Statistics

The Master of Arts in Statistics prepares students to perform in an information-rich, data-driven workforce that requires both general and specialized skills in statistical analysis. The 36-unit program, designed primarily for part-time study, covers essential elements of statistical studies with courses in probability, statistical computation and model building, experimental design, survival analysis, Bayesian statistics, and stochastic processes. Additionally, these courses, along with a required practicum, provide a foundation for further doctoral-level study in mathematics and statistics, or in other academic disciplines such as anthropology, biology, economics, political science, and psychology.

In addition to providing a solid theoretical foundation, the program also offers applied value, providing tools, strategies, and technical skills in areas such as predictive analytics and big data to help professionals in many fields analyze large volumes of data, make reliable and productive business decisions, and use technology efficiently. The program offers flexibility and a wide range of elective and applied courses that emphasize statistical analysis in mathematics, computer science, engineering, clinical investigation, biostatistics, economics and business. Students may choose from a broad-based pool of elective courses across disciplines, or they may organize elective course work and design the required practicum in one of the optional tracks that correspond to strong industry demand for statisticians: Biology and Health; Business and Finance; and Engineering and Materials.

Contact: Lisa Kuehne
Phone: 314-935-4226
Email: lmkuehne@wustl.edu
Website: http://ucollege.wustl.edu/programs/graduate/masters-statistics

Degree Requirements

Master of Arts in Statistics

The Master of Arts in Statistics is a 36-unit program that includes 15 units of required course work, 6 units of required thesis practicum, and 15 units of electives. Students may choose electives broadly from the list below or they have the option of organizing elective course work and designing the required thesis practicum in one of the suggested tracks in Biology and Health; Business and Finance; and Engineering and Materials. Candidates for this degree will have completed the calculus sequence (differential, integral, and multivariable calculus) as well as an intermediate statistics course such as Math 305 prior to beginning graduate study.

A maximum of 6 credits of related and comparable graduate-level course work may be transferred from another university or from a related graduate program at Washington University with the approval of the program director. These must be graduate-level units not used to fulfill undergraduate degree requirements. Transfer credit may be granted only for authorized courses for which the student received a grade of B or higher.

Required Courses (15 Units)

<table>
<thead>
<tr>
<th>Code</th>
<th>Title</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Math 593</td>
<td>Probability</td>
<td>3</td>
</tr>
<tr>
<td>Math 594</td>
<td>Mathematical Statistics</td>
<td>3</td>
</tr>
<tr>
<td>Math 529</td>
<td>Linear Algebra</td>
<td>3</td>
</tr>
<tr>
<td>or Math 5392</td>
<td>Advanced Linear Statistical Models</td>
<td></td>
</tr>
<tr>
<td>Math 539</td>
<td>Linear Statistical Models</td>
<td>3</td>
</tr>
<tr>
<td>Math 575</td>
<td>Statistical Computation</td>
<td>3</td>
</tr>
</tbody>
</table>

Total Units: 15

In the case that an equivalent course has been taken and proficiency in the course material has been demonstrated, other 500-level electives may be substituted in consultation with the adviser.

Required Thesis Practicum (3 Units)

<table>
<thead>
<tr>
<th>Code</th>
<th>Title</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>U20 Math 502</td>
<td>Thesis Practicum</td>
<td>3</td>
</tr>
</tbody>
</table>

Electives (18 Units)

Additional 500-level electives, selected from the list below, will be chosen by the student in consultation with University College, to make up the 36 units. Other 500-level electives may be selected in consultation with an adviser. Students may choose elective courses broadly, or follow one of the suggested tracks.

• U20 Math 5145 Advanced Theoretical Econometrics
• U20 Math 5161 Applied Econometrics
• Math 520 Experimental Design
• Math 534 Survival Analysis
• U20 Math 538 Measurement and Latent Trait Models
• Math 549 Numerical Applied Mathematics
• U20 Math 551 Advanced Probability I
• U20 Math 552 Advanced Probability II
• U20 Math 559 Bayesian Statistics
• Math 584 Multilevel Models in Quantitative Research
• U20 Math 585 Stochastic Processes

Biology and Health Optional Track

• Math 520 Experimental Design
• Math 522 Biostatistics
• Math 534 Survival Analysis
• Other courses with authorization

Business and Finance Optional Track
Courses


U20 Math 500 Independent Study
Credit 3 units.

