Statistics

Degree Offered

- Master of Science

Nature of the Program

The Department of Statistics offers a Master of Science (M.S.) in Statistics. The M.S. degree is intended to qualify the student to assume a professional role in educational, industrial, or governmental research projects; to teach in a college; or to undertake advanced training toward a doctorate in statistics or one of the quantitative fields of science.

Because many students receive baccalaureate degrees from colleges that do not offer undergraduate programs in statistics, and because historically statistics has been primarily a field of graduate education, a student does not need a degree in statistics to enter the degree program. A good background in mathematics, science, or engineering is reasonable preparation for graduate work in statistics.

The Department of Statistics also participates in the Combinatorial Computing and Discrete Mathematics (CCDM) Area of Emphasis within the Computer and Information Science Ph.D. Program or the Mathematics Ph.D. Program.

The Department of Statistics offers a Certificate in Applied Statistics for professionals or students who want to take applied statistics courses to enhance their quantitative skills and job opportunities.

FACULTY

CHAIR
- Mark V. Culp - Ph.D. (University of Michigan)
  Statistical Machine Learning, Computational Statistics, Semi-supervised and Multi-view Learning, Biometrics

PROFESSOR
- Kenneth J. Ryan - Ph.D. (Iowa State University)
  Experimental Design, Statistical Machine Learning, Biometrics

TEACHING ASSOCIATE PROFESSOR
- Huey Miin Lee - Ph.D. (Johns Hopkins University)
  Bioinformatics, Statistical Education

ASSISTANT PROFESSOR
- Stacey Culp - Ph.D. (University of Michigan)
  Experimental Design, Healthcare Applications

TEACHING INSTRUCTOR
- Anthony Billings - M.S. (West Virginia University); A.B.D. (Carnegie Mellon University)
  Statistical Computing, Statistical Modeling, Robust Estimation, Nonlinear Dynamic Systems, Statistical Education

PROFESSOR EMERITUS
- Erdogan Gunel - Ph.D. (State University of New York, Buffalo)
  Bayesian Inference, Biostatistics, Categorical Data Analysis
- E. James Harner - Ph.D. (Cornell University)
- William V. Thayne - Ph.D. (University of Illinois)
  Experimental Design, Statistical Genetics, Regression Analysis
- Edwin C. Townsend - Ph.D. (Cornell University)
  Experimental Design, Regression Analysis

ASSOCIATE PROFESSOR EMERITUS
- Daniel M. Chilko - M.S. (Rutgers University)
  Statistical Computing, Computer Graphics
• Gerald R. Hobbs Jr. - Ph.D. (Kansas State University)
  Biostatistics, Nonparametric Statistics, Regression Analysis

Admissions

ADMISSIONS AND PREREQUISITES FOR MASTER OF SCIENCE IN STATISTICS

Students are expected to know the material contained in the following courses or areas upon admission to the program. Otherwise, these deficiencies must be removed as early as possible in the student’s degree program under the terms specified by the Admissions and Standards Committee.

- Single and multivariable calculus (MATH 155, MATH 156, MATH 251, or equivalent)
- Linear or matrix algebra (MATH 441 or equivalent)
- Probability and statistics (STAT 215 or equivalent)
- Knowledge of a high-level programming language

ADMISSIONS AND PREREQUISITES FOR THE CERTIFICATE IN APPLIED STATISTICS

Admission to the Certificate of Applied Statistics (CAS) may be done at any time. Students who are currently admitted to or enrolled in a graduate degree program that want to earn the CAS should contact the Statistics Department to enroll in the certificate program. Students who want to pursue the CAS independent of a graduate degree program must be admitted as a non-degree graduate student prior to registering for the certificate program.

Those seeking admission to the CAS must have a minimum GPA of 2.75, have graduated from and accredited institution with a minimum of a Baccalaureate degree and successfully completed College Algebra. Single and Multi-variable Calculus are recommended.

The GRE General Test is not required for admission.

