BIOST 2000 TEACHING PRACTICUM Credit(s): 03.0

THIS COURSE WILL PROVIDE DOCTORAL STUDENTS WITH AN OPPORTUNITY TO OBTAIN TEACHING EXPERIENCE. THIS COURSE IS INTENDED FOR DOCTORAL STUDENTS DURING THEIR DISSERTATION STAGE. TEACHING EXPERIENCE WILL ENHANCE THE PROFESSIONAL GROWTH OF STUDENTS. STUDENTS WILL FURTHER DEVELOP ORAL AND WRITTEN COMMUNICATION SKILLS AND AN ART FOR EXPLAINING MATERIAL, WHICH IS AN INTEGRAL PART OF A BIOSTATISTICIAN'S CAREER.

BIOST 2011 PRINCIPLES STATISTICAL REASONING - RECITATION Credit(s): 00.0

CLASSROOM INSTRUCTION USUALLY ASSOCIATED WITH A LECTURE WHICH FACILITATES INTERACTION BETWEEN THE STUDENT AND THE INSTRUCTOR.

RECITATION for BIOST 2011 effective 2016, Term 2171.
(When enrolling in BIOST 2011 you will also enroll in one of the two required Recitations)

BIOST 2015 ELEMENTS STATISTICAL LEARNING Credit(s): 03.0

THE PURPOSE OF THE COURSE IS TO PRESENT THE THEORY AND PRACTICE OF STATISTICAL LEARNING ALGORITHMS, PLACING "STATISTICAL LEARNING" OR "DATA MINING" TECHNIQUES IN THE PROPER CONTEXT WITH REGARD TO THEIR ORIGINS IN SIMPLE CLASSICAL METHODS LIKE LINEAR REGRESSION, TO CLARIFY THE STRENGTHS AND WEAKNESSES FROM THEORETICAL AND PRACTICAL SIDES. "SUPERVISED LEARNING" TECHNIQUES STUDIED INCLUDE USING REGULARIZATION AND BAYESIAN METHODS, KERNEL METHODS, BASIS FUNCTION METHODS, NEURAL NETWORKS, SUPPORT VECTOR MACHINES, ADDITIVE TREES, BOOSTING, BOOTSTRAP-BASED METHODS. UNSUPERVISED LEARNING TECHNIQUES STUDIED INCLUDE CLUSTER ANALYSIS, SELF-ORGANIZING MAPS, INDEPENDENT COMPONENT ANALYSIS, AND PROJECTION PURSUIT.

BIOST 2016 SAMPLING DESIGN AND ANALYSIS Credit(s): 03.0

Prerequisite(s): BIOST 2011 or BIOST 2041 and BIOST 2093

THIS IS AN APPLIED STATISTICAL METHODS COURSE DESIGNED TO PROVIDE STUDENTS WITH A WORKING KNOWLEDGE OF INTRODUCTORY AND INTERMEDIATE-LEVEL SAMPLING DESIGNS AND ASSOCIATED METHODS OF STATISTICAL ANALYSIS ALONG WITH A BASIC UNDERSTANDING OF THEORETICAL UNDERPINNINGS. STUDENTS WILL ALSO LEARN SURVEY PROCEDURES IN THE SAS AND STATA SOFTWARE PACKAGES. EMPHASIS IS PLACED ON SAMPLING HUMAN POPULATIONS IN LARGE COMMUNITIES. LECTURE TOPICS INCLUDE: SIMPLE PROBABILITY SAMPLES, STRATIFIED SAMPLING, RATIO AND REGRESSION ESTIMATION, CLUSTER SAMPLING, SAMPLING WITH UNEQUAL PROBABILITIES, VARIANCE ESTIMATION AND WEIGHTING IN COMPLEX SURVEYS, TWO-PHASE SAMPLING, ESTIMATING POPULATION SIZE AND ESTIMATION OF RARE POPULATIONS AND SMALL AREAS. THE COURSE WILL CONSIST OF TWO WEEKLY 1.5 HOUR LECTURES INCLUDING TWO SPECIAL CLASSES DEVOTED TO USING SAS AND STATA.

