

Xihong Lin

Departments of Biostatistics and Statistics, Harvard University



Xihong Lin is a Professor of Biostatistics and Coordinating Director of the Program in Quantitative Genomics of Harvard T.H. Chan School of Public Health, and Professor of Statistics at Harvard Faculty of Arts and Sciences. Dr. Lin's research interests lie in development and application of statistical and computational methods for analysis of high-throughput genetic and genomic data in epidemiological, environmental and clinical studies, for analysis of complex exposure and phenotype data in observational studies, and statistical learning and inference methods for massive data.

September 29 and 30, 2022 | Free, Hybrid Event

General Lecture: Sep 29, 3:30-4:30 pm

Lessons Learned from the COVID-19 Pandemic: A Statistician's Reflection

In this talk, I will discuss my experience as a statistician involved in COVID-19 research in multiple capacities in the last two years, especially in the early phase of the pandemic. I will reflect on the challenges and the lessons I have learned in pandemic research regarding data collection and access, epidemic modeling and data analysis of Wuhan and US data, open science and real time dissemination of research findings, implementation science, media and public communication, and partnerships between academia, government, industry and civil society. I will also make several recommendations on preparing for the next stage of the pandemic and for future pandemics.



Technical Lecture: Sep 30, 3:30-4:30 pm

Ensemble Methods for Testing a Global Null Hypothesis

Testing a global null is a canonical problem in statistics and has a wide range of applications. In view of the fact of no uniformly most powerful test, prior and/or domain knowledge are commonly used to focus on a certain class of alternatives to improve the testing power, e.g., the class of alternatives in the scenario of the same effect sign or signal sparsity. However, it is generally challenging to develop tests that are particularly powerful against a certain class of alternatives. In this paper, motivated by the success of ensemble learning methods for prediction or classification, we propose an ensemble framework for testing that mimics the spirit of random forests to deal with the challenges. Our ensemble testing framework aggregates a collection of weak base tests to form a final ensemble test that maintains strong and robust power. The key component of the framework is to introduce a certain random procedure in the construction of base tests. We then apply the framework to four problems about global testing in different classes of alternatives arising from Whole Genome Sequencing (WGS) association studies. Specific ensemble tests are proposed for each of these problems, and their theoretical optimality is established in terms of Bahadur efficiency. Extensive simulations are conducted to demonstrate type I error control and power gain of the proposed ensemble tests. In an analysis of the WGS data from the Atherosclerosis Risk in Communities (ARIC) study, the ensemble tests demonstrate substantial and consistent power improvement compared to other existing tests.

For more information, please visit:
canssiontario.utoronto.ca/event/2022-dlss-xihong-lin/



CANSSI ONTARIO

700 University Avenue, 9th Floor, Suite 9005, Toronto, ON, M5G 1X6 | www.canssiontario.utoronto.ca

