

Zhili Qiao

☎ 515-817-3887 | ✉ zlqiao@iastate.edu | in Zhili Qiao

EDUCATION

- Iowa State University** 08/2019 - present
- Doctorate of Philosophy in Statistics, GPA: 4.0/4.0
 - **Research Interest:** machine learning, high-dimensional data analysis, clustering
- University of California, Davis** 09/2017 - 12/2018
- Master of Science in Statistics, GPA: 3.89/4.0
- Wuhan University** 09/2013 - 06/2017
- Bachelor of Arts in Economics & Bachelor of Science in Mathematics, GPA: 3.69/4.0
 - “Hongyi” Honors Program on Mathematical Economics/Finance, GPA: 3.86/4.0

SCHOLARSHIPS & AWARDS

- Nonclinical Biostatistics Scholarship** 01/2022
American Statistical Association
- George W. Snedecor Award** 08/2020
Most Outstanding Ph.D. Candidate in Statistics
Department of Statistics, ISU

WORKING EXPERIENCE

- Corteva Agriscience, Math and Tech Group** 05/2021 – 08/2021
Internship Johnston, Iowa
- Implemented a new type of equivalence test for safety assessment of genetically modified organisms in **R**.
 - Successfully replicated the simulation results from a paper; extended a method to multiple testing groups.
- Fosun International Limited, Big Data Service Group** 07/2018 – 08/2018
Internship Shanghai, China
- Participated in the construction of an automatic fraud-detection system for loan applicants.
 - My team built a machine learning credit scorecard model, by performing data cleaning and preprocessing, feature engineering, cross validation and model assembly.
- Graduate Student Assistant** 05/2020 – present
Department of Statistics Iowa State University
- Implemented machine learning models to predict plant phenotypes using transcriptome data.

WORKING PAPERS

- Poisson Hurdle Clustering for Sparse Count Data** 06/2020 - 09/2021
- We developed a model-based clustering method for count datasets with high sparsity.
 - A mixture of Poisson Hurdle distribution is utilized to deal with extra sparsity by modeling zero counts separately.
 - We use a combination of group coordinate descent and EM algorithm to iteratively estimate the optimal clusters.
- Dirichlet Multinomial Biclustering for Sparse Compositional Datasets** 07/2021 – present
- We developed a Dirichlet-Multinomial (DM) model-based biclustering algorithm to simultaneously cluster features and samples for highly sparse compositional datasets.
 - A piece-wise constant structure is implemented to deal with the common issue of identifiability that occurs in model-based biclustering methods.
 - Our method outperforms other competing methods both in simulation and on publicly available real datasets.
- Bayesian Feature Selection and Causal Inference for Group Structured Data** 01/2022 – present
- In high-dimensional regression with subgroups of observations, we try to identify differentially expressed features by assigning a group-fused Lasso penalty.
 - We use a fully Bayesian approach to do feature selection and construct confidence intervals for features of interest.

TEACHING EXPERIENCE

Teaching Assistant

- *2020 fall*: Statistical Methods I (Ph.D. core)
- *2020 spr, 2021 spr*: Introduction to Business Statistics (Lab instructor)
- *2020 spr*: Statistical Methods for Research Workers (graduate, Lab instructor)
- *2019 fall*: Computer Processing of Statistical Data (graduate)
- *2016 fall, 2017 spr*: Mathematical Analysis (undergraduate TA)

TECHNICAL SKILLS

Programming: Proficient with R, Python, SQL; experience with SAS, Matlab

Language: Chinese (native), English