DEPARTMENT OF GLOBAL BIOSTATISTICS AND DATA SCIENCE

Programs
Graduate Certificates

Graduate

Courses
Biostatistics (BIOS)
BIOS 6030 Introductory Biostat (3 Credit Hours)
Introduction to statistical methodology in the health field. Topics include presentation of data (graphs and tables), descriptive statistics, concepts of probability, estimation of parameters, hypothesis testing, simple linear regression, correlation, and the analysis of attribute data. It is recommended for students with any mathematical or statistical background and those needing a firm foundation in statistical methods either for their careers or preparation for further quantitative courses.

BIOS 6040 Intermediate Biostat (3 Credit Hours)
This is an intermediate course in applied biostatistics. The course covers Analysis of Variance and Multiple Regression and Correlation Analysis, and Logistic Regression. The focus will be on numerical computation and interpretation of results of statistical application using statistical packages. Elementary knowledge of the use of statistical computing packages is needed.

Prerequisite(s): (BIOS 6030 or SPHL 6050).

BIOS 6220 Database Management (3 Credit Hours)
An introduction to the principles and application of data management, techniques in data collection, data cleaning, data reporting, database design, and implementing databases for managing large data systems. After taking the course, students will be able to create databases with applications to public health intervention and surveillance, use SQL to administrate, manage, and retrieve data for statistical analysis. Prerequisite(s): Basic knowledge of MS Office.

BIOS 6240 Computer Packages-Spss (1 Credit Hour)
Prerequisite(s): BIOS 6030.
* May be taken concurrently.

BIOS 6300 Introduction To Arcgis (1 Credit Hour)
This course covers the elementary concepts and applications for mapping using the ArcGIS software. The course focuses on a wide variety of public health applications and is applicable to virtually all academic and professional settings where mapping is used. Each lecture begins with a PowerPoint presentation to introduce fundamental mapping concepts and is followed with in-class exercises to reinforce hands-on application. Two in-class, paper-based exams are given to monitor and assess students' understanding of the course concepts.

Prerequisite(s): (BIOS 6030 or SPHL 6050).
* May be taken concurrently.

BIOS 6800 Intro To Pub Hlthgis (3 Credit Hours)
The course is an introduction to desktop mapping and spatial analysis. The first part of the course covers geographic information systems (GIS) concepts and mapping using the ArcGIS software. The second part of the course covers introductory spatial analytical techniques, including spatial autocorrelation quantification, cluster analysis, and spatial modeling. The student will develop a public health GIS project that requires the synthesis of mapping and spatial analysis.

BIOS 7040 Statistical Inference I (3 Credit Hours)
The course is the first of a sequence in the theory of statistical interference and probability. The first part of the course covers probability theory; discrete, continuous, and exponential distribution functions; moment generating functions; and differentiation. The latter part of the course covers joint and marginal distributions and concepts of random samples. Students taking this course need to have completed at least one year of college calculus. Students will develop a project that synthesizes the course learning objectives through an applied course project. The course focuses on the theoretical underpinnings of biostatistics and improving understanding of statistical application and problem solving approaches.
BIOS 7050  Statistical Inference II  (3 Credit Hours)
The course is the second part of a sequence for introduction to statistical inference and probability. The first part of the course covers data reduction, point estimation, hypothesis testing, and interval estimation. The latter part of the course covers asymptotic evaluations, analysis of variance, and regression modes. The student will develop a project that synthesizes the course learning objectives through an applied course project. The course focuses on the theoretical underpinnings of biostatistics and improving understanding of statistical application and problem solving approaches.

**Prerequisite(s):** BIOS 7040.

BIOS 7060  Regression Analysis  (3 Credit Hours)
This is an advanced course on selected statistical techniques for analyzing data on multiple variables, both continuous and categorical. This course ultimately provides the student with insight into the application of regression techniques to the medical and health sciences. It focuses on statistical methodology with emphasis on selection of appropriate applications and interpretation of results. Elementary knowledge of the use of statistical computing package is needed.

**Prerequisite(s):** (BIOS 6030 or SPHL 6050) and BIOS 6040.

BIOS 7080  Design of Experiments  (3 Credit Hours)
This course deals with fundamental topics in design of experiments including principle theory of experimental designs (randomization, replication, and balance). It focuses the main elements of statistical thinking in the context of experimental design such as completely randomized design, randomized complete block design, experiments with two factors, factorial design, Latin Square, nested designs, repeated measurement design, and split-plot designs. Elementary knowledge of the use of statistical computing packages is needed.