To obtain a Master of Science in Statistics, the student must complete the course and comprehensive examination requirements. The student must maintain a minimum GPA of 3.0 and earn a grade of C- or better in all courses counting toward the degree.

Master of Science

MAJOR REQUIREMENTS

To obtain a Master of Science in Statistics, the student must complete the course and comprehensive examination requirements.

Minimum cumulative GPA of 3.0 is required.

Minimum grade of C- in all courses applied toward the degree

<table>
<thead>
<tr>
<th>Course</th>
<th>Title</th>
<th>Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAT 512</td>
<td>Statistical Methods 2</td>
<td>3</td>
</tr>
<tr>
<td>STAT 513</td>
<td>Design of Experiments</td>
<td>3</td>
</tr>
<tr>
<td>STAT 545</td>
<td>Applied Regression Analysis</td>
<td>3</td>
</tr>
<tr>
<td>STAT 555</td>
<td>Categorical Data Analysis</td>
<td>3</td>
</tr>
<tr>
<td>STAT 561</td>
<td>Theory of Statistics 1</td>
<td>3</td>
</tr>
<tr>
<td>STAT 562</td>
<td>Theory of Statistics 2</td>
<td>3</td>
</tr>
<tr>
<td>Electives</td>
<td>(STAT 462, any 500-, 600-, or 700-level STAT courses except STAT 511 or STAT 516)</td>
<td>9</td>
</tr>
</tbody>
</table>

Select either the non-thesis or thesis option

<table>
<thead>
<tr>
<th>Option</th>
<th>Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Thesis Option:</td>
<td>6</td>
</tr>
<tr>
<td>Electives (STAT 462, any 500-, 600-, or 700-level STAT courses except STAT 511 or STAT 516)</td>
<td>9</td>
</tr>
<tr>
<td>Thesis Option:</td>
<td></td>
</tr>
<tr>
<td>STAT 697</td>
<td>Research</td>
</tr>
<tr>
<td>Thesis</td>
<td>6</td>
</tr>
<tr>
<td>Thesis Defense</td>
<td>6</td>
</tr>
<tr>
<td>Comprehensive Examination</td>
<td>3</td>
</tr>
<tr>
<td>Total Hours</td>
<td>33</td>
</tr>
</tbody>
</table>

* Non-STAT electives require departmental consent.

EXAMINATIONS

Students must pass a written comprehensive examination on foundational material. The examination covers the theory taught in STAT 461 and STAT 462 and the applications taught in STAT 512, STAT 513, and STAT 545. The exam is given twice a year on the Thursday during the second full week following spring semester final exams and on the third Saturday in October. Students have a maximum of three attempts for this exam.
Major Learning Outcomes

STATISTICS

Graduate courses in statistics, and sequences of statistics courses leading to a Master of Science in Statistics or a Certificate in Applied Statistics, provide a foundation of statistical literacy, statistical reasoning, and statistical thinking. Our aim is for all of our students to be challenged and encouraged in their statistical course work. In particular, we enable our students to

• Appreciate the inherent variation and uncertainty of information, and understand that statistics can be a resource for improved decision making;
• Develop critical thinking skills for application of statistics;
• Effectively communicate the results of statistical analysis;
• Become responsible and competent practitioners of statistics in order to attain personal goals, either in a profession or in further educational experiences.

Certificate in Applied Statistics

CERTIFICATE CODE - CG29

The Certificate in Applied Statistics (CAS) is designed for professionals or students who want to take applied statistics courses to enhance their quantitative skills and job opportunities. The certificate will provide students with a solid foundation in statistical methodology, and depending on the elective courses selected, predictive analytics, statistical computing, or statistical theory. The flexibility in the certificate course work is intended to allow the student to select courses that will meet their needs, whether enhancing professional quantitative skills or research productivity.

Admissions to the CAS may be done at any time. Students who are currently admitted to or enrolled in a graduate degree program that are wishing to earn the CAS should contact the Statistics Department to enroll in the CAS. Students who wish to pursue the CAS independent of a graduate degree program must be admitted as a non-degree graduate student prior to registering their intent to earn the certificate.

