

Jin Zhu

(86+) 13570919327 | zhu37@mail2.sysu.edu.cn | Website | Github

EDUCATION

Sun Yat-Sen University

08/2019 – Present

School of Mathematics, supervised by Xueqin Wang

Ph.D. of Statistics

- Research Interests: statistical machine learning, non-Euclidean data analysis, reinforcement learning, applied machine learning, statistical software.

Sun Yat-Sen University

08/2016 – 06/2019

School of Mathematics, supervised by Xueqin Wang

Master of Statistics

Sun Yat-Sen University

09/2012 – 06/2016

School of Mathematics

Bachelor of Statistics

PUBLICATIONS AND PREPRINTS

Publications

- Yanhang Zhang, Junxian Zhu, **Jin Zhu**, Xueqin Wang (2022+). A Splicing Approach to Best Subset of Groups Selection. *INFORMS Journal on Computing* (Accepted). (Co-first author)
- Chengchun Shi, **Jin Zhu**, Ye Shen, Shikai Luo, Rui Song, Hongtu Zhu (2022+). Off-Policy Confidence Interval Estimation with Confounded Markov Decision Process. *Journal of the American Statistical Association* (Accepted).
- **Jin Zhu**, Xueqin Wang, Liyuan Hu, Junhao Huang, Kangkang Jiang, Yanhang Zhang, Shiyun Lin, Junxian Zhu (2022). abess: A Fast Best-subset Selection Library in Python and R. *Journal of Machine Learning Research* 23, no. 202 (2022): 1-7.
- **Jin Zhu**, Wangwei Wu, Yuting Zhang, Shiyun Lin, Yukang Jiang, Ruixian Liu, Heping Zhang, and Xueqin Wang. Computational analysis of pathological image enables interpretable prediction for microsatellite instability. *Frontiers in Oncology* 12 (2022).
- **Jin Zhu**, Wenliang Pan, Wei Zheng, Xueqin Wang (2021). Ball: An R Package for Detecting Distribution Difference and Association in Metric Spaces. *Journal of Statistical Software*, 97(6), 1–31.
- Junxian Zhu, Canhong Wen, **Jin Zhu**, Heping Zhang, Xueqin Wang (2020). A Polynomial Algorithm for Best-subset Selection Problem. *Proceedings of the National Academy of Sciences*, 117 (52) 33117-33123.
- Wenliang Pan, Xueqin Wang, Heping Zhang, Hongtu Zhu, **Jin Zhu** (2020). Ball Covariance: A Generic Measure of Dependence in Banach Space. *Journal of the American Statistical Association*, 115:529, 307-317. (Co-first author)
- **Jin Zhu**, Kunsheng Lv, Aijun Zhang, Wenliang Pan, Xueqin Wang (2019). Two-sample Test for Compositional Data with Ball Divergence. *Statistics and its Interface*, 12(2), 275-282.
- Yukang Jiang, Jianying Pan, Ming Yuan, Yanhe Shen, **Jin Zhu**, Yishen Wang, Yewei Li, Ke Zhang, Qingyun Yu, Huirui Xie, Huiting Li, Xueqin Wang, Yan Luo (2021). Segmentation of Laser Marks of Diabetic Retinopathy in the Fundus Photographs Using Lightweight U-Net. *Journal of Diabete Research*. doi:10.1155/2021/8766517.
- Minqiong Chen, Ting Tian, Wenliang Pan, Xueqin Wang, and **Jin Zhu** (2022). Paired-sample Tests for Homogeneity With/without Confounding Variables. *Statistics and Its Interface*, 15(3), 335-348.

Papers under Review/Revisions

- Xueqin Wang, **Jin Zhu**, Wenliang Pan, Junhao Zhu, Heping Zhang (2021). Nonparametric Statistical Inference via Metric Distribution Function in Metric Spaces. arXiv:2107.07317.
- Zhe Gao, **Jin Zhu**, Wenliang Pan, Yue Hu, Xueqin Wang (2022). A Conditional Strong Independence Screening for Non-Euclidean Responses.

