DEPARTMENT: **Biostatistics and Bioinformatics**

COURSE NUMBER: BIOS 570 **SECTION NUMBER:** **SEMESTER: Spring**

CREDIT HOURS: 2

COURSE TITLE: Methods in Statistical Genetics

CLASS HOURS: The first half of spring semester, Tuesday and Thursday, 3:00-4:50pm

INSTRUCTOR NAME Yijuan Hu, PhD

INSTRUCTOR CONTACT INFORMATION

EMAIL: yijuan.hu@emory.edu

PHONE: 404-712-4466

SCHOOL ADDRESS OR MAILBOX LOCATION: 1518 Clifton Rd NE, GCR Building, Room 342

OFFICE HOURS: Instructor**:** TBA

BRIEF COURSE DESCRIPTION

This is an introductory course for graduate students in Biostatistics, Bioinformatics, Epidemiology, Genetics, Computational Biology, and other related quantitative disciplines. The course will conduct a comprehensive survey of statistical methods for analysis of family-and population-based genetic data, including classical linkage analysis, population-based and family- based association analysis, genome- wide association studies (GWAS) and analysis of next-generation sequencing data. Because this course serves as a prerequisite to BIOS 770 Advanced Statistical Genetics, the focus of the course will be on identifying statistical problems, relating genetic concepts to statistical model assumptions, introducing the latest statistical methods, and ultimately preparing students for in-depth understanding/research of statistical methodologies on analysis of genetic data.

Prerequisites

BIOS506, EPID530, or equivalent, or instructor’s permission

Suggested Textbooks

The Fundamentals of Modern Statistical Genetics (2011) by Nan M. Laird and Christoph Lange
Statistical Methods in Genetic Epidemiology (2004) by Duncan C. Thomas
A Statistical Approach to Genetic Epidemiology (2006) by Andreas Ziegler and Inke R. Konig

EVALUATION

Participation (10%)

Homework (40%)

Midterm take-home exam (30%)

Final presentation (20%)

LIST SCHOOL LEVEL, DEPARTMENT, AND/ OR PROGRAM COMPETENCIES

- Use analytic reasoning and quantitative methods to address questions in public health and population-based research
- Perform the appropriate statistical analyses of study data
- Use computer statistical software for both data management and data analyses
- Assist in the interpretation of study results
- Interpret statistical results of biomedical studies effectively
- Apply existing statistical theory and methods to a broad range of medical or public health problems
- Conduct appropriate statistical analyses for a broad range of applications
- Communicate the results of statistical studies both orally and in writing to senior statisticians and other investigators

LIST LEARNING OBJECTIVES ASSOCIATED WITH THE COMPETENCIES

By the end of the course, students will be able to:

1. Translate genetic questions into statistical models
2. Identify software packages for analysis of genetic data
3. Identify appropriate statistical methods (implemented in software packages) for analysis
4. Understand the results produced by the genetic software packages
5. Communicate the results both orally and in writing to senior statisticians and other investigators.

Tentative Schedule and Topics

Lecture 1	**Basic Concepts of Molecular Genetics** • Mitosis and Meiosis, central dogma of molecular biology • Genetic variants (SNP, CNV), Mendel's Laws	
Lecture 2	**Basic Concepts of Population Genetics** • Hardy-Weinberg Equilibrium (HWE) • Recombination, Linkage Disequilibrium (LD) • Coalescent model	
Lecture 3	**Coalescent Model** • Allele frequency distribution	
Lecture 4	**Methods for Haplotype Frequency Estimation** • Maximum likelihood estimation (MLE), EM algorithm • Haplotype frequency estimation by EM algorithm • Haplotyping algorithms	Homework 1
Lecture 5	**Methods for Genotype Imputation** • HapMap project, 1000 Genomes project • Hidden Markov Model (HMM), MaCH	
Lecture 6	**Methods for Parametric Linkage Analyais of Binary Traits** • Map function • Direct count method, LOD score method	
Lecture 7	**Methods for Nonparametric Linkage Analyais of Binary Traits** • IBS vs. IBD • Affected sib pair method	
Lecture 8	**Methods for Linkage Analysis of Quantitative Traits** • Heritability, Kinship coefficient • Variance component methods	Homework 2
Lecture 9	**Methods for Family-Based Association Analysis** • Population stratification • Transimission Disequilibrium Test (TDT), FBAT	
Lecture 10	**Methods for Population-Based Association Analysis** • Genotype/allele-based test, Cochran-Armitage trend test • Logistic regression to include other covariates	
Lecture 11	**Methods for Genome-Wide Association Studies** • HapMap, LD block structure, tag SNPs • Genotype error, Quality control • Manhatton plot, QQ plot	Homework 3
Lecture 12	**Advanced Methods for GWAS** • Population stratification, Multiple testing, Winner's curse • Replication, Meta-analysis	
Lecture 13	**Methods for Next-Generation Sequencing and Epigenetics** • Sequencing technologies, read alignment, variant calling • Epigentics, Methylation data	
Lecture 14	**Methods for Rare Variants and Copy Number Variants** • Various rare variant tests • CNV data, detection and association testing	Homework 4