Students must earn a grade of C- in all courses applied to the CAS, and must earn at least an overall 3.0 GPA in the courses counted toward the certificate.

REQUIRED COURSES:

Students in the certificate program will complete a minimum of 15 credit hours of graduate level Statistics courses. The courses required for the completion of the CAS are defined below.

Students must earn a grade of C- in all courses applied to the CAS.

Students must earn at least an overall 3.0 GPA in the courses counted toward the certificate.

<table>
<thead>
<tr>
<th>Course</th>
<th>Title</th>
<th>Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAT 512</td>
<td>Statistical Methods 2</td>
<td>3</td>
</tr>
<tr>
<td>STAT 513</td>
<td>Design of Experiments</td>
<td>3</td>
</tr>
<tr>
<td>Electives (500, 600, 700-level STAT Courses)</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Total Hours</td>
<td></td>
<td>15</td>
</tr>
</tbody>
</table>

* Credit towards the Certificate is also given for STAT 461 and STAT 462.
** All courses applied to the certificate must be Statistics (STAT) courses; courses listed as equivalent to Statistics courses in the Catalog may not be counted.

COURSES

STAT 511. Statistical Methods 1. 3 Hours.
PR: MATH 126. Statistical models, distributions, probability, random variables, tests of hypotheses, confidence intervals, regression, correlation, transformations, F and Chi-square distributions, analysis of variance and multiple comparisons. (Equivalent to EDP 613 and PSYC 511.).

STAT 512. Statistical Methods 2. 3 Hours.
PR: STAT 511 or equivalent. Completely random, randomized complete block, Latin square, and split-plot experimental designs. Unplanned and planned multiple and orthogonal comparisons for qualitative and quantitative treatments and factorial arrangements. Multiple linear regression and covariance analysis. (Equivalent to EDP 614 and PSYC 512.).

STAT 513. Design of Experiments. 3 Hours.
PR: STAT 512 or equivalent. Expected mean squares, power of tests and relative efficiency for various experimental designs. Fixed, random, and mixed models. Use of sub-sampling, covariance, and confounding to increase power and efficiency.

STAT 516. Forensic Statistics. 3 Hours.
PR: STAT 215 or equivalent. Probabilistic and statistical evaluation of evidence in forensic science: concepts of uncertainty variation, discriminating power, coincidence/significance probabilities, historical overview, transfer evidence, DNA profiling, fingerprint identification, biometric identification, and case studies.
STAT 521. Statistical Analysis System Programming. 3 Hours.

STAT 522. Advanced Statistical Analysis System Programming. 3 Hours.
PR: STAT 521 or consent. Advanced topics in Statistical Analysis System (SAS); SAS SQL to generate reports, join tables, construct queries; SAS Macrolanguage basics; write/implement SAS macro programs. Prepares students for SAS Advanced Programmer Certification Exam.

STAT 523. Statistical Computing. 3 Hours.
PR: STAT 512. Monte Carlo methods; randomization, partitioning, and the bootstrap; identifying data structures, estimating functions, including density functions; statistical models of dependencies. R programming.

STAT 531. Sampling Theory and Methods. 3 Hours.
PR: STAT 511 or consent. Survey components, methods of sampling for finite and infinite populations, single and multi-stage procedures, confidence limits for estimating population parameters, sample size determination, area sampling sources of survey error, and basic inference derived from survey design.

STAT 540. Introduction to Exploratory Data Analysis. 3 Hours.
PR: An introductory statistics course. Basic ways in which observations given in counted and measured form are approached. Pictorial and arithmetic techniques of display and discovery. Methods employed are robust, graphical, and informal. Applications to social and natural sciences. (Alternate years.).

STAT 541. Applied Multivariate Analysis. 3 Hours.
PR: STAT 511 or equivalent. Introduction to Euclidean geometry and matrix algebra; multiple and multivariate regression including multiple and canonical correlation; the k-sample problem including discriminant and canonical analysis; and structuring data by factor analysis, cluster analysis, and multi-dimensional scaling.