(A working knowledge of first-term level calculus (e.g. BIOST 2081) is recommended but not required.)

BIOST 2018 STATISTICAL FDS BIOINF DATA MINING Credit(s): 03.0

COURSE INTRODUCES DATA ANALYSIS METHODS WIDELY USED OR GAINING USE IN BIOINFORMATICS. METHODS DEAL WITH PREDICTION, CLASSIFICATION, OPTIMIZATION, AND CLUSTERING; INCLUDE CLASSIFICATION TREES, FLEXIBLE VARIETIES OF DISCRIMINANT ANALYSIS INCLUDING SUPPORT VECTOR MACHINES, EM ALGORITHM AND MONTE CARLO MARKOV CHAIN, BOOTSTRAP AND BAGGING, BOOSTING AND SELF-ORGANIZING MAPS. METHODS ARE IN CONTEXT OF PRINCIPLES AND MODELS OF STATISTICAL SCIENCE, WITH EMPHASIS ON BAYESIAN METHODS. EXAMPLES ARE FROM MICROARRAYS, ANALYSIS OF GENETIC NETWORKS, PROTEONICS, COMPUTATIONAL PHARMACOLOGY AND RESEARCH TEXT MINING.
BIOST 2021 SPECIAL STUDIES  
Credit(s): 01.0 to 15.0  
QUALIFIED STUDENTS MAY UNDERTAKE ADVANCED WORK OR RESEARCH WITH THE APPROVAL AND UNDER THE GUIDANCE OF A MEMBER OF THE STAFF.

BIOST 2025 BIOSTATISTICS SEMINAR  
Credit(s): 01.0  
BIOMETRY SEMINARS INTRODUCE THE STUDENTS TO CURRENT HEALTH PROBLEMS INVOLVING THE APPLICATION AND DEVELOPMENT OF BIOSTATISTICS METHODS AND THEORY.

BIOST 2040 ELEMENTS OF STOCHASTIC PROCESSES  
Credit(s): 03.0  
Prerequisite(s): College Calculus and BIOST 2043  
COVERS GENERATING FUNCTIONS AND CONVOLUTIONS OF RANDOM VARIABLES, THE POISSON AND COMPOUND POISSON DISTRIBUTIONS, BRANCHING PROCESSES, RANDOM WALK, AND THE GAMBLER'S RUIN PROBLEM, MARKOV CHAINS, AND SIMPLE BIRTH AND DEATH PROCESSES.

BIOST 2041 INTRO TO STATISTICAL METHODS 1  
Credit(s): 03.0  
DISCUSSES TECHNIQUES FOR THE APPLICATION OF STATISTICAL THEORY TO ACTUAL DATA. TOPICS INCLUDE PROBABILITY THEORY, ESTIMATION OF PARAMETERS, AND TESTS OF HYPOTHESIS FOR BOTH THE DISCRETE AND CONTINUOUS CASE.

(Effective fall 2015, term 2161: When enrolling in BIOST 2041 you will also enroll in a required Recitation.)

BIOST 2041 INTRO TO STATISTICAL METHODS 1 - RECITATION  
Credit(s): 00.0  
CLASSROOM INSTRUCTION USUALLY ASSOCIATED WITH A LECTURE WHICH FACILITATES INTERACTION BETWEEN THE STUDENT AND THE INSTRUCTOR.

(RECITATION for BIOST 2041, effective fall 2015, term 2161:  
You will also enroll in the required Recitation.)

BIOST 2042 INTRO TO STATISTICAL METHODS 2 - RECITATION  
Credit(s): 00.0  
CLASSROOM INSTRUCTION USUALLY ASSOCIATED WITH A LECTURE WHICH FACILITATES INTERACTION BETWEEN THE STUDENT AND THE INSTRUCTOR.

