Jiangjie Zhou,

Department of Biostatistics, Peking University, Beijing, China
actionsafe@pku.edu.cn
+86 17778316315

Education

2022 – 2025 **M.Phi. Biostatistics, Peking University** , Beijing, China.

2018 – 2022 **B.Med. Preventive Medicine, Peking University** , Beijing, China.
Courses: Anatomy, Physiology, Biochemistry, Pathology, Clinical Medicine, Medical Statistics …
GPA: 3.5/4.0 (85/100), top 30%

Minor: Mathematics and Applied Mathematics, Peking University , Beijing, China.
Courses: Mathematical Analysis, Advanced Algebra, Real Analysis, Complex Analysis, Probability Theory, Applied Stochastic Process, Bayesian Inference, Numerical Optimization, Deep Learning …

Honors

2020-2021 Jiang Zehan Cup Mathematical Modeling Competition, 3rd prize
Award for Scientific Research, Peking University
Chugai scholarship for excellent medical students
Share Cup Innovation Competition, excellence prize

2019-2020 Award for Scientific Research, Peking University
Award for Future Medical Star, Peking University
Eisai pharmaceutical scholarship

2018-2019 Award for Academic Excellents, Peking University

Research Interests

Statistics: Recurrent Event Analysis, Bayesian Inference, Mixture Model, Causal Inference.

ML/AI: Deep Learning (e.g., Variational Autoencoder, Diffusion Model, Large Language Model), Reinforcement Learning, Artificial General Intelligence.

Research Experience

2024.04-2024.07 **Application of Deep Learning in Interval-Censored Data Analysis (revision and model implementation)**
Developed a neural network with monotonic splines for baseline cumulative hazard function modelling. Modeled nonlinear relationships between covariates and survival outcome with deep neural networks. Evaluated variable importance with SHAP. Implemented with Python and R, and created a web-based application for data analysis.

Research Experience (continued)

2023.12-2024.04 **Adverse Drug Reaction (ADR) Risk Prediction Based on Longitudinal Data (model implementation and conducting simulation studies)**
Predicted drug adverse reaction risks based on time-dependent covariates, generated simulated data based on real-world application scenarios, and evaluated multiple methods, including joint models, landmark models, and dynamic Bayesian networks, in the simulated data. Selected the best-performing model for application in real data analysis.

2023.07-2023.12 **Estimation of Rare Disease Incidence Based on Health Insurance Databases with Missing Data (model implementation and conducting simulation studies)**
Generated simulated data based on real-world application scenarios and evaluated multiple methods, including traditional statistical methods, Bayesian framework-based methods, and machine learning-based methods, in the simulated data.

2022.12-2023.06 **The Diagnosis of Malignant Pleural Effusion Based on Stacking Model (co-author, data collection and analysis)**
Studied Stacking models' performance in right-censored data using SHAP for variable importance. Compared ISE (mean squared errors) with Cox, CoxBoost, and Random Forest. Published one paper.

2022.12-2023.03 **Association of Air Pollution Exposure and Breast Cancer Mortality in Inner Mongolia (first-author)**
Cleaned health insurance data with R, matched patients to air pollution exposure levels. Implemented cure models using C++ and R. Developed an interactive interface with JavaScript. Published one paper.

2020.04-2021.04 **Public Sentiment Analysis of COVID-19 Prevention and Control Measures Based on Social Media (co-author, data collection and analysis)**
Scraped over 10 million Weibo posts using Scrapy, stored in MongoDB. Conducted keyword extraction, LDA topic modeling, and sentiment analysis with Python. Visualized results using JavaScript. Published one conference paper.

2020.04-2020.09 **Quality Evaluation of HPV Vaccine-Related Online Messages (co-author, data collection and analysis)**
Scraped articles about HPV vaccines from WeChat using Python, cleaned and tokenized data, generated word clouds, and scored article readability. Published one paper.

2019.11-2020.05 **Application of Machine Learning Methods in Rapid Tumor Screening (first-author)**
Applied machine learning methods (SVM, Random Forests, Neural Networks, Ensemble Learning) to predict tumors using the National Population Health Data Science Center dataset. Tuned models and compared performance. Achieved AUC > 0.9. Published one paper, won third prize at the Population Health Cup.

2018.12 – 2019.06 **Investigation of Service Content and Providers of Online Consultation Websites (first-author)**
Wrote Python scripts to scrape 800,000 doctor profiles from online platforms and store them in MySQL. Cleaned data with Python and SQL, generated Excel reports, and conducted statistical analysis with R. Published one paper, providing foundational data for platform optimization.

Research Publications

Journal Articles

1. **J. Zhou** and B. Liang, "Semiparametric regression for partially observed panel count data," *Submitted*, 2024.
2. **J. Zhou**, B. Liang, M. Su, J. Li, and S. Wang, "Long-term exposures to pm2.5 and its components increase the breast cancer mortality in the region of inner mongolia, china: A retrospective study based on mixture cure model," *BMC Cancer (Submission ID: 596206fa-8cb6-47a2-b21d-5bf295a87072)*, 2024.
3. W. Jiang, **J. Zhou**, and B. Liang, "An improved dunnett's procedure for comparing multiple treatments with a control in the presence of missing observations," *Mathematics*, vol. 11, no. 14, 2023, ISSN: 2227-7390. DOI: 10.3390/math11143233.
4. X. Ou, **J. Zhou**, B. Liang, *et al.*, "Prognostic factors analysis of metastatic recurrence in cervical carcinoma patients treated with definitive radiotherapy: A retrospective study using mixture cure model," *Cancers*, vol. 15, no. 11, 2023, ISSN: 2072-6694. DOI: 10.3390/cancers15112913.
5. J. Wang, **J. Zhou**, H. Wu, Y. Chen, and B. Liang, "The diagnosis of malignant pleural effusion using tumor-marker combinations: A cost-effectiveness analysis based on a stacking model," *Diagnostics*, vol. 13, no. 19, 2023, ISSN: 2075-4418. DOI: 10.3390/diagnostics13193136.
6. **J. Zhou** and S. F. Wang, "Introduction of landmarking approach and its application in dynamic prediction," *Chinese Journal of Epidemiology*, vol. 43, 2022.
7. **J. Zhou**, W. Zhang, S. F. Wang, and L. M. Li, "A study on the online medical consulting websites based on personal computers," *Chinese Journal of Epidemiology*, vol. 42, 2021.
8. W. Z. Wang, J. L. Lyu, M. T. Li, *et al.*, "Quality evaluation of hpv vaccine-related online messages in china: A cross-sectional study," *Human Vaccines Immunotherapeutics*, vol. 17, 2020.
9. **J. Zhou**, S. F. Wang, and L. M. Li, "Application of python web crawler technology in infodemiology," *Chinese Journal of Epidemiology*, vol. 41, 2020.

Books and Chapters

1. L. X. Zhang and H. B. Han, Eds., *Introduction to Health Data Science*. Beijing, China: Peking University Medical Press, 2022, ISBN: 978-756-5926-61-7.

Employment

2019.12 – 2020.2 **Frontend Dev (intern)** Meinian Da Jiankang(Group)Co.,Ltd., Beijing, China.

2024.08 – 2024.09 **Teaching Fellow (part-time)** GEC Academic, Beijing, China.

Skills

Languages TOEFL(Test Date: 2024/07/02): 102/120 (Reading: 29, Listening: 27, Speaking: 22, Writing: 24); CET-6

Coding C/C++, Java, Python, R, JavaScript, LaTeX, …