STAT 543. Bioinformatics Data Analysis. 3 Hours.
PR: STAT 512 or equivalent. Statistical analyses of high-throughput experiments using data visualization, clustering, multiple testing, classification and other unsupervised and supervised learning methods. Data processing, including background adjustment and normalization. Case studies.

STAT 545. Applied Regression Analysis. 3 Hours.
PR: STAT 512 or equivalent. Matrix approach to linear and multiple regression, selecting the best regression equation, model building, and the linear models approach to analysis of variance and analysis of covariance.

STAT 547. Survival Analysis. 3 Hours.
PR: STAT 512. Survival model methodology, including model selection for incomplete data with censored, truncated, and interval censored observations. Applications to many real life problems using R.

STAT 551. Nonparametric Statistics. 3 Hours.
PR: STAT 511 or equivalent. Distribution-free procedures of statistical inference. Location and scale tests for homogeneity with two or more samples (related or independent); tests against general alternatives.

STAT 555. Categorical Data Analysis. 3 Hours.
PR: STAT 512 or equivalent. Bivariate association for ordinal and nominal variables, models for categorical or continuous responses as a special case of generalized linear models, methods for repeated measurement data, exact small-sample procedures.

STAT 561. Theory of Statistics 1. 3 Hours.
PR: MATH 251. Probability and random variables, univariate and multivariate distributions, expectations, generating functions, marginal and conditional distributions, independence, correlation, functions of random variables, including order statistics, limiting distributions, and stochastic convergence.

STAT 562. Theory of Statistics 2. 3 Hours.
PR: STAT 561. Techniques of point and interval estimation; properties of estimates including bias, consistency, efficiency, and sufficiency; hypothesis testing including likelihood ratio tests and Neyman-Pearson Lemma; Bayesian procedures; analysis of variance and nonparametrics.

STAT 582. Statistical Consulting. 1 Hour.
PR: STAT 513 or Consent. Statistical consulting principles and procedures. The entire consulting experience, including design, models, communication skills, ethics, tracking, and documentation, is presented in a series of case studies, including student presentations and reports on assigned cases.

STAT 590. Teaching Practicum. 1-3 Hours.
PR: Consent. Supervised practice in college teaching of statistics. Note: This course is intended to insure that graduate assistants are adequately prepared and supervised when they are given college teaching responsibility. It will also present a mechanism for students not on assistantships to gain teaching experience. (Grading may be S/U.).

STAT 591. Advanced Topics. 1-6 Hours.
PR: Consent. Investigation in advanced topics not covered in regularly scheduled courses.

STAT 593. Special Topics. 1-6 Hours.
A study of contemporary topics selected from recent developments in the field.

STAT 595. Independent Study. 1-6 Hours.
STAT 595. Independent Study. 1-6 HR. Faculty supervised study of topics not available through regular course offerings.
STAT 623. Data Technologies. 3 Hours.
PR: STAT 512 or consent. R data manipulation and processing. Topics include: R operators, functions, data structures, and objects; R data input and output, package development, and text processing; R interfaces to XML and SQL databases.

STAT 624. High Performance Analytics. 3 Hours.
PR: STAT 623. High performance and data-stream computing using R. Topics include: parallel R packages; Hadoop clusters; MapReduce R scripting; shared R network spaces; beyond-memory data analysis; data-stream modeling and visualization.

STAT 641. Multivariate Statistical Theory. 3 Hours.
PR: STAT 541, and STAT 561 or consent. Euclidean vector space theory and matrix algebra, multivariate normal sampling theory, the theory of the multivariate general linear hypothesis including multivariate regression, MANOVA, and MANCOVA, and the theory of factor analysis.

STAT 645. Linear Models. 3 Hours.
PR: STAT 545 and (STAT 462 or STAT 562) or consent. Multivariate normal distribution, distribution of quadratic forms, linear models, general linear hypotheses, experimental design models, components of variance for random effects models.