(RECITATION for BIOST 2042  
(Effective spring 2016, term 2164)

BIOST 2042 INTRO TO STATISTICAL METHODS 2  
Credit(s): 03.0  
Prerequisite(s): BIOST 2041  
MORE TECHNIQUES ARE GIVEN FOR THE APPLICATION OF STATISTICS TO ACTUAL DATA WITH EMPHASIS ON DISTRIBUTION-FREE AND MULTIVARIATE METHODS. INTERPRETATION OF RESULTS AND CONCEPTS WILL BE STRESSED.

(Effective spring 2016, term 2164: When enrolling in BIOST 2042 you will also enroll in a required Recitation.)

BIOST 2043 INTRO TO STATISTICAL THEORY 1  
Credit(s): 03.0  
Prerequisite(s): College Calculus  
COVERS JOINT, MARGINAL, AND CONDITIONAL PROBABILITIES; DISTRIBUTIONS OF RANDOM VARIABLES AND FUNCTIONS OF RANDOM VARIABLES; EXPECTATIONS OF RANDOM VARIABLES AND MOMENT GENERATING FUNCTIONS; LAW OF LARGE NUMBERS; CENTRAL LIMIT THEOREM.
**INTRO TO STATISTICAL THEORY 2**

**Credit(s):** 03.0  
**Prerequisite(s):** BIOST 2043

Covers elements of statistical inference; sampling distributions of estimators; Rao-Cramer's inequality; problems of testing statistical hypotheses; Neyman-Pearson lemma; likelihood ratio tests.

**ANALYSIS OF COHORT STUDIES**

**Credit(s):** 03.0  
**Prerequisite(s):** BIOST 2042 and BIOST 2049

This introductory applied course in statistical modeling focuses on maximum likelihood and related regression methods for the analysis of cohort data. Topics include generalized linear models, generalized estimating equations, and generalized linear mixed models. The course emphasizes logistic and Poisson regression, and discrete survival models with time-dependent covariates. Students analyze several cohort data sets, assess the adequacy of their models, and interpret their results.

**APPLIED REGRESSION ANALYSIS**

**Credit(s):** 03.0  
**Corequisite(s):** BIOST 2042

Deals with basic analytic techniques of regression analysis with special emphasis on valid interpretations of results using such techniques. Analysis with computer packaged programs is stressed.

**STATISTICAL ESTIMATION THEORY**

**Credit(s):** 03.0  
**Prerequisite(s):** BIOST 2042 and BIOST 2044

Fisher's information; Rao-Cramer inequality and sufficient statistics; Bhattacharyya bounds; Rao-Blackwell theorem; methods of moments; the method of maximum likelihood; Newton-Raphson method and Rao's scoring for parameters; estimation of several parameters; order statistics and life-testing problems; estimation with censored data and survival analysis.

**MULTIVARIATE ANALYSIS**

**Credit(s):** 03.0  
**Prerequisite(s):** BIOST 2044

Multivariate normal distribution, estimation of the mean vector and covariance matrix, distributions and uses of simple, partial and multiple conelation correlation coefficients, the generalized T² statistic, the distribution of the sample generalized variance, multivariate analysis of variance and the multivariate Behrens-Fisher problem. Multivariate methods are applied to repeated measures analysis, factor analysis, and discriminant analysis. The beginning of the course emphasizes theory. Later applications and computational methods are emphasized. Lectures are of current and classical literature. Prerequisite: Biost 2044 or permission of instructor.

**SURVIVAL ANALYSIS**

**Credit(s):** 03.0  
**Prerequisite(s):** BIOST 2042 and BIOST 2044

Introduces the student to the design considerations and statistical analysis of clinical trials. Covers the theoretical aspects of various models in reliability theory and the proportional hazards model, as well as the more applied problems of interpreting specific data sets and designing large-scale trials.

