Xingjian Ma

xingjian.ma@uconn.edu
860-634-1048

Education

University of Connecticut **Sep 2024 – Present**
PhD Student in Statistics, Department of Statistics *Storrs, CT, U.S.*

- *Courses: Design of Experiments, Statistical Inference, Linear Models*

Renmin University of China **Sep 2022 – Jun 2024**
Master of Science in Applied Statistics, Institute of Statistics and Big Data *Beijing, China*

- *Courses: Clinical Trial Statistical Methodology, Causal Inference, Survival Analysis*

Xi'an Jiaotong University **Sep 2018 – Jun 2022**
Bachelor of Economics in Finance, School of Economics and Finance *Xi'an, China*

- *Courses: Probability Theory, Mathematical Statistics, Econometrics*

Research Experience

Department of Statistics, Uconn **Sep 2025 – Present**
Dissertation Research, Advised by Prof. Haiying Wang and Prof. Jun Yan *Storrs, CT, U.S.*

Topic: Rerandomization Methods for Experimental Design

- Conducting a comprehensive review of rerandomization methodologies, synthesizing core theoretical developments, and identifying open research questions in the field.
- Investigating the adverse impact of high-dimensional covariates on rerandomization, including strategies to mitigate high-dimensional imbalance and analysis of the associated statistical properties.
- Developing principled variable selection strategies under rerandomization to identify influential covariates while maintaining design validity and statistical efficiency.
- Building a high-performance R package for rerandomization-based experimental design and analysis, incorporating recent theoretical advances with scalable and user-friendly implementations.

Department of Public Health Sciences, UConn Health **Sep 2025 – Present**
Research Assistant *Farmington, CT, U.S.*

Topic: Longitudinal Analysis of Depression Outcomes in Large-Scale Medicare Claims Data

- Processed multi-billion Medicare claims datasets using high-performance big-data tools in R, including `fst` and `data.table`, enabling efficient extraction and linkage of depression patient cohorts across multiple years.
- Constructed a longitudinal cohort of Medicare beneficiaries with depression by integrating claims histories, chronic condition indicators, and demographic covariates.
- Applied survival analysis methods to identify key clinical and demographic factors associated with the onset and progression of depression among Medicare patients.

Center for Population Health, UConn Health **Jan 2025 – Jun 2025**
Research Assistant *Farmington, CT, U.S.*

Topic: Survival and Machine Learning Analysis of Suicide Risk Using Clinical Data

- Conducted data mining on patient demographics and clinical records, employing logistic regression and Weibull Cox proportional hazards models to identify key factors associated with suicide rates.
- Implemented survival analysis techniques to address missing data challenges and assess time-to-event outcomes related to suicide risk.
- Employed multiple machine learning techniques, including Random Forest, SVM, and Neural Networks, to predict suicide rates and improve forecasting accuracy.

Connecticut Children's Medical Center **Aug 2024 – Jun 2025**
Research Assistant *Hartford, CT, U.S.*

Topic: Time-Series Forecasting of Emergency Department Patient Volume

- Preprocessed and visualized emergency department census data from Connecticut Children's Medical Center, using heatmap calendar plots to reveal key periodic patterns and temporal trends in patient flow.
- Collected external covariates (temperature, air pollution, water quality) from public APIs, integrated them with census data, and performed data merging for downstream modeling and analysis.
- Utilized Fourier series to construct a harmonic dynamic regression model to estimate different periodic cycles, combined with an ARIMAX framework to model non-periodic short-term variations, achieving effective forecasting performance.

Institute of Geriatric Medicine, Beijing Hospital **Oct 2022 – May 2023**
Research Assistant *Beijing, China*

Topic: Change-Point Analysis of Physiological in in Aging Populations

- Streamlined and preprocessed population data from Beijing Hospital, and utilized line plots and cubic spline regression plots to identify significant change points in various physiological function variables.
- Estimated knots in linear spline models using the Newton-Raphson algorithm, specifically tailored to capture critical shifts related to the aging process.
- Developed advanced plotting functions to visually validate results and engaged in thorough consultations with hospital doctors, ensuring a robust integration of statistical findings with clinical experiences.

Institute of Statistics and Big Data, RUC **Mar 2023 – Jun 2024**
Dissertation Research, Advised by Prof. Yang Liu *Beijing, China*

Topic: Covariate-Adaptive Randomization for Multi-Arm Experiment with Unequal Allocations

- Developed improved covariate-adaptive randomization procedures for multi-arm clinical trials, proposing novel randomization schemes that enhance covariate balance while preserving randomization integrity.
- Studied theoretical properties of CAR procedures by deriving asymptotic distributions of covariate imbalances under general multi-arm designs with unequal allocation ratios.

Teaching Experience

Renmin University of China, Institute of Statistics and Big Data **Sep 2023 – Jan 2024**
Teaching Assistant for Course Clinical Trial Methodology *Beijing, China*

Publications

- Ma, X., Aseltine, R. H., Chiam, T., deMayo, R., Xie, Y., & Yan, J. Forecasting Emergency Department Census: Leveraging Periodic Patterns and External Covariates.(*Submitted to The American Journal of Emergency Medicine*)
- Xie, Y., Aseltine, R. H., Chiam, T., deMayo, R., Huang, Y., Ma, X., & Yan, J. (2025). Modeling Emergency Department Volume-at-Risk Using a Discrete Extreme Value Approach. (*Submitted to The Journal of the Royal Statistical Society: Series C*)

Projects

Parkinson's Disease Progression Prediction | *Python* **May 2023 – Jun 2023**

- Engineered time-aware clinical features from high-dimensional proteomic data, reducing 1,200+ variables to 25 informative predictors, enhancing model interpretability and performance.
- Developed a hybrid model combining LightGBM (multi-class classification) and a neural network and applied strict time-series cross-validation to ensure robust generalization on future patient visits.

Covariate-Adjusted Randomization in Group Sequential Design | *R, Rcpp* **Feb 2023 – Dec 2023**

- Incorporated covariates in a multivariate regression analysis to aid in estimating the variance of treatment effects, further optimizing the efficiency of interim analysis and reducing the expected sample size.
- Conducted numerical simulation experiments to compare the test efficiency of clinical trials under different α spending functions and covariate regression methods to obtain the optimal design plan.

Deep Reinforcement Learning | *Python (PyTorch)* **Oct 2022 – Dec 2022**

- Developed AI-driven algorithms employing CNN architectures (e.g., AlexNet, VggNet, ResNet) for enhanced interactive game image analysis, incorporating data augmentation and transfer learning techniques to boost model robustness.
- Utilized reinforcement learning strategies, specifically the Q-learning algorithm, to continuously train and refine the model based on dynamic game scenarios, resulting in improved decision-making and adaptability.

Honors & Awards

- Outstanding Graduate of Renmin University of China **Apr 2024**
- Graduate Excellent First Class Scholarship, Renmin University of China **Oct 2023**
- Honorary Title of Outstanding Student Cadre, Renmin University of China **Oct 2023**

Technical Skills

Programming: R (dplyr, tidyverse, ggplot2); Python (Numpy, Pandas, PyTorch); Rcpp; SAS; Git; Linux
Languages: Chinese(Native); English(Fluent - TOEFL: 100 (R28 L26 S22 W24))