Zixiao (Jolene) Wang

+1 (410) 375 8197 | zixiaowang@fas.harvard.edu | Google Scholar | Linkedin

Education

Harvard Univeristy MA, U.S.
PhD in Methods in Epidemiology *Aug 2024-May 2029*

- Advisor: **Professor James M. Robins**.
- Selected courses: MIT 9.520/6.7910: Statistical Learning Theory and Applications; STAT 216: Topics in High-Dimensional Probability Theory

Johns Hopkins University MD, U.S.
Master of Biostatistics *Sep 2022 - May 2024 (Expected)*

- GPA: 4.0/4.0
- Selected courses: Methods in Biostatistics I/II/III/IV, Essentials of Probability and Statistical Inference I/II/III/IV: Probability, Real Analysis
- Teaching TA for Undergraduate Biostatistics course (AS.280.345.01.FA23)

Shanghai Jiao Tong University Shanghai, China
Bachelor of Science in Resource and Environment Science *Sep 2018 - Jun 2022*

- GPA: 3.8/4.0 (rank 3/33)
- Selected courses: Real Analysis, C++, Simulation Methods and Its Applications (MCMC), Linear Algebra, Probability and Statistics, Calculus, Biostatistics and experimental design, Frontiers in Neuroscience and Informatics Methods
- Selected awards: Undergraduate-Merit Scholarship, SJTU (Top 5)
- 2021 Summer School of the Center for Statistical Science at Peking University

Publication

- **Wang, Zixiao**, AmirEmad Ghassami, and Ilya Shpitser. *Identification and Estimation for Nonignorable Missing Data: A Data Fusion Approach.* In Forty-first International Conference on Machine Learning. (**ICML 2024**)
- **Wang, Zixiao**, Yi Feng, and Lin Liu (2022). *Book Review: Semiparametric Regression with R*: Journal of the American Statistical Association 117.540 (2022): 2283-2287.
- Li, Boyang, Yuxuan Wang, **Zixiao Wang**, Xinyue Li, Shannon Kay, Geoffrey L. Chupp, Hongyu Zhao, and Jose L. Gomez. *Shared genetic architecture of blood eosinophil counts and asthma in UK Biobank.* ERJ Open Research 9, no. 4 (2023).
- Niu, Zongwu, **Zixiao Wang**, Yongxing Shen. *An Asynchronous Variational Integrator for Contact Problems Involving Elastoplastic Solids.* Acta Mech. Solida Sin. (2024).

Research Experience

Johns Hopkins University, Department of Computer Science MD, U.S.
*Research Assistant, Professor **Ilya Shpitser**: Causal inference, missing data, graphical models Oct 2022 – Mar 2024*

Project 1: Missing Data

- Introduced a **data fusion approach** in settings where data is missing not at random **(MNAR)**, but an auxiliary data that is missing at random (MAR) is available.
- Provided **identification** results under two complementary models , as well as an inverse probability weighting **(IPW)** estimator for the identified parameter, outcome mean in primary domain, in each model.
- Developed **semi-parametrically efficient estimator**. Accepted by **ICML 2024**.

Project 2: Marginal Structural Models in Cardiac Surgery

- Addressed **statistical modeling** of clinical outcomes when diagnostic testing for the underlying disease condition is infrequent and often missing in practice.
- Leveraged **marginal structural modeling** to investigate optimal testing strategy. Conducted sensitivity analyses to test the robustness of our results to model assumptions: **Markov assumption; Unmeasured confounding**.
- Collaborated with the Johns Hopkins Cardiac Surgery Research Laboratory.

Project 3: Causal Inference Package Development: Ananke

- Constructed modules for finding the **optimal backdoor adjustment set** in a general DAG, which yields non-parametric estimators of the interventional mean with the smallest asymptotic variance

Johns Hopkins University, Department of Biostatistics MD, U.S.

*Research Assistant, Professor **Hongkai Ji**: Data Science, Computational Biology, Genomics* *Jul 2023 – Mar 2024*

Project 1: Pipeline Development of NextGen CUT&Tag

- Implemented **cutting-edge NextGen CUT&Tag workflow (Bash/Snakemake)** on HPC Linux server: demultiplexing fastq based on DNA barcodes alignment using Bowtie2, and peak calling by SEACR and MACS2.
- Perfromed differential analysis, explored peak patterns on the UCSC genome browser, and generated TSS/TES plots and peak attribution for comprehensive data analysis.
- **Derived methodology for technology-specific data pattern**: novel peak calling method based on fragment decomposition using fragment length, connected target pairs to genomic bins for both validated technology and further biological insights; updated method of enrichment heatmaps based on different sets of peak performance.
- Collaborated with Dr. Heng Zhu's lab, pharmacology, molecular sciences and oncology, Johns Hopkins Medicine

Project 2: Deep Learning Approach to DNA Base Prediction

- Derived **tissue-specific DNA pattern analysis** based on Expression quantitative trait loci (eQTL) data. Utilized **deep learning base method ChromBPNet** for prediction of ATAC-Seq data, focusing on predicting base-resolution chromatin patterns through modifications on DNA sequences.
- Re-modelized novel deep learning approaches, within the current, aiming to model the prediction from DNA sequence to ATAC-Seq data comprehensively and ultimately to gene expression. Provided insights that contribute to the advancement of precision medicine.

Yale University, Department of Biostatistics Remote

Undergraduate Research Assistant, Professor Hongyu Zhao: Statistical genomics *Mar 2021 - Nov 2021*

- Estimated the genetic correlation and causal effect between thousands of diseases and COVID-19 infection, hospitalization, and severity, using GNOVA and several **Mendelian Randomization** models
- Designed an SVD-based model to identify major SNP clusters in the framework of Mendelian Randomization; found multiple pathways between several distinct causal mechanisms.
- Provided GWAS of doctor-diagnosed asthma and blood eosinophil counts in the UK Biobank; ran heritability estimation and genetic correlation using BOLT; Performed conditional analysis using GCTA on nearby signals.

SJTU-YALE Joint Center for Biostatistics and Data Science Shanghai, China

Undergraduate Research Assistant, Professor Lin Liu: Statistics *Apr 2021 - Mar 2022*

- Investigated and analyzed the framework for addressing unmeasured confounding concerning Individualized Causal Effects and **Structural Nested Mean Models (SNMMs).**
- Conducted an extensive **review of the book** titled "Semiparametric Regression with R"; contributed paper writing that has been accepted by the Journal of the American Statistical Association (JASA).

Shanghai Jiao Tong University, UM-SJTU Joint Institute Shanghai, China

Undergraduate Research Assistant, Professor Yongxing Shen: Computational Mechanics *May 2020 - Mar 2021*

- Investigated the application of the **asynchronous variational integration (AVI) method** within the context of dynamical systems, with a focus on its utility in drop test simulations. This endeavor aimed to expedite numerical calculations associated with drop test simulations and enhance the overall accuracy of simulation outcomes.
- **Developed a spacetime discretization framework** and derived corresponding equations for the simulation of elastic and plastic solids. Implemented this algorithm using **MATLAB**, accounting for the diverse critical time steps inherent to various materials. Examined the influence of collision angles and vertices on simulation outcomes.

Working Experience

IQVIA, Global Service Department Shanghai, China

Intern *May. 2022 - Jul. 2022*

- Conducted **interviews with doctors nationwide** to gauge opinions on the current diagnosis and treatment of cirrhosis of the liver. Analyzed the current medicine market, emphasizing pricing, risk evaluation, and qualitative data from interviews

Additional Information

- Skilled in Linux, R, Python, MATLAB, C++