Lin Gui

5747 S. Ellis Avenue, Jones 203/204, Chicago, IL, 60637
Email: ybnbxb@gmail.com; Website: https://gl-ybnbxb.github.io/

EDUCATION

- Ph. D. in Statistics, The University of Chicago, Chicago, USA 2020- Present
- M.S. in Statistics, The University of Chicago, Chicago, USA 2018-2020
- B.S. in Statistics, University of Science and Technology of China, Hefei, China 2014- 2018

RESEARCH INTERESTS

- Statistical Inference and Multiple Testing
- Generative Models
- Causal Inference and Machine Learning
- Biostatistics

PUBLICATIONS

- **Aggregating Dependent Signals with Heavy-Tailed Combination Test**
  Lin Gui, Yuchao Jiang, Jingshu Wang
  Preprint
- **Concept Algebra for Score-Based Conditional Models**
  Zihao Wang, Lin Gui, Jeffrey Negrea, Victor Veitch
  NeurIPS 2023
- **Causal Estimation for Text Data with (Apparent) Overlap Violations**
  Lin Gui, Victor Veitch
  ICLR 2023
- **Detecting Multiple Replicating Signals using Adaptive Filtering Procedures**
  Jingshu Wang, Lin Gui, Weijie J. Su, Chiara Sabatti, Art B. Owen
  The Annals of Statistics 50 (4), 1890-1909

RESEARCH

- **A Theoretical and Practical Analysis of the Heavy-Tailed Combination Test for Global Test With Correlated Hypotheses**
  - Undertook comprehensive theoretical evaluations to decipher the intricacies of the state-of-the-art Cauchy combination test and its expansion, termed the heavy-tailed combination test, tailored for the global test with correlated hypotheses.
  - Conducted empirical studies, offering a general practical guideline for the Cauchy/heavy-tailed combination test
  - Enhanced the heavy-tailed combination test into a multiple testing procedure adept at controlling the family-wise error rate (FWER) and introduced a shortcut for this closed testing procedure.
  - Implemented the method on genetic data to address real-world challenges.
- **Gene Transcription Mechanism in 3D Genome**
  - Designed a statistical tool aimed at identifying gene pairs with a high likelihood of co-transcription due to shared enhancers (close to each other). A core challenge is accommodating the 3D dynamic structures of chromosomes. Thus, spatial details of gene pairs must be analyzed using 3D genome data, such as Hi-C data.

CONFERENCES AND PRESENTATIONS

- **ICML 2023 Workshop SPIGM and SCIS** Jul. 2023
  - Concept Algebra for Score-Based Conditional Models
- **ICLR 2023** May. 2023
  - Causal Estimation for Text Data with (Apparent) Overlap Violations
- **2021 Joint Statistical Meetings** Aug. 2021
  - Detecting Multiple Replicating Signals Using Adaptive Filtering Procedures

CODING SKILLS

- R, Python, MATLAB, SQL
- PyTorch, Numpy, Pandas

HONORS & AWARDS

- Nominee, The 37th. Guo Moruo Scholarship (The highest honor at USTC) 2017
- Winner, Outstanding Student Scholarship, USTC 2016-2017
- Winner, China National Scholarship, USTC 2015