BIOSTATISTICS (BSTA)

BSTA 5110 Biostatistics in Practice
Biostatistics in Practice offers Biostatistics students an opportunity to acquire and demonstrate proficiency in statistical collaboration, data analysis and scientific writing. The project is defined by several elements: A scientific question or hypothesis arising in medical research; the statistical methodology needed to address the question; the development of a study design and/or analysis of a relevant data set; and a summary of the results of these analyses. In most cases, a collaborating medical scientist provides the research question and the data. The student, under the supervision of a biostatistics faculty member, identifies the appropriate statistical methods and conducts the analysis. The analysis should be sufficiently extensive and detailed to support a manuscript publishable in the medical literature. Enrollment open to Biostatistics student only.
Fall
1 Course Unit

BSTA 5500 Applied Regression and Analysis of Variance
An applied graduate level course in multiple regression and analysis of variance for students who have completed an undergraduate course in basic statistical methods. Emphasis is on practical methods of data analysis and their interpretation. Covers model building, general linear hypothesis, residual analysis, leverage and influence, one-way anova, two-way anova, factorial anova. Primarily for doctoral students in the managerial, behavioral, social and health sciences. Permission of instructor required to enroll.
Fall
Also Offered As: PSYC 6110, STAT 5000
1 Course Unit

BSTA 6100 Biostatistical Methods for Epidemiology
This one-semester course is designed to provide a strong foundation in biostatistical methods for epidemiologic research, intended for students entering a PhD program in epidemiology. Covered topics include introductory probability theory, estimands, large-sample theory, hypothesis testing, confidence intervals, linear regression, generalized linear models, models for correlated data, and survival analysis, all with a throughline of likelihood-based inference. The course will consist of interactive lectures and labs designed to develop hands-on analytic skills.
Prerequisite: Single variable calculus, Prior coursework in statistics at the undergraduate level
Fall
Corequisite: EPID 7010
1 Course Unit

BSTA 6200 Probability I
This course covers Elements of (non-measure theoretic) probability necessary for the further study of statistics and biostatistics. Topics include set theory, axioms of probability, counting arguments, conditional probability, random variables and distributions, expectations, generating functions, families of distributions, joint and marginal distributions, hierarchical models, covariance and correlation, random sampling, sampling properties of statistics, modes of convergence, and random number generation. Two semesters of calculus (through multivariate calculus), linerar algebra, or permission of the instructor to enroll.
Fall
1 Course Unit

BSTA 6210 Statistical Inference I
This class will cover the fundamental concepts of statistical inference. Topics include sufficiency, consistency, finding and evaluating point estimators, finding and evaluating interval estimators, hypothesis testing, and asymptotic evaluations for point and interval estimation.
Prerequisite: If course requirements not met, permission of instructor.
Spring
Prerequisite: BSTA 6200
1 Course Unit

BSTA 6220 Statistical Inference II
This This class will cover the fundamental concepts of statistical inference. Topics include sufficiency, consistency, finding and evaluating point estimators, finding and evaluating interval estimators, hypothesis testing, and asymptotic evaluations for point and interval estimation.
Fall
Prerequisite: BSTA 6210
1 Course Unit

BSTA 6300 Statistical Methods and Data Analysis I
This first course in statistical methods for data analysis is aimed at first-year Biostatistics students. It focuses on the analysis of continuous data. Topics include descriptive statistics (measures of central tendency and dispersion, shapes of distributions, graphical representations of distributions, transformations, and testing for goodness of fit); populations and sampling (hypotheses of differences and equivalence, statistical errors); one- and two-sample t tests; analysis of variance; correlation; nonparametric tests on means and correlations; estimation (confidence intervals and robust methods); categorical data analysis (proportions; statistics and test for comparing proportions; test for matched samples; study design); and regression modeling (simple linear regression, multiple regression, model fitting and testing, partial correlation, residuals, multicollinearity). Examples of medical and biologic data will be used throughout the course, and use of computer software demonstrated. Multivariable calculus and linear algebra and permission of instructor required to enroll.
Fall
Prerequisite: BSTA 6200
1 Course Unit