**Prerequisite(s):** (BIOS 6030 or SPHL 6050) and BIOS 6040.

BIOS 7150  Categorical Data Analysis  (3 Credit Hours)
Fundamental concepts and methods for analysis of categorical outcomes. Topics include analysis of 2-way tables, unconditional and conditional logistic regression, power and sample size computation, and modeling of dependent categorical outcomes via mixed models and GEE methods. Course covers the mathematical basis of the statistical procedures but the emphasis is on application of the methods using statistical software and interpretation of results. Elementary knowledge of the use of statistical computing packages is needed.

**Prerequisite(s):** (BIOS 6030 or SPHL 6050) and BIOS 6040.

BIOS 7220  Nonparametric Statistics  (3 Credit Hours)
Nonparametric inferential statistical methods are introduced. Topics include single, paired, independent, and multiple sample hypothesis testing and confidence interval methods; non parametric regression and correlation methods; categorical data and measures of concordance. Elementary knowledge of the use of statistical computing packages is needed.

**Prerequisite(s):** (BIOS 6030 or SPHL 6050) and BIOS 6040.

BIOS 7250  Principles of Sampling  (3 Credit Hours)
This course introduces core principles of survey sampling, with emphasis on sampling plans, methods of estimating unknown parameters of population and subdomain, and techniques for calculating precisions of the estimators. Topics include: basic concepts in survey sampling, simple random sampling; stratified random sampling; systematic sampling; one-, two-, and multi-stage cluster sampling; probability proportionate to size sampling. Elementary knowledge of the use of statistical computing packages is needed.

**Prerequisite(s):** (BIOS 6030 or SPHL 6050).

BIOS 7270  Asymptotic Inference  (3 Credit Hours)
This course provides an introduction to the fundamental tools and concepts of asymptotic statistics, or the large sample theory, without requiring measure theory. Students who take this course should have the general knowledge of statistical inference and working knowledge of statistical software such as SAS or R. The course includes and introduction on the limit theory of random variables, classic likelihood-based asymptotic theory and the asymptotic theory of nonparametric statistics, including the theory of U-statistics and smoothing methods. It also covers resampling methods, including permutation tests, bootstrap, and Jackknife.

**Prerequisite(s):** MATH 6070 and 6080.

BIOS 7300  Survival Data Analysis  (3 Credit Hours)
Topics include analysis of survivorship data including estimation and comparison of survival curves, regression methods in the analysis of prognostic and etiologic factors, concepts of competing risks, and the analysis of clinical trial data. Software used for problem solving. Emphasis placed on the application of methods to the analysis of public health data with examples of clinical trials, cancer survivorship, and other data sets for which there is partial follow-up of subjects. Elementary knowledge of the use of statistical computing packages is needed.

**Prerequisite(s):** (BIOS 6030 or SPHL 6500) and BIOS 6040.
BIOS 7380 Bayesian Inference (3 Credit Hours)
This course examines theoretical foundations and applications of Bayesian paradigm, including Bayes’ theorem, prior distribution, likelihood function, deriving posterior distributions, and point and interval estimations. A variety of topics are covered, which encompass Bayesian inference for single- and multi-parameter models, linear regression, hierarchical models, and commonly used Gibbs sampler and Metropolis-Hastings algorithm. Assessment of convergence, the evaluation of models, and the presentation of the results are also illustrated. Real world examples drawn from medical research are used to show practicality of Bayesian approach, particularly how to update beliefs and make inferences from observed data. Elementary knowledge of the use of statistical computing packages is needed.

Prerequisite(s): (BIOS 6030 or SPHL 6050) and BIOS 6040.

BIOS 7400 Clinical Trials (3 Credit Hours)
Covers design, implementation, analysis and reporting of clinical trials. Topics encompass trial design, hypothesis formulation and testing, methods of randomization, ethics, sequential trials, sample size determination, blinding, subject recruitment, data collection and management, quality control, monitoring outcomes and adverse events, interim analysis, statistical methods in analyzing trial data, and addressing scientific issues in reporting and interpreting trial results. Elementary knowledge of the use of statistical computing packages is needed.

Prerequisite(s): (BIOS 6030* or SPHL 6050*) and BIOS 6040.

* May be taken concurrently.

BIOS 7550 Epigenetics and Epigenomics (3 Credit Hours)
This course exposes students to the underlying biological basis behind current research in the area of epigenetics and epigenomics. Students will review the principles and recent progresses in epigenetic regulation of physiology and patho-physiology, along with current and emerging techniques and methodologies for assessing epigenomic features. Students will present and discuss recent cutting-edge and epigenetic and epigenomic studies.

Prerequisite(s): TRMD 6010*.

* May be taken concurrently.

