2024-25

BIOSTATISTICS

Degree Offered: Doctor of Philosophy

MCW GRADUATE SCHOOL

Program Description

The PhD program in Biostatistics is designed for students having strong quantitative and computing skills with interest in applying cutting-edge biostatistical research techniques to biological and medical sciences. In this program, students will receive in-depth training on theoretical/methodological/collaborative research in biostatistics and the use of state-of-the-art software. The Division of Biostatistics is a highly collaborative unit and provides lots of opportunities to participate in numerous research projects within MCW and its affiliates.

Admission Requirements

In addition to the general Graduate School admission requirements, this program has additional specific requirements.

Any graduate of an accredited college or university with an undergraduate degree in mathematics or closely related fields with strong preparation in mathematics is eligible for admission. Applicants are expected to have completed and performed well in courses in advanced calculus, linear/matrix algebra, and scientific programming. Those who haven't done so may be considered for admission to the program upon approval of the biostatistics admission committee, and if admitted, these requirements must be completed during the first year of study. In addition, the applicant must have strong interest in pursuing statistical research in biomedical sciences.

Fields of Study

- Survival and competing risks data analysis
- Big data analysis
- High dimensional data analysis
- Bayesian statistics and Bayesian machine learning
- Clinical trials
- Machine learning
- Statistical genetics
- Bioinformatics
- Personalized medicine, causal inference
- Image data analysis
- Missing data

Credits Required to Graduate

60 credits minimum

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.

BIOE 10444 Research Ethics Discussion Series. *1 credit.*
Prerequisite: 10222 Ethics and Integrity in Science.
The course is directed by members of the Bioethics Faculty and provides facilitated discussions of a series of topics in research ethics. Discussions are led by members of the Basic Science faculty and are focused on ethical issues that commonly come up in biomedical research. The course is meant to not only reinforce the basic ethics taught in the online course Ethics and Integrity in Science, which is a prerequisite, but also to explore the gray areas of the individual topics. The intent is to offer students illustrative examples of ethical issues that might arise in their careers, to emphasize the ethical principles that apply in such situations, and the provide practical guidance on how these types of situations should be correctly handled. This course is offered as a discussion series. Students are expected to attend and participate in the discussion. Bioethics 10444 is offered during the spring terms only.

BIOS 04214 Design and Analysis of Clinical Trials. *3 credits.*
Prerequisites: 04231 Statistical Models and Methods I or concurrent registration
This course covers topics in clinical trials including the clinical trial protocol; sources of bias in treatment evaluation; blinding; randomization; sample size and power calculation; phase I, phase II, phase III and hybrid trials; covariate adjustment; interim analysis; stochastic curtailment; safety monitoring; factorial and crossover designs.

BIOS 04220 Research Seminar. *1 credit.*
Students present plans for an analysis of research projects and research data. Projects and examples from classical and current literature are discussed by students and faculty.

BIOS 04221 Biomedical Applications and Consulting. *3 credits.*
Prerequisites: 04231 Statistical Models and Methods I
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 04222 Statistical Consulting. *1 credit.*
Prerequisites: 04231/04231 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 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 graphical displays and/or statistical analysis. It will cover the techniques necessary to write R functions and SAS macros for developing new/modified data analysis methods. Students are expected to be somewhat facile in the use of computers before they take this course. UNIX/Linux and working with a cluster are also emphasized. SQL, C/C++ and LaTeX will be introduced as time permits.

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/signed-rank tests, simple linear regression model and inference, one-way ANOVA, two-way ANOVA, Kruskal-Wallis one-way ANOVA. SAS/R will be used throughout the course.

BIOS 04232 Statistical Models and Methods II. *3 credits.*
Prerequisite: 04231 Statistical Models and Methods I
This course will cover various regression models for independent and correlated data. Topics include multiple linear regression, model diagnostics, variable selection, influence/leverage, outliers, collinearity, transformation, GLM including logistic and Poisson regression, overdispersion, GEE, mixed models, and GLMM. SAS/R will be used throughout the course.

