Yale SCHOOL OF PUBLIC HEALTH

Biostatistics

Opening Black Boxes: Enhancing Statistical Rigor in Genomics Data Science

Jingyi Jessica Li, PhD

Associate Professor

Department of Statistics

University of California, Los Angeles

12:00 Noon Eastern time, Tuesday, October 11, 2022

47 College Street, Room 106 A&B

ABSTRACT

The rapid development of genomics technologies has propelled fast advances in genomics data science. While new computational algorithms have been continuously developed to address cutting-edge biomedical questions, a critical but largely overlooked aspect is the statistical rigor. In this talk, I will introduce our recent work that aims to enhance statistical rigor by addressing three issues:

1. Large-scale feature screening (i.e., enrichment and differential analysis of high-throughput data) relying on ill-posed p-values;
2. Double-dipping (i.e., statistical inference on biasedly altered data);
3. Gaps between black-box generative models and statistical inference.