SOFTWARES

- abess: Fast Best-Subset Selection (in PyPI and R CRAN).
- Ball: Statistical Inference and Sure Independence Screening via Ball Statistics (in PyPI and R CRAN).
- cdcsis: Conditional Distance Correlation Based Feature Screening and Stastical Inference (in R CRAN).
- robustlm: Robust Variable Selection with Exponential Squared Loss (in R CRAN).

INVITED TALKS

- A Fast Best-Subset Selection Library. Southwest Jiaotong University Doctoral Forum, 2021.
- Ball: An R Package for Detecting Distribution Difference and Association in Metric Spaces. Young Statisticians Forum, 2020.

PARTICIPATED GRANTS

- Outstanding Graduate Student Innovation and Development Program of Sun Yat-Sen University (Co-PI), 2019-2020. “Conditional Independence in Metric Spaces”.
- National Nature Science Foundation of China (Participated), 2018-2021. “Conditional Independence and Its Application”.
- The Science and Technology Program of Guangzhou (Participated), 2020-2022. “Statistical Inference in Metric Spaces”.

INTERNSHIP EXPERIENCE

Cutting-edge algorithm engineer 06/2020 – 06/2021

Didi artificial intelligence labs (part time, remote internship). Mentors: Hongtu Zhu and Shikai Luo. Guangzhou, China

- Evaluate commercial strategies with the data collected by another strategy, known as an off-policy evaluation task. Implement an evaluation method that enjoys semi-parametric statistical efficiency even when unobserved confounders impact the strategy. Provide a Wald-type statistic to test the value difference between strategies.
- Implement a synthetic difference in differences (SDID) method for assessing the average treatment effect. Conduct simulation studies to validate the advantage of SDID by comparing the difference in difference method and synthetic control method. Wrap a Python library to facilitate the online deployment.

Recommendation algorithm engineer 06/2018 – 08/2018

Alibaba Cloud (full time). Mentor: Xiaoguang Wang Hangzhou, China

- Recommend technology news to the users of the Alibaba-cloud App. Evaluate the matching scores between news reports and users by a classification model with features extracted from the user and news.
- Deploy the whole procedure into a recommendation platform, and make it automatically update by week.

Big-data algorithm engineer 07/2017 – 01/2018

Zuzuche (part time) Guangzhou, China

- Develop algorithm to smooth the altitude data and detect the abnormal drifting altitude data based on the statistical methods for time series analysis.
- Quantify the possibility of a vehicle on the viaduct from the perspective of height and trend. From the height perspective, use a mixture model to learn the altitude distribution of viaduct and flat land; and use the Kendall-tau statistic to assess the possibility of an upward/downward trend in altitude.

AWARDS AND HONORS

- 2021 Guangdong Industrial Intelligent Manufacturing Innovation Competition – Intelligent Algorithm Competition (Ranked 4% out of more than 4000 teams).
- Best Student Paper Awards, the 4th National Statistical Doctoral Forum, 2020.
- National Scholarship, China, 2020.
- Outstanding Graduate Students of Guangdong Province, 2019.
- Best Students Paper Awards, Guangdong Mathematical Society, 2019.
- Youchu Zhong Specialized Awards (Outstanding Student Award), 2018.
- Outstanding Undergraduates of Sun Yat-Sen University, 2016.
- First-class Scholarship for Outstanding Students of SYSU, 2014 and 2015.
- National Encouragement Scholarship, China, 2013, 2014, and 2015.
- Honorable Winner in 2014 Mathematical Contest in Modelling, 2014.

MISCELLANEOUS

Service	Referee for Journal of the American Statistical Association, R Journal.
Languages	R, Python, C++, SQL, L ^A T _E X.
Communication	Chinese, English.