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

Spring 2019 – Course Syllabus

Time and Location: Monday & Wednesday 9:00am – 10:30am, 418 Blockley Hall

Instructor:
Mingyao Li, Ph.D.
Office: 213 Blockley Hall
Phone: 215-746-3916
Email: mingyao@pennmedicine.upenn.edu

Office hour: appointment by email

Course 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 linkage analysis, population-based and family-based genetic association analysis, haplotype analysis, genome-wide association studies (GWAS), analysis of copy number and structural variations, next-generation sequencing data (DNA), analysis of gene expression data, and single cell RNA sequencing 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 520-521), or permission of the instructor.

Evaluation:
1. Two homework assignments (**30%** of the final grade). Late homework is NOT accepted.

2. One paper presentation (**25%** of the final grade). The presenter is expected to clearly explain the content of the assigned paper and lead discussion during class. Please feel free to meet with the instructor before your presentation.

3. Final project (**40%** of the final grade). A final project is required at the end of the semester. Students will submit a written report (at least 2,000 words in the format of a scientific article with Introduction, Methods, Results, and Discussion), and do an in-class presentation. For the final project, students can 1) analyze real data using existing methods or new methods proposed by yourself, or 2) evaluate existing methods by simulations, 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.

4. Lecture attendance and class participation (**5%** of the final grade).

Recommended Textbooks (not required):

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.