

Data Science Seminar

Hosted by the Department of Mathematics and Statistics

- Date: Tuesday, April 21, 2026
- Time: 12:15pm - 1:15pm
- Room: Whitney Hall 100E
- Speaker: Dr. Yiqun Chen (Johns Hopkins University)
- Title: AI for (Bio)statistics and (Bio)statistics for AI.

Abstract

The rise of AI models has offered much promise to augment expensive human labels with cheap model-generated ones. However, the model predictions could introduce systematic errors, false confidence, and estimates that quietly drift from the truth. This is one half of a broader story about the deep interplay between AI and statistics. In the first half of this talk, we walk through a family of statistical frameworks — drawing on classical ideas from missing data, measurement error, and survey sampling — that use a small number of human-verified labels to correct for the errors in a large pool of noisy AI-generated labels. We show that the core problem reduces to an elegant problem: estimating a conditional expectation, and we connect it to decades of progress in noisy-label and ensemble learning. In the second half, we turn it around and show how statistical principles help AI make better predictions in high-dimensional biological settings. Specifically, we describe how ideas from (bio)statistics improve AI-driven analysis of perturb-seq experiments. Together, these two directions illustrate a productive loop: AI creates new applied opportunities for classical statistical tools, and relevant statistical thinking makes AI work better in the sciences.

Biography of the speaker: Yiqun Chen (he/him) is an Assistant Professor of Biostatistics and Computer Science (courtesy) at Johns Hopkins University. He was previously a data science postdoctoral fellow at Stanford and received his PhD from the University of Washington. His research focuses on statistical inference and evaluation frameworks for modern biomedical datasets, with recent work exploring creative ways to leverage large language models for scientific data analysis and discovery.

From:

<http://www2.math.binghamton.edu/> - **Department of Mathematics and Statistics, Binghamton University**

Permanent link:

<http://www2.math.binghamton.edu/p/seminars/datasci/042126>

Last update: **2026/04/15 15:09**