**INTROD GNOMC ANAL 1: APPLCS**

**Credit(s):** 03.0  
**Prerequisite(s):** BIOST 2094

This 3-credit course is a graduate level introduction and overview of modern high-throughput genomic data analysis. It is designed for graduate students in biostatistics and human genetics who are interested in the technology and elementary data mining of high-throughput genomic data (including but not limited to classical expression arrays, various array-based applications, next-generation sequencing and proteomics). The course is also helpful for biology students with basic quantitative training (e.g. two elementary statistics courses and R programming) who have interests in understanding the intuition and logic underlying the data analysis methods. R is the major language used in the course.
<table>
<thead>
<tr>
<th>Course Code</th>
<th>Course Title</th>
<th>Credit(s)</th>
<th>Prerequisites</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIOST 2056</td>
<td>INTRO TO DIAG TEST EVAL &amp; ROC</td>
<td>0.3</td>
<td>BIOST 2041 and BIOST 2042 and BIOST 2043 and BIOST 2044</td>
</tr>
<tr>
<td>BIOST 2058</td>
<td>SCIENTFC COMMUNICATION SKILLS</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td>BIOST 2061</td>
<td>LIKELIHOOD THEORY &amp; APPLICATN</td>
<td>0.2</td>
<td>BIOST 2044</td>
</tr>
<tr>
<td>BIOST 2062</td>
<td>CLINICAL TRIALS: METHDS &amp; PRACT</td>
<td>0.3</td>
<td></td>
</tr>
<tr>
<td>BIOST 2065</td>
<td>ANALYSIS OF INCOMPLETE DATA</td>
<td>0.3</td>
<td>BIOST 2049 and BIOST 2051 and BIOST 2061</td>
</tr>
<tr>
<td>BIOST 2066</td>
<td>APLD SURVIVAL ANAL METHS &amp; PRA</td>
<td>0.3</td>
<td>BIOST 2042 and BIOST 2049</td>
</tr>
<tr>
<td>BIOST 2077</td>
<td>SPECIAL TOPICS</td>
<td>0.3</td>
<td></td>
</tr>
</tbody>
</table>
BIOST 2078 INT GNOMC ANAL 2: THRY ALGRTHM
Prerequisite(s): BIOST 2041 and BIOST 2042 and BIOST 2055 and BIOST 2094
Credit(s): 03.0

This course is a graduate level course to introduce theories and algorithms for statistical analysis of high-throughput genomic data. Emphasis will be given to high-dimensional data analysis and theories behind the commonly used methods. It is designed for graduate students who already have sufficient statistical background, have basic knowledge of various high-throughput genomic experiments and wish to learn advanced statistical theories for bioinformatics and genomics research. [Prequisites: Biost 2041 and 2042 or equivalent; proficiency in R programming (Biost 2094 Statistical Computing in R) and high-throughput genomic data analysis experiences (Biost 2055).]

BIOST 2081 MATHEMATICAL METHODS FOR STAT
Credit(s): 03.0

Differentiation and integration of functions of several variables. Infinite sequences and series. Fundamentals of matrix algebra. Class examples and homework problems will emphasize applications to probability and statistics.

(Enrollment requirement: Biostatistics (PhD; MPH; MS) effective fall 2015, term 2161.)

BIOST 2083 LINEAR MODELS
Prerequisite(s): BIOST 2044
Credit(s): 03.0

Acquaints students with linear model techniques for analyzing both balanced and unbalanced data. The topics covered include generalized inverses, orthogonal contrasts with unbalanced data, and analysis of covariance. Analysis with computer packaged programs is discussed.

BIOST 2086 APPLIED MIXED MODELS ANALYSIS
Prerequisite(s): BIOST 2083
Credit(s): 03.0

Mixed model analysis provides a new approach to modeling which allows one to relax the usual independence assumptions and take into account complicated data structures. This course will consider all types of mixed models into a general framework and consider the practical implications of their use. Topics will include: normal mixed models, generalized mixed models, and mixed models for categorical data, repeated measures data analysis and cross-over trials with mixed models. Software for fitting mixed models will be discussed.

