

BSTA 787: Methods for Statistical Genetics and Genomics in Complex Human Diseases Spring 2020

Description: This is an introductory course for graduate students in Biostatistics, Epidemiology, Computational Biology and other BGS disciplines, which will cover statistical methods for the analysis of genetic and genomic data. Topics covered include twin studies, linkage analysis, population-based and family-based association, haplotype analysis, genome-wide association studies (GWAS), analysis of copy number and structural variations, next-generation sequencing data (DNA), analysis of bulk RNA-seq data, single-cell RNA-seq data, and spatial transcriptomics data. Students will be exposed to the latest methodology and computer tools in statistical genetics and genomics.

Prerequisites: Introductory graduate-level courses in statistics or biostatistics (e.g., BSTA 630-631, EPID 520521), or permission of the instructor.

Instructors: Rui Feng, PhD (office: Blockley 209), ruifeng@upenn.edu
Mingyao Li, PhD (office: Blockley 213), mingyao@upenn.edu

Lecture Time: 9:00-10:30am Monday and Wednesday

Lecture Location: Blockley 701

Office Hour: by appointment

Evaluation:

1. Two homework assignments (20% of the final grade). Submit your homework on Canvas in word or pdf. Late homework is NOT accepted.
2. One paper presentation (30% of the final grade). The presenter is expected to clearly explain the content of the assigned paper and lead discussion (about 45 minutes) during class. Please meet with the instructor before your presentation.
3. Final project (40% of the final grade). Students will submit a written report (at least 2,000 words in the format of a scientific article with Abstract, Introduction, Methods, Results, and Discussion) at the end of the semester. Students will give a 15-minute presentation on their projects in class on May 4, 6, and 11. Students can 1) analyze real data using existing methods or new methods proposed by yourself, or 2) evaluate existing methods by simulations and applications to real data, or 3) do a comprehensive literature review on a specific topic. You are welcome to discuss your project with the instructor. The topic of the project needs to be approved by the instructor by April 3, 2020.
4. Lecture attendance and class participation (10% of the final grade).

Recommended (not Required) Textbook:

1. A Statistical Approach to Genetic Epidemiology: Concepts and Applications. Andreas Ziegler, Inke R. Koenig (2006).
2. Handbook on Analyzing Human Genetic Data: Computational Approaches and Software. Shili Lin, Hongyu Zhao (2010). <http://www.springerlink.com/content/978-3-540-69263-8>
3. Statistical Analysis of Next Generation Sequencing Data. Somnath Datta, Daniel Nettleton (2014). <http://www.springer.com/us/book/9783319072111>.
4. Computational Methods for Single-Cell Data Analysis. Guo-Cheng Yuan (2019). <https://www.springer.com/gp/book/9781493990566>.