Zhengwei Song

Tel: (332) 256-6895 | Email: zs2539@cumc.columbia.edu | Personal Website

EDUCATION

Columbia University Mailman School of Public Health, New York, NY, USA May 2024 (Expected)

Master of Science in Biostatistics (GPA: 3.96)

- Relevant Coursework: (Advanced) Probability, Biostatistical Methods, Epidemiology, Data Sciences, Survival Analysis, Randomized Clinical Trial, Data Mining, Advanced Statistical Computing

University of Manchester, Manchester, UK Jul 2021

Bachelor of Science in Mathematics and Statistics (GPA: 3.65, 1st in class)

- Relevant Coursework: Real Analysis, Statistical Inference, Markov Chain, Martingales

Shandong University, Jinan, China Jun 2021

Bachelor of Science in Mathematics (GPA: 3.62, top 30%)

- Relevant Coursework: Advanced (Linear) Algebra, Calculus, Geometry

WORK EXPERIENCES

Biostatistics Intern, Medical Scientific Affairs Dept, Roche Diagnostics (Shanghai) Apr 2022 – Sep 2022

- Developed statistical methods for analyzing clinical trial data, including the development of novel approaches to address specific research questions and issues with existing methodologies
- Collaborated with 10 medical team members and provided statistical support in phase 4 clinical trials by co-developing analytical plans, performing analyses, interpreting results, and summarizing findings into concise reports that are understandable to non-statisticians
- Co-developed an R package (jmpost) of linear mixed effects models for the tumor size over time by Bayesian inference using Hamiltonian Monte Carlo method

Data Analyst Intern, Information System Dept, Sina Weibo (Beijing) Oct 2021 – Apr 2022

- Scraped & wrangled user data, and created visualization (user portraits) for rankings in the entertainment operations, and presented final statistics for several popular TV series, variety shows, and documentary, to provide data support for social media influencers and internal operations
- Maintained Hive SQL and data warehouse services

RESEARCH EXPERIENCES

Multi-omics Clustering and Mediation Analysis in Alzheimer's Disease Mar 2023 – Present

Annie Lee Lab, Dept of Neurology, Columbia University

- Detect Alzheimer's pathologies-related patient subgroups and ageing population subgroups from conducting the latest cluster analyses (SNF, CCA, etc.) from five omics data across various cohorts, to yield molecular information leading to precision interventions
- Identify disease-related novel loci from the adaptive gene-environment interaction test, to discover trait genes
- Implement mediation analysis among various pathologies (vascular, neurodegenerative, etc.), gene expression, and Alzheimer's Disease
- Draft manuscript on mediation analysis which is under consideration for publication

Black-Scholes Pricing Model Data Simulation by Multilevel Monte-Carlo Method Sep 2020 – Jun 2021

Kody Law Lab, Dept of Mathematics, University of Manchester, UK

- Utilized the Euler-Maruyama method and stochastic differential equations to generate thousands of simulated asset price paths for the calculation and analysis of European option payoffs and present values.
- Implemented a cutting-edge Multi-Level Monte Carlo method for efficient option pricing under the Black-

Scholes model, significantly reducing computational overhead while maintaining accuracy, providing quantitative support for risk management and strategic decision-making

ADDITIONAL EXPERIENCE

Teaching Assistant, Columbia University (New York) Sep 2023 – Dec 2023 (Expected)

- Hold weekly office hours, grade homework, attend the instructor's flipped class for 25 students Q&A for Mathematical Statistics and 20-30 students for Biostatistical Methods

Edible Tableware based on Finite Element Analysis Apr 2019 – Apr 2020

Song Yu Lab, School of Mathematics, Shandong University, China

- Led a team of five with diverse academic backgrounds and secured full funding (around $900)
- Designed and produced a chopstick-like mold by SolidWorks software according to finite element analysis theories
- Connected and partnered with local restaurants and bars for testing mechanical characteristics

PUBLICATIONS

"Genome-wide gene-based study in multi-ethnic cohorts identifies genes that interact with vascular risk factors in Alzheimer's disease (AD)" (with Annie Lee), under review for *Alzheimer's disease and Parkinson's disease Conference,* 2024.

SKILLS

Computer: R (tidyverse, caret, survival, lme4, gee, httr, bioconductor, etc.), SAS, Shiny, SQL, C, MATLAB, Unix, Microsoft Office, AutoCAD

Tests: t, z, ANOVA, chi-squared, Fisher's exact, McNemar's, Log-rank, sign, Wilcoxon signed-rank & rank-sum, etc.

Modeling / Machine Learning: linear, generalized linear (logistic, Poisson), weighted least squares, mixed effect, GEE, survival (Cox, Stratified PH, AFT), decision tree, random forest, boosting, K-NN, cubic splines, local regression, GAM, MARS, LDA, QDA, NB, SVM, clustering (K-means, Hierarchical, spectral, etc.), PCA, LASSO, Elastic net, Ridge, PCR, PLS, cross validations

Simulation & Optimization: Data generation, Newton-Raphson, EM, bootstrapping, MCMC

AWARDS

AAAI 2022 Security AI Challenger VIII Award (44 out of 3692) Jan 2022

Enactus Annual Outstanding Individual (Top 10%) Dec 2018

Shandong University Third-level Academic Scholarship, 2017-2018 (Top 30%) Oct 2018

EXTRACURRICULAR ACTIVITIES

Treasurer, Enactus, Shandong University Oct 2018 – Oct 2019

- Tutored students in budget planning for every program
- Responsible for the entire budget arrangement of the organization

Secretary, Association of International Exchange, Shandong University Oct 2017 – Oct 2018

- Assisted in events and activities for the international student exchange
- Trained students to avoid and bridge the gap of culture shock