BIOST 2087 BIOST CONSULTING PRACTICUM
Credit(s): 01.0

Provides advanced students (second-year masters and doctoral) with exposure and practical experience in consulting on the biostatistical aspects of research problems arising in the biomedical or other allied fields. Students initially under the supervision of a faculty member participate in discussions with investigators leading to the design and/or analysis of a current research problem.

BIOST 2087 BIOST CONSULTING PRACTICUM - LAB
Credit(s): 00.0

Laboratory

BIOST 2093 SAS DATA MANAGEMENT & ANALYSIS
Corequisite(s): BIOST 2041
Credit(s): 02.0

The goal of this course is to provide students with an understanding of the SAS program environment as well as the skills needed to use SAS as a tool to conduct research, prepare data, and perform analyses. Upon completion of the course the student will have an understanding of SAS at the intermediate level. The course covers the utility of SAS as a data management, data manipulation, and data analysis tool. The focus will not be statistical analysis, but rather how to use SAS as a programming tool. Emphasis will be placed on program code writing. Concepts will be illustrated with numerous examples from basic descriptive analysis.

(Note: Students should have a basic understanding of the PC computer environment with some exposure to the Windows operating system.)
BIOST 2094  ADVANCED R COMPUTING  Credit(s):  02.0

Prerequisite(s):  BIOST 2041 and BIOST 2043

An advanced statistical computing course using R designed for graduate level biostatistics students with programming experience in R or other low-level languages such as C, C++, Java, and/or Fortran. Experience in SAS and/or Stata does not qualify. The course will cover topics, including but not limited to, R in modeling and optimization, advanced R graphics, functional programming, object-oriented field guide, efficient computing in R, GUI for R-shiny, embedding C/C++, R package/documentation, Julia, Github etc. The course will also include real life application for students to practice the programming techniques learned in class.

[Revised course: for biost students; title, course description, pre-reqs changed, effective 4/30/17.]

BIOST 2096  NUMERICAL METHODS IN BIOSTATISTICS  Credit(s):  03.0

Prerequisite(s):  BIOST 2044 and BIOST 2049

THE PURPOSE IS TO FAMILIARIZE STUDENTS WITH A BROADER RANGE OF NUMERICAL METHODS WHICH ARE USEFUL IN BIOSTATISTICAL RESEARCH. SELECTED COMPUTATIONAL TECHNIQUES USED IN STATISTICAL RESEARCH WILL BE COVERED. SOME BACKGROUND WILL BE PROVIDED TO FACILITATE UNDERSTANDING OF A FEW NUMERICAL ALGORITHMS WIDELY USED IN STATISTICS. THE FOLLOWING ARE COVERED: RECURRENCE RELATIONS, POWER SERIES AND ASYMPTOTIC EXPANSIONS, GENERATING PSEUDO-RANDOM DEVIATES, BASIC SIMULATION METHODOLOGY, SOLUTIONS OF NONLINEAR EQUATIONS, NEWTON'S METHOD, VECTOR AND MATRIX NORMS, LINEAR REGRESSION AND MATRIX INVERSION.

BIOST 3010  RESEARCH AND DISSERTATION PHD  Credit(s):  01.0 to 15.0

DISSERTATION CREDITS FOR QUALIFIED DOCTORAL STUDENTS IN THE DEPARTMENT OF BIOSTATISTICS.

FTDR 3999  FULL-TIME DISSERTATION RESEARCH  Credit(s):  00.0

DOCTORAL CANDIDATES WHO HAVE COMPLETED ALL CREDIT REQUIREMENTS FOR THE DEGREE, INCLUDING ANY MINIMUM DISSERTATION REQUIREMENTS, AND ARE WORKING FULL-TIME ON THEIR DISSERTATIONS MAY REGISTER FOR THIS COURSE. WHILE THE COURSE CARRIES NO CREDITS AND NO GRADE, STUDENTS WHO ENROLL IN "FULL-TIME DISSERTATION STUDY" ARE CONSIDERED BY THE UNIVERSITY TO HAVE FULL-TIME REGISTRATION STATUS.