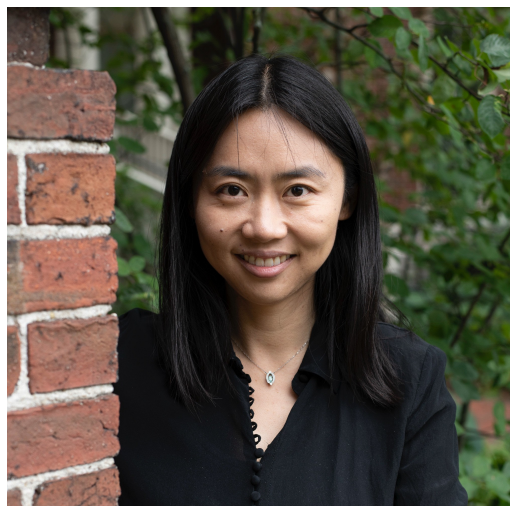


Department of Statistics Seminar Series



Jingyi (Jessica) Li
Department of Biostatistics
Fred Hutchinson Cancer Center

Nullstrap: A Simple, High-Power, and Fast Framework for FDR Control in Variable Selection for Diverse High-Dimensional Models

Tuesday, November 4th, 2025

12:10-1:30 PM

Hill 552

Zoom Meeting: Meeting ID: 983 2864 8232

Password: 665123

<https://rutgers.zoom.us/j/98328648232?pwd=gcHifQ9OouKZnUY0EHLe8sjNXaV1oj.1>

Light refreshments will be served in Hill 452, 11:40am

Abstract:

Balancing false discovery rate (FDR) control with high statistical power remains a central challenge in high-dimensional variable selection. While several FDR-controlling methods have been proposed, many degrade the original data -- by adding knockoff variables or splitting the data -- which often leads to substantial power loss and hampers detection of true signals. We introduce Nullstrap, a novel framework that controls FDR without altering the original data. Nullstrap generates synthetic null data by fitting a null model under the global null hypothesis that no variables are important. It then applies the same estimation procedure in parallel to both the original and synthetic data. This parallel approach mirrors that of the classical likelihood ratio test, making Nullstrap its numerical analog. By adjusting the synthetic null coefficient estimates through a data-driven correction procedure, Nullstrap identifies important variables while controlling the FDR. We provide theoretical guarantees for asymptotic FDR control at any desired level and show that power converges to one in probability. Nullstrap is simple to implement and broadly applicable to high-dimensional linear models, generalized linear models, Cox models, and Gaussian graphical models. Simulations and real-data applications show that Nullstrap achieves robust FDR control and consistently outperforms leading methods in both power and efficiency.

Bio:

Jingyi Jessica Li is a statistician and computational biologist whose work bridges data science and biomedical research. She develops reliable and interpretable methods to analyze complex biological data, with a focus on understanding how genes function and are regulated in health and disease. Her research has helped uncover hidden patterns in gene expression and emphasized the importance of statistical rigor—ensuring that scientific discoveries are trustworthy, even when working with noisy or low-quality data. She strongly believes that statistics is not just a supporting tool but a core driver of progress in biology and medicine. Dr. Li is Professor and Program Head of Biostatistics at Fred Hutchinson Cancer Center, where she holds the Donald and Janet K. Guthrie Endowed Chair in Statistics, and Affiliate Professor of Biostatistics at the University of Washington. Previously, she was Professor of Statistics and Data Science at UCLA, with joint appointments in Biostatistics, Computational Medicine, and Human Genetics. She earned her Ph.D. in Biostatistics from the University of California, Berkeley, and her B.S. in Biological Sciences from Tsinghua University. Her contributions have been recognized with the NSF CAREER Award, Sloan Research Fellowship, ISCB Overton Prize, COPSS Emerging Leaders Award, Guggenheim Fellowship, and the Mortimer Spiegelman Award.