ZHEJIA DONG

zhejia_dong@brown.edu | +1 401 263 5252 | 121 South Main Street, Providence, Rhode Island 02903, USA

EDUCATION

Brown University **Providence, USA**
Master of Science, Biostatistics *Sept. 2021 – June 2023*

- GPA 4.0/4.0
- Coursework: Linear Models, Statistical Inference I & II, Statistical programming with R, Bayesian Statistical Methods, Generalized Linear Model

Xi'an Jiaotong University (XJTU, C9 League) **Xi'an, China**
Bachelor of Science with Honors, Mathematics and Applied Mathematics *Sept. 2017 – July 2021*

- Thesis：Robust estimation in linear regression with response missing at random (Best undergraduate thesis award, 1/121)
  Advisor: Prof. Liya Fu
- Overall GPA: 88.2/100 | Major GPA: 90.1/100 | Rank: 1/33
- Coursework: Functions of Real Variables (100), Stochastics Process (98), Advanced Algebra and Analytic Geometry (96), Probability (96), Functional Analysis (96), Mathematical Modeling (96), Mathematical Statistics (96), Data Analysis & Statistical Software (95)
- Honors: Outstanding Graduates, Merit-based Scholarship Recipient (2018, 2019, 2020), Outstanding Student (2018, 2019, 2020)

RESEARCH EXPERIENCE

Bias Amplification and Cancellation of Unmeasured Confounders **Providence, U.S.**
Research Assistant, Prof. Youjin Lee, Dpt. of Biostatistics, Brown University *Nov. 2021 – Present*

- Identified the instrumental variable is a bias amplifier in a new data structure that the measured confounder is associated with the instrumental variable and unmeasured confounders exist.
- Proved the analytical results of the amplified bias by including the instrumental variable in the regression model, performed simulation, and showed the numerical results are consistent with promised bias by the formula.

Robust estimation in linear regression with response missing at random **Xi'an, China**
Bachelor Thesis, Research Assistant, Prof. Liya Fu, Dpt. of Statistics, XJTU *Oct. 2020 – Aug. 2021*

- Proposed a new covariate balancing propensity score-based estimator for linear regression coefficients and population mean, proved the asymptotical consistency of the proposed estimator
- Derived the estimated selected probability by covariates balancing propensity score method with Robit model as the assumed selected probability model
- Formulated the imputed M estimator by the augment inverse weighted imputation with the estimated selected probability and lnch function as the objective function
- Performed numerical simulation and compared the new M-estimation approach with existing approaches like the maximum likelihood estimation, the M-estimation under the complete case, the weighted M-estimation, by the limited sample performance when the data obeys different distributions through Bias and RMSE, demonstrated our method is superior to the other estimators

Deep Reinforcement Learning Theories and Improvements of Two Algorithms **(Remote)**
Research Assistant, Prof. Pietro Liò, Dpt. of Computer Science and Technology, University of Cambridge *July 2020 – Oct. 2020*

- Studied Deep Reinforcement Learning and Graph Neural Networks (GNN) theory, and implemented Deep Q Networks (DQN) algorithm, Deep Deterministic Policy Gradient (DDPG) algorithm, Twin Delayed Deep Deterministic Policy Gradient (TD3) algorithm in gym environment 'Pendulum-v0' by Python and Tensorflow
- Improved DDPG algorithm, in which the max steps per episode were determined by a specific gym environment with a discount factor set to correspond to the max steps; the algorithm was able to solve the task after 157 episodes of training in LunarLanderContinuous-v by soft and delay updates
- Proposed an algorithm based on the TD3 algorithm by introducing the Ornstein-Uhlenbeck process as the off-policy exploration strategy to generate noise for each action
- Evaluated our algorithm and the original TD3 algorithm in four gym environments by comparing the corresponding average rewards gain during training and testing episodes and determined that our algorithm performs better in MountainCarContinuous-v0, and similarly in Pendulum-v0, and BipedalWalker-v3 than the original TD3 algorithm

Analysis of M-estimation Objective Functions **Xi'an, China**

Research Assistant, Prof. Liya Fu, Dpt. of Statistics, XJTU *Nov. 2019 – June 2020*

- Studied multiple M-estimation methods for errors following normal, t, Cauchy, mixture distributions, and high leverage cases; compared the performance of different objective functions such as Huber, Hampel, Welsch to estimate the parameters of interest
- Implemented a data-driven procedure automatically adjusts the value of the tuning constant to provide robustness against outliers
- Adapted an algorithm for multiple tuning parameter settings that can be applied to estimating objective functions
- Performed numerical simulation where the multiple tuning parameters data-driven algorithm was employed on different estimating objective functions; evaluated performance through bias and MSE
- Applied methods to real datasets, further compared the performance of different methods, and concluded that Welsch and Tukey functions and Huber with data-driven approach have the overall best performance among most cases and are consistent for different sample sizes

INTERNSHIP EXPERIENCE

Tencent Technology (Beijing) Co., Ltd **Beijing, China**

Data Analyst Intern, Data Center *May – July 2020*

- Obtained the number of thumbs, fans, answers, articles, and other data of Tencent's 55 official accounts on Zhihu by Python, R and Excel; conducted word segmentation and word cloud analysis for all articles and answers (~8.6 million words) for data visualization
- Conducted the Spearman Test for correlation of the rare word percent with the popularity of the article and verified the hypothesis that easier-to-understand articles and answers were more popular in Zhihu, and accomplished a completed data analysis report suggesting optimizing proposals for the operations department
- Assisted the Tencent medical team to perform data analysis for COVID-19, obtained data from national and provincial CDC, and drew real-time maps, bar charts, and pie charts to visualize the number of confirmed, suspected, dead, and recovered cases
- Supported the Timi Studio Group of King of Glory in conducting classification research on 100 heroes of the game; applied Random Forest and High Correlation Filter to reduce the dimensionality of datasets and used GMM and K-means to conduct cluster analysis on heroes to guide subsequent character updates

EXTRACURRICULAR ACTIVITIES

***Founder and Leader*, Living Water Welfare Organization, Xi'an & Wuhan, China** *Feb. – Mar. 2020*

- Launched a public welfare organization and donated medical supplies to the township hospitals in Hubei province at the outbreak of the COVID-19
- Collected 92,976 RMB in total, and bought 9 batches of goggles, alcohol, protective suits, masks, and other medical supplies, which were donated to 23 township hospitals, including Huarong Health Center in Ezhou City

***Teaching Assistant*, Academic Counselling Centre of Nanyang College, XJTU** *Oct. 2018 – June 2021*

- Held office hours for courses such as Calculus, Linear Algebra, Probability and Statistics
- Tutored one-to-one 16 students in math with an average improvement of 15 points; provided 6 pre-test lectures

***Department Manager*, Yingzai Volunteer Club, XJTU** *Sept. 2017 – June 2018*

- Organized campus tours for more than 800 junior and senior high school students from less developed areas to visit the campus.
- Taught History and Geography courses in Yehu Junior Middle School of Lantian County every Saturday, and held various extracurricular activities, such as football matches.

SKILLS

Languages: English (Proficient), Chinese (Native)

Computer Skills: R, Python, MATLAB, Stata, SPSS, Maple, VB.NET , C++, Tableau, LaTeX