BSTA 6320 Statistical Methods for Categorical and Survival Data
This is the second half of the methods sequence, where the focus shifts to methods for categorical and survival data. Topics in categorical include defining rates; incidence and prevalence; the chi-squared test; Fisher's exact test and its extension; relative risk and odds-ratio; sensitivity; specificity; predictive values; logistic regression with goodness of fit tests; ROC curves; the Mantel-Haenszel test; McNemar's test; the Poisson model; and the Kappa statistic. Survival analysis will include defining the survival curve, censoring, and the hazard function; the Kaplan-Meier estimate, Greenwood's formula and confidence bands; the log rank test; and Cox's proportional hazards regression model. Examples of medical and biologic data will be used throughout the course, and use of computer software demonstrated.
Spring
Prerequisite: BSTA 6200 AND BSTA 6210 AND BSTA 6300
1 Course Unit

BSTA 6510 Linear Models
This course extends the content on linear models in BSTA 630 and BSTA 632 to more advanced concepts and applications of linear models. Topics include the matrix approach to linear models including regression and analysis of variance, general linear hypothesis, estimability, polynomial, piecewise, ridge, and weighted regression, regression and collinearity diagnostics, multiple comparisons, fitting strategies, simple experimental designs (block designs, split plot), random effects models, Best Linear Unbiased Prediction. Applications of methods to example data sets will be emphasized.
Spring
Prerequisite: BSTA 6200 AND BSTA 6300 AND BSTA 6210 AND BSTA 6320
1 Course Unit

BSTA 6560 Longitudinal Data Analysis
This course covers both the applied aspects and methods developments in longitudinal data analysis. In the first part, we review the properties of the multivariate normal distribution and cover basic methods in longitudinal data analysis, such as exploratory data analysis, two-stage analysis and mixed-effects models. Focus is on the linear mixed-effects models, where we cover restricted maximum likelihood estimation, estimation and inference for fixed and random effects and models for serial correlations. We will also cover Bayesian inference for linear mixed-effects models.The second part covers advanced topics, including nonlinear mixed-effects models, GEE, generalized linear mixed-effects models, nonparametric longitudinal models, functional mixed-effects models, and joint modeling of longitudinal data and the dropout mechanism. If course requirements are not met, permission of instructor required.
Fall
Prerequisite: BSTA 6210 AND BSTA 6320 AND BSTA 6510
1 Course Unit

BSTA 6610 Design of Interventional Studies
This course is designed for graduate students in statistics or biostatistics interested in the statistical methodology underlying the design, conduct, and analysis of clinical trials and related interventional studies. General topics include designs for various types of clinical trials (Phase I, II, III), endpoints and control groups, sample size determination, and sequential methods and adaptive design. Regulatory and ethical issues will also be covered. Prerequisite: If course requirement not met, permission of instructor required.
0.5 Course Units

BSTA 6700 Statistical Computing
This course concentrates on computational tools, which are useful for statistical research and for computationally intensive statistics. Through this course you will develop a knowledge base and skill set of a wide range of computational tools needed for statistical research. Topics include computer storage, architecture and arithmetic; random number generation; numerical optimization methods; spline smoothing and penalized likelihood; numerical integration; simulation design; Gibbs sampling; bootstrap methods; and the EM algorithm. Prerequisite: If course requirements not met, permission of instructor required.
Prerequisite: BSTA 6510 AND BSTA 6200 AND BSTA 6210
1 Course Unit

BSTA 6990 Lab Rotation
Student lab rotation.
0-4 Course Units

BSTA 6990A Lab Rotation
Student lab rotation.
0-2 Course Units

BSTA 6990B Lab Rotation
Student lab rotation.
0-2 Course Units

BSTA 7000 Seminal Ideas & Controversies in Statistics
This graduate-level seminar explores the driving forces behind modern statistical thought by engaging directly with foundational papers and debates that have shaped the field. Anchored in Roderick Little's Seminal Ideas and Controversies in Statistics, the course emphasizes conceptual understanding and historical context rather than formal proofs. Three central themes structure the semester: 1. Philosophical Foundations of Inference • Likelihood principle • Bayesian vs. frequentist paradigms • Fisher's fiducial inference • Hypothesis testing debates (Fisher vs. Neyman–Pearson) 2. Methodological Developments • Regression and shrinkage estimation • Multiple comparisons • Robust inference methods 3. Statistical Design & Randomization • Experimental design foundations • Randomization as the basis of inference • Modern perspectives on causal estimation Students will analyze seminal and modern papers, present key controversies, and articulate their relevance to contemporary statistical practice. Recommended Prerequisites: Graduate-level courses in probability, statistical inference, and regression.
Fall, even numbered years only
1 Course Unit

