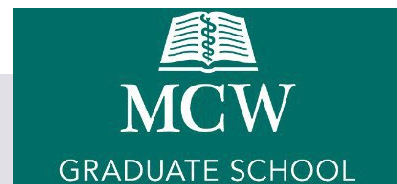


2023-24

BIostatistics & Data Science

Degree Offered: Master of Arts



Program Description

The Master of Arts program in Biostatistics and Data Science provides a learning experience focused on solid theoretical foundation and practical experience. Robust course offerings, active engagement in statistical consulting, and a capstone project create ample opportunities to develop essential analytical skills. Consulting projects ranging from the simplest statistical summaries to the most complex protocols and data collection schemes allow students to get experience of working with real data analysis projects from start to finish. This hands-on experience will enable students to synthesize the acquired knowledge and integrate various courses they have taken. In the process, students will create a portfolio which demonstrates competency in data analysis, statistical programming, consulting experience with non-statisticians, oral and written communication skills.

Admission Requirements

In addition to the general [Graduate School admission requirements](#), this program has specific requirements:

- Any graduate of an accredited college or university with an undergraduate degree in mathematics, statistics, or related field with strong preparation in mathematics is eligible for admission.
- Prior coursework in calculus (including integrals, such as Calculus II), probability and/or statistics, linear/matrix algebra, and computer programming experience.

Credits Required to Graduate

31 credits

Program Credit Requirements

The curriculum consists of eight required biostatistics courses which have been identified as an essential knowledge base for all students in the program. Also required, is an Ethics and Integrity in Science course. The capstone project course can be taken throughout multiple semesters but at least 3 credit hours are required for graduation. The program allows for students to choose two or more elective courses which best reflect their personal interests. Students may pursue the degree on a full-time or part-time basis.

Required Courses

BIOE 10222 Ethics and Integrity in Science. *1 credit.*

This course provides the basis for understanding the ethical issues related to basic scientific and medical research, including animal and human subject research, fraud, and misconduct, and governmental, institutional, and researcher responsibilities. Bioethics 10222 is offered during the spring and summer terms only.

BIOS 04221 Biomedical Applications and Consulting. 3 credits.

Prerequisites: 04231/04232 Statistical Models and Methods I & II

Theory of consulting, communication and statistical techniques most often used in consulting and biomedical applications, practical experience in the real consulting setting and writing statistical reports.

BIOS 04224 Biostatistical Computing. 3 credits.

Prerequisites: 04231 Statistical Models and Methods I or concurrent registration

This course will cover the details of manipulating and transforming data required for statistical analysis. Topics include reshaping the data from a per-case to a per-event within a case and vice-versa. It will also cover the techniques necessary to write functions and macros in both SAS and R for developing new/modified data analysis methods. How to use R packages and C/C++ codes in R will also be covered. The LaTeX document production system is also introduced.

BIOS 04231 Statistical Models and Methods I. 3 credits.

Prerequisite: Three semesters of calculus and one semester of linear algebra

This course will cover statistical techniques for basic statistics. Topics include one-sample/two-sample tests, analyses for count data and contingency tables, basic nonparametric methods including sign, rank-sum and signed-rank tests, simple linear regression model and inference, checking model assumptions, model diagnostics, correlation analysis, one-way analysis of variance, Kruskal-Wallis one-way ANOVA, simple logistic regression, and weighted linear regression. SAS/R will be used throughout the course.

BIOS 04232 Statistical Models and Methods II. 3 credits.

Prerequisite: 04231 Statistical Models and Methods I

Factorial, nested, split-plot and repeated measures designs, multiple regression and variable selection, multiple comparisons, logistic regression, discriminant analysis, principal components and factor analysis, rates and proportions, introduction to survival analysis.

BIOS 04233 Introduction to Statistical and Machine Learning. 3 credits.

Prerequisite: 04232 Statistical Models and Methods II

This course will provide an introduction to statistical learning. Core topics include variable selection, penalized linear regression such as lasso, dimension reduction including principal component analysis, flexible regression techniques including kernel smoothing/smoothing splines/generalized additive models/regression trees, support vector machine, clustering, and random forests. Other topics that can be covered include but are not limited to ridge regression, group lasso, fused lasso, adaptive lasso, SCAD, Bayesian lasso, Bayesian group lasso, Bayesian CART, BART, neural network, feature screening, graphical models, and quantile regression.

BIOS 24150 Bioinformatics in Omics Analysis. 3 credits.

Prerequisites: 04231 Statistical Models and Methods I and Biostatistical Computing, or consent of instructor

The course aims to introduce modern statistical and computational methods in high-throughput omics data analysis. The first half of the course focuses on fundamental statistical and computational methods applicable in different types of high-throughput omics data. The second half covers selected important topics in bioinformatics and aims to give students a systematic view of the omics data analysis. The goals of the course include: (1) to motivate students from quantitative fields into omics research (2) to familiarize students from biological

fields with a deeper understanding of statistical methods (3) to promote inter-disciplinary collaboration atmosphere in class. Students are required to have a basic statistical training (i.e., elementary statistics courses, basic calculus, and linear algebra) and basic programming proficiency (R programming is required for homework and the final project).

BIOS 24160 Concepts in Probability and Statistics. 3 credits.