U20 Math 520 Experimental Design
A first course in the design and analysis of experiments, from the point of view of regression. Factorial, randomized block, split-plot, Latin square, and similar design. Prerequisite: CSE 131 or 200, Math 3200, or permission of instructor. Credit 3 units. Art: NSM

U20 Math 522 Biostatistics
A second course in elementary statistics with applications to life sciences and medicine. Review of basic statistics using biological and medical examples. New topics include incidence and prevalence, medical diagnosis, sensitivity and specificity, Bayes’ rule, decision making, maximum likelihood, logistic regression, ROC curves, and survival analysis. Prerequisites: Math 3200 or a strong performance in Math 2200 and permission of the instructor. Credit 3 units. Art: NSM

U20 Math 529 Linear Algebra
Introduction to the linear algebra of finite-dimensional vector spaces. Includes systems of equations, matrices, determinants, inner product spaces, spectral theory. Prerequisite: Math 310 or permission of instructor. Math 309 is not an explicit prerequisite but students should already be familiar with such basic topics from matrix theory as matrix operations, linear systems, row reduction, and Gaussian elimination. Material on these topics in early chapters of the text will be covered very quickly. Note: Not equivalent to L24 429.
Credit 3 units.

U20 Math 534 Survival Analysis
Life table analysis and testing, mortality and failure rates, Kaplan-Meier or product-limit estimators, hypothesis testing and estimation in the presence of random arrivals and departures, and the Cox proportional hazards model. Techniques of survival analysis are used in medical research, industrial planning, and the insurance industry. Prerequisites: CSE 131 or 200, Math 309 and 3200, or permission of the instructor.
Same as L24 Math 434
Credit 3 units. A&S IQ: NSM Art: NSM

U20 Math 535 Statistical Learning: An Introduction to Data Mining
This course is an introduction to applications of statistical learning to big data sets. Topics include assessing model accuracy, linear v. logistic regression, cross validation and resampling, shrinkage and regularization (lasso) methods, decision trees and other tree-based methods, and clustering methods such as K-means, hierarchical clustering, and support vector machines. We also cover data mining for massive data sets, such as association rule mining. Linear regression will be reviewed. The course provides skills and experience for careers in statistical and machine learning, and for positions such as data scientist, data analyst, applied statistician, and data-savvy manager. Prerequisites: U20 Math 594 Mathematical Statistics or permission of instructor, and introductory-level programming (R, SAS, or Python).
Credit 3 units.

U20 Math 539 Linear Statistical Models
Theory and practice of linear regression, analysis of variance (ANOVA) and their extensions, including testing, estimation, confidence interval procedures, modeling, regression diagnostics and plots, polynomial regression, collinearity and confounding, model selection, geometry of least squares, etc. The theory will be approached mainly from the frequentist perspective and use of the computer (mostly R) to analyze data will be emphasized. Prerequisites: CSE 131 or 200, Math 3200 and a course in linear algebra (such as Math 309 or 429), or permission of instructor.
Credit 3 units. Art: NSM

U20 Math 5392 Advanced Linear Statistical Models
Review of basic linear models relevant for the course; generalized linear models including logistic and Poisson regression (heterogeneous variance structure, quasilikelihood); linear mixed-effects models (estimation of variance components, maximum likelihood estimation, restricted maximum likelihood, generalized estimating equations); generalized linear mixed-effects models for discrete data, models for longitudinal data, optional multivariate models as time permits. The computer software R will be used for examples and homework problems. Implementation in SAS will be mentioned for several specialized models. Prerequisites: Math 439 and a course in linear algebra (such as Math 309 or 429), or consent of instructor.
Same as L24 Math 4392
Credit 3 units. A&S IQ: NSM Arch: NSM Art: NSM
U20 Math 549 Numerical Applied Mathematics
Computer arithmetic, error propagation, condition number and stability; mathematical modeling, approximation and convergence; roots of functions; calculus of finite differences; implicit and explicit methods for initial value and boundary value problems; numerical integration; numerical solution of linear systems, matrix equations, and eigensystems; Fourier transforms; optimization. Various software packages may be introduced and used. Prerequisites: CSE 200 or 131 (or other computer background with permission of the instructor); Math 217 and 309.
Same as L24 Math 449
Credit 3 units.

