonresponse.

Prerequisite: STATS 120C

STATS 265. Causal Inference. 4 Units.
Various approaches to causal inference focusing on the Rubin causal model and propensity-score methods. Topics include randomized experiments, observational studies, non-compliance, ignorable and non-ignorable treatment assignment, instrumental variables, and sensitivity analysis. Applications from economics, politics, education, and medicine.

Prerequisite: STATS 200C and STATS 210

STATS 270. Stochastic Processes. 4 Units.
Introduction to the theory and application of stochastic processes. Topics include Markov chains, continuous-time Markov processes, Poisson processes, and Brownian motion. Applications include Markov chain Monte Carlo methods and financial modeling (for example, option pricing).

Prerequisite: STATS 120C

Overlaps with MATH 271A, MATH 271B, MATH 271C.

STATS 275. Statistical Consulting. 4 Units.
Training in collaborative research and practical application of statistics. Emphasis on effective communication as it relates to identifying scientific objectives, formulating a statistical analysis plan, choice of statistical methods, and interpretation of results and their limitations to non-statisticians.

Prerequisite: STATS 203 or STATS 212

Repeatability: May be taken for credit 2 times.

STATS 280. Seminar in Statistics. 0.5 Units.
Periodic seminar series covering topics of current research in statistics and its application.

Grading Option: Satisfactory/unsatisfactory only.

Repeatability: May be repeated for credit unlimited times.

Restriction: Graduate students only.
STATS 281A. Introduction to Probability and Statistics I. 4 Units.
Introduction to basic principles of probability and statistical inference. Axiomatic definition of probability, random variables, probability distributions, expectation.

Restriction: Graduate students only.

Concurrent with STATS 120A.

STATS 281B. Introduction to Probability and Statistics II. 4 Units.
Introduction to basic principles of probability and statistical inference. Point estimation, interval estimating, and testing hypotheses, Bayesian approaches to inference.

Restriction: Graduate students only.

Concurrent with STATS 120B.

STATS 281C. Introduction to Probability and Statistics III. 4 Units.
Introduction to basic principles of probability and statistical inference. Contingency table analysis, linear regression, analysis of variance, model checking.

Restriction: Graduate students only.

Concurrent with STATS 120C.

STATS 295. Special Topics in Statistics. 4 Units.
Studies in selected areas of statistics. Topics addressed vary each quarter.

Repeatability: Unlimited as topics vary.

STATS 298. Thesis Supervision. 2-12 Units.
Individual research or investigation conducted in preparation for the M.S. thesis option or the dissertation requirements for the Ph.D. program.

Repeatability: May be repeated for credit unlimited times.

STATS 299. Individual Study. 2-12 Units.
Individual research or investigation under the direction of an individual faculty member.

Repeatability: May be repeated for credit unlimited times.