STAT 682. Statistics Practicum. 1 Hour.
PR: STAT 582. Statistical consulting on university-related research projects under the direction of a statistics faculty member.

STAT 689. Professional Field Experience. 1-6 Hours.
PR: Consent. (May be repeated up to a maximum of 18 hours). Prearranged experiential learning program, to be planned, supervised, and evaluated for credit by faculty and field supervisors. Involves temporary placement with public or private enterprise for professional competence development.

STAT 690. Teaching Practicum. 1-3 Hours.
PR: Consent. Supervised practice in college teaching of statistics. Note: This course is intended to insure that graduate assistants are adequately prepared and supervised when they are given college teaching responsibility. It also provides a mechanism for students not on assistantships to gain teaching experience. (Grading may be S/U.).

STAT 691. Advanced Topics. 6 Hours.
PR: Consent. Investigation of advanced topics not covered in regularly scheduled courses.

STAT 692. Directed Study. 1-6 Hours.
Directed study, reading, and/or research.

STAT 693. Special Topics. 1-6 Hours.
A study of contemporary topics selected from recent developments in the field.

STAT 694. Seminar. 1-6 Hours.
Special seminars arranged for advanced graduate students.

STAT 695. Independent Study. 1-6 Hours.
Faculty supervised study of topics not available through regular course offerings.

STAT 696. Graduate Seminar. 1 Hour.
PR: Consent. Each graduate student will present at least one seminar to the assembled faculty and graduate student body of his or her program.

STAT 697. Research. 1-15 Hours.
PR: Consent. Research activities leading to thesis, problem report, research paper or equivalent scholarly project, or a dissertation. (Grading may be S/U.).

STAT 698. Thesis or Dissertation. 1-6 Hours.
PR: Consent. This is an optional course for programs that wish to provide formal supervision during the writing of student reports (698), or dissertations (798). Grading is normal.

STAT 699. Graduate Colloquium. 1-6 Hours.
PR: Consent. For graduate students not seeking coursework credit but who wish to meet residency requirements, use of the University's facilities, and participate in its academic and cultural programs. Note: Graduate students who are not actively involved in coursework or research are entitled, through enrollment in their department's 699/799 Graduate Colloquium to consult with graduate faculty, participate in both formal and informal academic activities sponsored by their program, and retain all of the rights and privileges of duly enrolled students. Grading is P/F; colloquium credit may not be counted against credit requirements for masters programs. Registration for one credit of 699/799 graduate colloquium satisfies the University requirement in the semester in which graduation occurs.

STAT 745. Data Mining. 3 Hours.
PR: STAT 545 or equivalent. Development of predictive models for large datasets, including logistic and linear models, regression and classification trees, and neural networks. Data preparation, including imputation and filtering.

STAT 761. Theoretical Statistics 1. 3 Hours.
PR: STAT 562 or consent. Advanced statistical theory including: consistent estimators; limiting distributions; asymptotic properties; goodness-of-fit tests; maximum likelihood estimation, moment generating functions; properties of statistical tests and procedures for finite-dimensional and infinite-dimensional spaces.
STAT 762. Theoretical Statistics 2. 3 Hours.
PR: STAT 761. Continuation of STAT 761 including: asymptotic optimality, contiguity of probability measures, local asymptotic normality of likelihood ratio test, Bayesian estimation, general linear models estimation and testing, and kernel smoothing methods in density and regression estimation.

STAT 763. Stochastic Processes. 3 Hours.
PR: STAT 561. Modeling of random phenomenon occurring over time, space, or time and space simultaneously. Modern techniques, such as the martingale decomposition, are applied to different statistical models.

STAT 791. Advanced Topics. 1-6 Hours.
PR: Consent. Investigation of advanced topics not covered in regularly scheduled courses.

STAT 797. Research. 1-15 Hours.
PR: Consent. Research activities leading to thesis, problem report, research paper or equivalent scholarly project, or a dissertation. (Grading will be S/U).

STAT P101. . 3 Hours.
STAT P211. . 3 Hours.
STAT T211. . 3 Hours.