BSTA 7500 Statistical Methods for Risk Prediction and Precision Medicine
This is an advanced elective course for graduate students in Biostatistics, Statistics, Epidemiology, and other BGS disciplines. It will cover various topics for evaluating the performance of biomarkers to predict risk of clinical or disease outcomes, specifically including relative, absolute and competing risks for binary and time-to-disease outcomes; statistical inference for quantifying predictive accuracy with binary and time-to-event outcomes; statistical methods and inference for case-control study designs; Efficient study design issues for biomarker evaluation. This course is designed to help students 1) understand various concepts of risk in the medical literature; 2) understand various statistical methods for evaluating prediction performance of biomarkers and diagnostic tests and for designing efficient biomarker studies; 3) improve the ability to read critically papers published in statistical and medical journals on related topics; and 4) develop research ideas for risk prediction. Upon successfully completing this course, students will be able to: 1) Conduct statistical analysis for evaluating prediction performance of biomarkers and diagnostic tests; 2) Have a better ability to read and understand papers published in statistical and medical journals on related topics; and 3) Be well prepared to work on related topics for dissertation.
Fall
Prerequisite: BSTA 6300 AND BSTA 6320
0.5 Course Units

BSTA 7510 Statistical Methods for Neuroimaging
This course is intended for students interested in both statistical methodology, and the process of developing this methodology, for the field of neuroimaging. This will include quantitative techniques that allow for inference and prediction from ultra-high dimensional and complex images. In this course, basics of imaging neuroscience and preprocessing will be covered to provide students with requisite knowledge to develop the next generation of statistical approaches for imaging studies. High-performance computational neuroscience tools and approaches for voxel- and region-level analyses will be studied. The multiple testing problem will be discussed, and the state-of-the art in the area will be examined. Finally, the course will end with a detailed study of multivariate pattern analysis, which aims to harness patterns in images to identify disease effects and provide sensitive and specific biomarkers. The student will be evaluated based on 3 homework assignments and a final in-class presentation. Prerequisite: If course requirement not met, permission of instructor required.
Prerequisite: BSTA 6210 AND BSTA 6510
1 Course Unit

BSTA 7540 Advanced Survival Analysis
This advanced survival analysis course will cover statistical theory in counting processes, large sample theory using martingales, and other state of the art theoretical concepts useful in modern survival analysis research. Examples in deriving rank-based tests and Cox regression models as well as their asymptotic properties will be demonstrated using these theoretical concepts. Additional potential topics may include competing risk, recurrent event analysis, multivariate failure time analysis, joint modeling of survival and longitudinal data, sample size calculations, multi-state models, and complex sampling schemes involving failure time data. In addition to satisfying course prerequisites, permission of instructor is required.
Fall
Prerequisite: BSTA 6220
1 Course Unit

BSTA 7770 Statistical Methods for Meta-Analyses
This graduate-level Biostatistics course will introduce the fundamentals of statistical methods for meta-analyses. It will cover key principles of meta-analysis and the statistical rationales behind the analytic models, including univariate meta-analysis, multivariate meta-analysis, meta-analysis of diagnostic test accuracy, network meta-analysis, and multivariate network meta-analysis. Beyond these commonly used models, the course will cover statistical methods and software that investigate and correct for biases in systematic reviews such as publication bias, outcome reporting bias. Advanced statistical inferential tools such as publication bias, outcome reporting bias. Advanced statistical inferential tools such as composite likelihood, pseudolikelihood, integrated likelihood methods, EM algorithms will be introduced. In addition, the course will also cover some practical steps in systematic review including search strategies, data abstraction methods; quality assesment; and writing a meta-analysis report. The course is composed of a series of weekly lectures and small group discussions. Students will be expected to attend weekly lectures, participate in class discussions, review assigned readings, complete homework assignments, and conduct a real-world meta-analysis with a clinically meaningful problem. Fundamentals of Biostatistics background or permission of instructor required to enroll.
1 Course Unit

BSTA 7800 The Science of Science and Innovation
""I fully agree with you about the significance and educational value of history and philosophy of science. So many people today - even professional scientists - seem to me like somebody who has seen thousands of trees but has never seen a forest."" "— Albert Einstein
The increasing burden of knowledge in biomedical science has led training and coursework to focus on the many trees within a specific area of research. While understandable, this narrowed scope means that scientists themselves are often unaware of historical, economic, and social forces that structure the enterprise in which they work. This course aims to illuminate these dynamics. Tapping into the many emerging metasciences—the science of science, economics of science, philosophy of science, etc—we will embark on a slow zoom in from a 1000 foot view, moving gradually from the perspective of governments, to funders, to practitioners, to trainees.
Spring
1 Course Unit