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, random forests, and deep learning. 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, boosting, and quantile regression.

BIOS 04240: Biostatistics Inference I. *3 credits.*
This course introduces fundamental concepts and theory in statistics. Topics cover set theory, random variables, distributions of random variables, transformation and expectations, exponential families, joint/marginal/conditional distributions for multiple random variables, random samples, and convergence concepts.

BIOS 04241: Biostatistics Inference II. *3 credits*.
This course introduces advanced concepts and theory in probability and statistics. Topics cover the sufficiency and likelihood principles, methods of moments, maximum likelihood estimators, Bayes estimators, methods of evaluating estimators, likelihood ratio tests, Bayesian tests, most powerful tests, interval estimation, and asymptotic evaluations.

BIOS 04275 Applied Survival Analysis. *3 credits.*
Prerequisites: 04231 Statistical Models and Methods I
Parameters of interest in time-to-event studies; censoring and truncation; survival and competing risks outcomes; univariate estimation, including the Kaplan-Meier, Nelson-Aalen, and Aalen-Johansen estimators; tests comparing two or more samples, including the log rank test and Gray's test; semi-parametric regression, including the Cox proportional hazards, Aalen's additive, and Fine-Gray models; multi-state modeling; methods for high dimensional data analysis and machine learning may also be covered.

BIOS 04285 Introduction to Bayesian Analysis. *3.5 credits.*
Prerequisites: 04231 Statistical Models and Methods I
This course introduces basic concepts with computational tools for Bayesian statistical methods extolling the dynamism of the likelihood, prior, posterior and predictive distributions. Topics covered include one and two sample inference, regression models and comparison of several populations with normal, dichotomous and count data. An introduction to modern Bayesian software such as NIMBLE and Stan along with a basic understanding of Markov chain Monte Carlo (MCMC) via Gibbs sampling, Metropolis-Hastings, slice sampling and state-of-the-art techniques such as hybrid/Hamiltonian MCMC variants like No U-Turn Sampler (NUTS).

BIOS 04295 Reading and Research. *1-9 credit(s).*
The course of study for Reading and Research is designed by each student with his/her advisor to focus on readings in literature in the student's field, to build bibliographic resources for the dissertation, and to conduct supervised, independent research.

BIOS 04313 Advanced Statistical Computing. *3 credits.*
Prerequisites: 04232 Statistical Models and Methods II, Statistical Inference II, 04224 Biostatical Computing*
This is a course for performing sophisticated statistical simulation studies using computer. We will introduce some numerical algorithms which are useful for statistical modeling and data analysis. The topics include: random number generations, acceptance/rejection method, unconstrained/constrained likelihood optimizations, Kuhn-Tucker conditions, regularized regressions, the EM algorithm, numerical integration, Gaussian quadrature, Monte-Carlo method (importance sampling), Bootstrap, Permutation, Jackknife, Cross-validation.

BIOS 04363 Advanced Statistics I. *3 credits.*
*Prerequisites: 04232 Statistical Models and Methods II, Statistical Inference II**
This course covers both the theoretical framework and practical aspects of likelihood-based inference. Topics covered include main elements of likelihood inference, properties of likelihood, exponential families and generalized linear models, large sample properties of likelihood-based inference, likelihood-based regression models, generalized estimating equations, and conditional, marginal likelihood, and asymptotics of penalized regression.

BIOS 04365 Linear Models I. *3 credits.*
*Prerequisites: Statistical Inference II**
Review of matrix algebra and vector spaces, multivariate normal distribution and quadratic forms, least squares estimation, testing nested models, weighted least squares, one-way ANOVA, testing contrasts, multiple comparison, partial and multiple correlation coefficients, polynomial regression, lack-of-fit tests.

