BCB725 Introduction to Statistical Genetics

(offered Spring every odd year)

Course Description

This is an introductory course for graduate students in Computational Biology, Bioinformatics, Biostatistics, Genetics, Statistics, Epidemiology, and other related quantitative disciplines. The course will cover statistical methods for the analysis of family and population based genetic data. Topics covered will include classical linkage analysis, population-based and family-based association analysis, haplotype analysis, genome-wide association studies, basic principles in population genetics, imputation-based analysis, pathway-based analysis, admixture mapping, analysis of copy number variations, and analysis of massively parallel sequencing data. Students will be exposed to the latest statistical methodology and computational tools on gene mapping for complex human diseases. We will also have guest lecturers covering the fundamentals of mice genetics; and other special topics.

Instructors

Yun Li, PhD
Department of Genetics
Department of Biostatistics
Rm5090 GMB (Genetic Medicine Building)
Rm 4115F McGavran Greenberg
(919) 843-2832
yunli@med.unc.edu

Ethan M Lange, PhD
Department of Genetics
Department of Biostatistics
Rm5111 GMB
(919) 966-3356
ethan_lange@med.unc.edu

When   WF 2-3:15pm

Where   3074 Bondurant on Wednesdays
                        1074 Bondurant on Fridays

Office Hours   WF 1-2pm, or by appointment

Prerequisites

Introductory graduate-level courses in statistics or biostatistics (e.g., BCB720, Bios660 or 661) or permission of the instructor.

Recommended Textbooks (not required)

1. The Fundamentals of Modern Statistical Genetics. Nan M. Laird, Christoph Lange. (2011). (UNC ebook available)
2. Handbook on Analyzing Human Genetic Data: Computational Approaches and Software. Edited by Shili Lin and Hongyu Zhao. (2010) (UNC ebook available)
3. A Statistical Approach to Genetic Epidemiology: Concepts and Applications. Andreas Ziegler, Inke R. Koenig (2006).

Evaluation

1. Written homework assignments (40% of the final grade).
2. A final project which will entail a written report and in-class presentation (25-30 minutes) (40% of the final grade). You can 1) analyze a real data using existing methods, or 3) evaluate existing methods by simulations, or 3) develop new methods.
3. Course evaluations (mid-term and final, total 10% of the final grade).
4. Lecture attendance and class participation (10% of the final grade).

Notes:
➢ Late assignment will NOT be accepted.
➢ Some programming is expected for the assignments and final project. Students can choose the programming language they are familiar with (e.g., C/C++, Java, R, Perl, Python, shell scripts). Computing skills are essential for analysis of genetics data. Basic UNIX/LINUX tutorial and example scripts will be posted to sakai course website.

Honor code

We all follow the UNC Honor Code (honor.unc.edu), and suspected violations are reported to the University for investigation. This class involves homework assignments and individual final projects. The general rules for home assignments (unless otherwise explicitly put in writing) are:
1. Students must complete their own work (obviously). In this class, discussions with classmates are encouraged and limited collaboration is allowed. But please submit your own work.
2. Any codes presented in class, on sakai course website, or on Dr. Li's software page (http://www.unc.edu/~yunmli/software.html) can be freely used, re-used, or modified, unless we specifically state that another type of approach must be used.
3. Codes or solutions from elsewhere (i.e. found on the internet) are not allowed without permission from the instructors.
If you seem stuck or need clarification, the safest and most efficient way is to email any questions to us, and then we typically respond to the entire class (without identifying the questioner).

Class Schedule:

Lecture	Date	Instructor	Topic
1	Jan 11, Wed	Li	Basic Genetics Concepts and Terminologies
2	Jan 13, Fri[1]	Li	Hardy-Weinberg Equilibrium, Basic Population Genetics
3	Jan 18, Wed	Li	Coalescent Models
4	Jan 20, Fri[2]	Li	Review of Likelihood Theory
5	Jan 25, Wed	Li	Allele Frequency Estimation, Population-based Association
6	Jan 27, Fri	Li	Linkage Disequilibrium
7	Feb 1, Wed	Lange	EM Algorithm in Genetics
8	Feb 3, Fri	Lange	Haplotype Frequency Estimation
9	Feb 8, Wed	Li	Inferring Haplotype from Genotypes
10	Feb 10, Fri	Tzeng	Haplotype-based Association Tests
11	Feb 15, Wed	Lange	Linkage Analysis for Binary Traits
12	Feb 17, Fri	Lange	Linkage Analysis for Quantitative Traits
13	Feb 22, Wed	Lange	Family-based Association Analysis
14	Feb 24, Fri	Lange	GWAS: Basics
15	Feb 29, Wed	Li	GWAS: Genotype Imputation
16	Mar 2, Fri	Lange	GWAS: Meta-Analysis
	Mar 7, Wed		Spring Break
	Mar 9, Fri		
17	Mar 14, Wed	Li	GWAS: Pathway Analysis
18	Mar 16, Fri	Li	MPS: Basics and Read Alignment
19	Mar 21, Wed	Li	MPS: SNP Genotype Calling
20	Mar 23, Fri	Li	MPS: Design for Association Studies
21	Mar 28, Wed	Wu	Analysis of Rare Genetic Variants
22	Mar 30, Fri	TBD	Special Topic: TBD
23	Apr 4, Wed	TBD	Special Topic: TBD
	Apr 6, Fri[3]		Good Friday
24	Apr 11, Wed	Valdar	Special Topic: Mice Genetics
25	Apr 13, Fri	Valdar	Special Topic: Mice Genetics
	Apr 18, Wed		Final Project Presentation
	Apr 20, Fri		Final Project Presentation

[1] Last day to register: Jan 13, Fri. After that and under extenuating circumstances, talk to Sausyty Hermreck (sausyty_hermreck@med.unc.edu).
[2] Last day to drop class with tuition refund: Jan 23, Mon.
[3] Last day to drop class.