BSTA 7820 Stat Meth/Incomplet Data
Selected topics from public health and biomedical research where "Big data" are being collected and methods are being developed and applied, together with some core statistical methods in high dimensional data analysis. Topics include dimension reduction, detection of novel association in large datasets, regularization and high dimensional regression, ensemble learning and prediction, kernel methods, deep learning and network analysis. R programs will be used throughout the course, other standalone programs will also be used. Prerequisite: If course requirement not met, permission of instructor required.
Spring
1 Course Unit

BSTA 7870 Methods for Statistical Genetics and Genomics in Complex Human Disease
This is an advanced elective course for graduate students in Biostatistics, Statistics, Epidemiology, Bioinformatics, Computational Biology, and other BGS disciplines. This course will cover statistical methods for the analysis of genetics and genomics data. Topics covered will include genetic linkage and association analysis, analysis of next-generation sequencing data, including those generated from DNA sequencing and RNA sequencing experiments. Students will be exposed to the latest statistical methodology and computer tools on genetic and genomic data analysis. They will also read and evaluate current statistical genetics and genomics literature. Prerequisite: If course requirements not met, permission of instructor required.
Prerequisite: (BSTA 6300 AND BSTA 6320) OR (EPID 5200 AND EPID 5210)
1 Course Unit

BSTA 7890 Big Data
Selected topics from public health and biomedical research where "Big data" are being collected and methods are being developed and applied, together with some core statistical methods in high dimensional data analysis. Topics include dimension reduction, detection of novel association in large datasets, regularization and high dimensional regression, ensemble learning and prediction, kernel methods, deep learning and network analysis. R programs will be used throughout the course, other standalone programs will also be used. Prerequisite: If course requirement not met, permission of instructor required.
Prerequisite: BSTA 6210 AND BSTA 6220
1 Course Unit

BSTA 7900 Causal Inference in Biomedical Research
This course considers approaches to defining and estimating causal effects in various settings. The potential-outcomes approach provides the framework for the concepts of causality developed here, although we will briefly consider alternatives. Topics considered include: the definition of effects of scalar or point treatments; nonparametric bounds on effects; identifying assumptions and estimation in simple randomized trials and observational studies; alternative methods of inference and controlling confounding; propensity scores; sensitivity analysis for unmeasured confounding; graphical models; instrumental variables estimation; joint effects of multiple treatments; direct and indirect effects; intermediate variables and effect modification; randomized trials with simple noncompliance; principal stratification; effects of time-varying treatments; time-varying confounding in observational studies and randomized trials; nonparametric inference for joint effects of treatments; marginal structural models; and structural nested models. Prerequisite: If course requirement not met, permission of instructor required.
Not Offered Every Year
Prerequisite: BSTA 6200 AND BSTA 6210 AND BSTA 6220 AND BSTA 6300 AND BSTA 6310
1 Course Unit

BSTA 7980 Advanced Topics in Biostatistics
This course is designed for second-year PhD students in Biostatistics. The goal is to provide an in-depth exploration of special topics within the field of biostatistics. The course covers a range of advanced statistical methods and their applications in various biostatistical domains, including clinical trials, causal inference, survival analysis, genetics and genomics, neuroimaging, and health informatics. The course emphasizes the unique aspects of these topics and their significance in biomedical research and public health. Throughout the course, ten faculty members will deliver presentations, each focusing on a special topic.
Fall
Prerequisite: BSTA 6200 AND BSTA 6210 AND BSTA 6300 AND BSTA 6510 AND BSTA 6320
0.5 Course Units

BSTA 8990 Pre-Dissertation Research
0.5-3 Course Units

BSTA 9200 Guided Dissertation Research
0.5-3 Course Units

BSTA 9900 Master's Thesis
Students in the PhD and MS programs will develop and conduct advanced research under the supervision of a biostatistics faculty mentor before completing an MS thesis.
Fall
1-3 Course Units

BSTA 9950 Dissertation
Ph.D. students enroll in this course after passing their candidacy exam. They work on their dissertation full-time under the guidance of their dissertation supervisor and other members of their dissertation committee.
Fall or Spring
3 Course Units