BIOS 04384 Statistical Genetics. *3 credits.*
*Prerequisites: 04365 Linear Models I, Statistical Inference II**
This course will cover the fundamental concepts in population genetics and statistical models and methods on genetic linkage and association mapping studies. Topics include Mendelian inheritance, Hardy-Weinberg equilibrium, linkage disequilibrium, allele identity-by-descent (IBD), inbreeding and kinship coefficients, statistical models for quantitative traits, heritability, genetic variance components, linkage analysis, association analysis, case-control analysis, haplotype association analysis, adjust for familial relatedness and population structure, analysis of genetic rare variants, polygenic risk score analysis, LD score regression, statistical fine mapping.

BIOS 04385 Advanced Bayesian Analysis. *3.5 credits.*
Prerequisites: Introduction to Bayesian Analysis, Applied Survival Analysis
A combination of advanced Bayesian principles, tools and methods. Emphasis is on modern computations for parametric and nonparametric models with a deeper dive into NIMBLE/Stan and state-of-the-art sampling techniques, convergence diagnostics, goodness-of-fit, etc. Topics include Bayes factors, HPD regions, conjugate/non-informative priors, the generalized linear models, hierarchical/mixed models, multivariate data, restricted parameter spaces/time-to-event analysis with censored data, Dirichlet Process Mixtures, Gaussian Processes, Bayesian Additive Regression Trees (BART), advanced computational techniques like stochastic gradient descent and illustrative examples of Bayesian analyses for complex biomedical data.

BIOS 04386 Theory of Survival Analysis. *3 credits.*
*Prerequisites: 04275 Applied Survival Analysis, Statistical Inference II**
This course will provide students with a solid foundation in both classical and modern theory of survival analysis including mathematical theory of counting processes, martingales, and empirical processes; asymptotic properties for estimation of the survival, cumulative hazard, and cumulative incidence functions; extensions of k-sample nonparametric tests to survival data; sample size and power calculation; proportional hazards, additive hazards, and proportional subdistribution hazards regression models; multivariate survival analysis; methods for high dimensional data analysis and machine learning may also be covered.

BIOS 04399 Doctoral Dissertation. *1-9 credit(s).*
This course is required for the completion of the PhD degree. The PhD candidate must submit a dissertation based on original research of a high scholarly standard that makes a significant contribution to knowledge in their chosen field.

BIOS 24150 Bioinformatics in Omics Analysis. *3 credits.*
Prerequisites: 04231 Statistical Models and Methods I and 04224 Biostatistical Computing, or consent of instructor.
The course aims to introduce modern statistical and computational methods in high-throughput 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).

*Taken at UW-Milwaukee

Courses taken at UW-Milwaukee
MTHSTAT 871 Mathematical Statistics I. *3 credits.*
MTHSTAT 872 Mathematical Statistics II. *3 credits.*

Required Courses as Needed

BIOS 04002 Master's Thesis Continuation. *0 credits*.
Prerequisite: 04299 Master's Thesis
This is a form of registration available to students who have completed all of the required coursework, including thesis credits but have not yet completed the writing of the Thesis. Continuation status is limited to three consecutive terms following the completion of Thesis credits.

BIOS 04003 Doctoral Dissertation Continuation. *0 credits*.
Prerequisite: 04339 Doctoral Dissertation
This is a form of registration available to students who have completed all of the required coursework, including dissertation credits but have not yet completed the writing of the Dissertation. Continuation status is limited to three consecutive terms following the completion of Dissertation credits.

BIOS 04299 Master's Thesis. *1-6 credits.*
Students in the Ph.D. degree program who cannot or elect not to complete that program may be allowed to transfer to the Master's program. This transfer must be approved by the student's advisor, the Program Director, the Chair, and the Graduate School. To transfer to the Master's Program, the student must be in good academic standing according to regulations established by the Graduate School.

Elective Courses

A minimum of six credit hours of graduate-level electives in a non-statistical field such as biological/medical science, mathematics, and computer science is a requirement for a candidate seeking the PhD degree in Biostatistics.