BIOS 7650 Stat Learning in Data Science (3 Credit Hours)
This course provides detailed overviews over the evaluation and application of statistical learning theories and techniques for inference and prediction in data science, particularly for biological and public health data. Topics include linear and nonlinear models, resampling techniques, tree-based methods, unsupervised learning such as clustering, support vector machine, graphical models, etc. Working on real and/or simulated data through assignments, students will apply the knowledge learned and practice their skills in solving various biological and public health problems, such as sequence alignment, gene prediction, subtype identification and classification, and disease risk and prognosis prediction. Discussion on model assessment and selection are also included. Elementary knowledge of the use of statistical computing packages is needed.

Prerequisite(s): (BIOS 6030 or SPHL 6050) and BIOS 6040.

BIOS 7990 Masters Independent Studies (1-3 Credit Hours)
Masters students and advisor select a topic for independent study and develop learning objectives and the expected written final product.

BIOS 8000 Doctoral Journal Club (0 Credit Hours)
This course is intended to improve students’ ability in interpreting, evaluating, critiquing, presenting, and communicating the elements, concepts, findings, and implications from current Biostatistics and Bioinformatics research literatures in a seminar setting. All enrolled students will be expected to give at least one oral presentation and participate in the student-led discussions. Feedback to each presenter will be given orally, in writing and/or through e-mails by faculty and peer students. At the end of the course, students will gain experience in assessing the value of research findings from selected publications to biostatistics and bioinformatics research.

BIOS 8200 Causal Inference for Biomedical Informatics (3 Credit Hours)
This course covers basic concepts and selected state-of-the-art statistical methods and theory of causal inference for biomedical informatics. It will empower students to draw causal conclusions and make predictions by mining data from observational and experimental studies. Topics include: targeted machine learning, structural equation modeling, Mendelian randomization, and heteroskedastic genomic prediction. Elementary knowledge of the use of statistical computing packages is needed.

Prerequisite(s): (BIOS 6030 or SPHL 6050) and BIOS 6040, 7060 and MATH 6080.

BIOS 8270 Asymptotic Inference (3 Credit Hours)
Prerequisite(s): BIOS 7040* and 7050*.

* May be taken concurrently.

BIOS 8350 Clustered & Longitudinal Data Analysis (3 Credit Hours)
This is an advanced course in analysis of clustered and longitudinal data, with or without missing values. Students will compute power and sample size for clustered and longitudinal data using generalized linear mixed effect models and estimating equations. Class discussion, lecture, and assignments emphasize application of methods to the analysis of public health data with examples of clinical trials and epidemiological observational studies. Use of standard statistical software and methods required. Elementary knowledge of the use of statistical computing packages is needed.

Prerequisite(s): (BIOS 6030 or SPHL 6050) and (BIOS 6040 or 7060).
BIOS 8500 Monte Carlo and Bootstrapping (3 Credit Hours)
This hands-on course introduces the methods used for Monte Carlo simulations and nonparametric bootstrapping. Students learn how to design, program, and interpret a simulation study, uses of bootstrapping for estimation and inference, jackknifing, and other resampling methods. Monte Carlo Markov Chain methods and Bayesian inference in Monte Carlo methods will be introduced. This is an advanced, computer-intensive course, so knowledge of programming language (SAS or R preferred) as well as ability to work independently are required.

Prerequisite(s): (BIOS 7060, 7080, 7150, 7220 or 7300).

BIOS 8800 Applied Data Analysis (3 Credit Hours)
This is an advanced methods course for hands-on data analysis and management. Students use real datasets to formulate testable hypotheses, interrogate and clean data, implement appropriate strategies for missing data, design and perform appropriate analyses, and keep written documentation of their analyses. Students also learn how to interpret and effectively report the results of statistical analyses, both orally and in writing. Use of a statistical software package is required.

BIOS 8820 Multivariate Methods (3 Credit Hours)
This is a doctorate level course that covers techniques used to conduct analysis with more than one outcome variable. The focus will be on association methods and predictive models between multiple independent and multiple dependent variables. Additionally, the students will learn techniques for variable reduction, path models, and factor analysis. Students will conduct numerical computation and interpretation of results of statistical application using statistical packages.

BIOS 8990 Doctoral Independent Study (1-3 Credit Hours)
Doctoral students and advisors select a topic for independent study and develop learning objectives and the expected final written product.

BIOS 9970 Dissertation (0 Credit Hours)
Doctoral candidates who have defended their prospectus and are engaged in research.

BIOS 9980 Master's Thesis Research (0 Credit Hours)
MS Students engaging in thesis research.

BIOS 9990 Dissertation Research (2 Credit Hours)
Doctoral students who have completed course work but not defended their prospectus.