Prerequisites: Calculus I and II

The course is designed for graduate students who have a background in statistics but would benefit from a review of the basic concepts in probability and statistics. It focuses on the properties of random variables including distributions, expectations, and variability measures. Topics in inferential statistics covered in this course include estimation, hypotheses testing methods for categorical data tabulation and analysis. It also includes an overview of statistical techniques based on simulations and resampling. Key features of Bayesian analysis will be covered as well. After completion of the course, students should be well prepared for taking more advanced courses in statistics, both theoretical and applied.

BIOS 24297 Capstone Project. 1-3 credits.

Prerequisites: 04232 Statistical Models and Methods II

The course is the culmination of the MA program in Biostatistics. Students will complete a project integrating their statistical analysis, data science, and application domain knowledge. The project results in a written report and presentation which will improve students' ability to communicate effectively about statistics and data science in written and oral form using both technical and nontechnical language. In addition, the project will enable students to expand their professional portfolio of coding samples, written reports, and presentations.

Elective Courses

BIOS 04214 Design and Analysis of Clinical Trials. 3 credits.

Prerequisites: 04231 Statistical Models and Methods I or concurrent registration

This course covers issues in clinical trials including the clinical trial protocol, sources of bias in clinical trials, blinding, randomization, sample size calculation; phase I, phase II, phase III and hybrid trials; interim analysis, stochastic curtailment, Bayesian designs, and administrative issues in study design.

BIOS 04222 Statistical Consulting. 3 credits.

Prerequisites: 04231/04232 Statistical Models and Methods I & II

This course is designed for students to gain experience in statistical consulting by working with the biostatistics faculty members on various consulting projects.

BIOS 04275 Applied Survival Analysis. 3 credits.

Prerequisites: 04231 Statistical Models and Methods I

The following topics will be covered in this course: Basic parameters in survival studies; Censoring and truncation, competing risks; Univariate estimation including the Kaplan-Meier and Nelson-Aalen estimator; tests comparing two or more populations, the log rank test; Semi-parametric regression, the Cox model; Aalen's Additive hazards regression model; regression diagnostics.

BIOS 04285 Introduction to Bayesian Analysis. 3 credits.

Prerequisites: 04231 Statistical Models and Methods I

This course introduces basic concepts and computational tools for Bayesian statistical methods. Topics covered include one and two sample inference, regression models and comparison of several populations with normal, dichotomous and count data.

PUBH 18201 Principles of Public Health Data and Epidemiology. 3 credits.

The Principles of Public Health Data and Epidemiology course examines the basic epidemiological concepts required by a health professional, including disease transmission and prevention; morbidity, vital statistics, and mortality; screening tests; the natural history of disease; clinical trials; cohort study designs; case control and cross-sectional study designs; measures and risks in populations; disease causation; and epidemiology as it relates to program evaluation and public policy.

Principles of Public Health Data and Epidemiology provides the student with an understanding of the distribution and determinants of health and disease in population groups. Epidemiology is considered the basic science of public health, and, as such, provides the foundation for many other courses in the MPH program; it is particularly germane to Community Health Assessment and Environmental Health.

PUBH 18258 Advanced Epidemiological Methods. 3 credits.

Epidemiologic Methods builds on introductory epidemiology courses by providing a more in depth understanding of fundamental epidemiologic principles presented in introductory epidemiologic courses such as study design and bias. In addition, Epidemiologic Methods emphasizes more advanced concepts needed in establishing causal relationships from observational data. It is particularly relevant to students who intend to conduct studies investigating the occurrence and determinants of diseases or who wish to be sophisticated consumers or critics of epidemiologic research conducted by others. The course emphasizes practical application of Epidemiologic Methods to real world problems.

PUCH 19150 Introduction to Epidemiology. 3 credits.

The course provides: 1) an overview of epidemiologic concepts; 2) an introduction to the approaches and techniques that are used to measure and monitor health status in populations; 3) an introduction to study designs to assess disease prevention and intervention; and 4) an introduction to clinical research study designs that elucidate causative factors for disease.

PUCH 19210 Health and Medical Geography. 3 credits.

Geography and physical and social environments have important implications for human health and health care. This course will explore the intersections among geography, environments, and public health, with an emphasis on geographical analysis approaches for health data, to address two key questions: (1) How can concepts from geography help us to better understand health and well-being? (2) How can geographic tools, such as Geographic Information Systems (GIS) be used to address pressing questions in health and medical research? Students will become acquainted with theories and methods from health and medical geography through readings, discussion, Geographic Information Systems (GIS) laboratory exercises, and the completion of a focused course project. Throughout the semester we will use the concepts and techniques of the discipline of geography to

investigate a variety of health-related topics, and laboratory exercises will center on common health and medical geography research questions. Course projects will allow students to develop a deep understanding of the geographical nature of a health problem of their choosing and will incorporate both literature review and the analysis of geographical data.

PUCH 19229 Survey Research Methods. 3 credits.

Survey Research Methods is a graduate-level, 3-credit hour course that introduces students to the broad concepts of survey design, conduct, and analysis. Students will gain a detailed and comprehensive understanding of questionnaire design, sampling, data collection, survey nonresponse, and analysis of survey data. The course will include lectures, reading assignments, class discussions, individual and group presentations, and exams.

CTSI 20151 Introduction to Epidemiology. 3 credits.

This course provides an introduction to the concepts, principles, and research methods specific to epidemiology. Students will learn about population health, how to select appropriate study designs for collecting evidence for medical practice, how to summarize evidence for medical practice and how to translate evidence into medical practice. By the end of the course, students should be able to apply the skills learned to assess the health of a population, describe determinants of health, and select an appropriate study design to evaluate population health.

Contact information: gradschool@mcw.edu | (414) 955-8218