U20 Math 560 Multivariate Statistical Analysis
A modern course in multivariate statistics. Elements of classical multivariate analysis as needed, including multivariate normal and Wishart distributions. Clustering; principal component analysis. Model selection and evaluation; prediction error; variable selection; stepwise regression; regularized regression. Cross-validation. Classification; linear discriminant analysis. Tree-based methods. Time permitting, optional topics may include nonparametric density estimation, multivariate regression, support vector machines, and random forests. Prerequisite: Multivariable calculus (Math 233), linear or matrix algebra (Math 429 or Math 309), multivariable-calculus-based probability and mathematical statistics (Math 493, Math 494) and linear models (Math 439). Prior knowledge of R at the level introduced in Math 439 is assumed.
Same as L24 Math 460
Credit 3 units.

U20 Math 561 Time Series Analysis
A time series is a set of data that consists of measurements taken repeatedly along a time scale. Examples include hourly temperature readings, the price of a stock over the course of a day, and monthly counts of numbers of airline passengers. Such data includes dependence and requires a different approach than those covered in linear models courses. This course will introduce methods used to analyze such complicated data. The emphasis will be on real-world applications and data analysis using statistical software. Prerequisites: Math 593 and Math 594; or permission of the instructor. Programming will be done with R. Students should have completed Math 124 or equivalent experience.
Credit 3 units.

U20 Math 593 Probability
Mathematical theory and application of probability at the advanced undergraduate level; a calculus based introduction to probability theory. Topics include the computational basics of probability theory, combinatorial methods, conditional probability including Bayes’ theorem, random variables and distributions, expectations and moments, the classical distributions, and the central limit theorem.
Credit 3 units.

U20 Math 594 Mathematical Statistics
Theory of estimation, minimum variance and unbiased estimators, maximum likelihood theory, Bayesian estimation, prior and posterior distributions, confidence intervals for general estimators, standard estimators and distributions such as the Student-t and F-distribution from a more advanced viewpoint, hypothesis testing, the Neymann-Pearson Lemma (about best possible tests), linear models, and other topics as time permits.
Credit 3 units.

U20 Math 596 Thesis Practicum II
Credit 3 units.

U20 Math 584 Multilevel Models in Quantitative Research
This course covers statistical model development with explicitly defined hierarchies. Such multilevel specifications allow researchers to account for different structures in the data and provide for the modeling of variation between defined groups. The course begins with simple nested linear models and proceeds on to non-nested models, multilevel models with dichotomous outcomes, and multilevel generalized linear models. In each case, a Bayesian perspective on inference and computation is featured. The focus on the course will be practical steps for specifying, fitting, and checking multilevel models with much time spent on the details of computation in the R and Bugs environments. Prerequisite: Math 2200, Math 3200, Poli Sci 581, or equivalent.
Same as L32 Pol Sci 584
Credit 3 units.

U20 Math 575 Statistical Computation
Introduction to modern computational statistics. Pseudo-random number generators; inverse transform and rejection sampling. Monte Carlo approximation. Nonparametric bootstrap procedures for bias and variance estimation; bootstrap confidence intervals. Markov chain Monte Carlo methods; Gibbs and Metropolis-Hasting sampling; tuning and convergence diagnostics. Cross-validation. Time permitting, optional topics include numerical analysis in R, density estimation, permutation tests, subsampling, and graphical models. Prior knowledge of R at the level used in Math 494 is required. Prerequisite: Math 233, 309, 493, 494 (not concurrently); acquaintance with fundamentals of programming in R.
Same as L24 Math 475
Credit 3 units.

U20 Math 596 Thesis Practicum II
Credit 3 units.