

Qian Zhao

1444 Lederle Graduate Research Center
740 N. Pleasant St
Amherst, MA, 01003-9306
United States of America

Cellphone: (872) 202-2644
Email: qianzhao@umass.edu
Github: <https://github.com/zq00>
Website: <https://qianzhao14.netlify.app/>

Employment

Assistant Professor	Statistics, University of Massachusetts, Amherst	2024-Present
Postdoctoral Scholar	Biomedical Data Science, Stanford University	2022-2024

EDUCATION

Ph.D.	Statistics, Stanford University	2022
M.S.	Statistics, The University of Chicago	2016
B.S.	Physics, Fudan University, Shanghai, China	2014

RESEARCH INTERESTS

High-dimensional statistical inference, Statistical genetics, Nonparametric density estimation, Data science education, Reproducible research workflow, Precision medicine

RESEARCH

Assistant professor, University of Massachusetts, Amherst <i>Model inference under sampling bias</i>	2024-Present
--	--------------

Develop statistical methods to infer coefficients of a linear mixed effect model for a case-control sample; develop statistical methods to analyze effects of genetic variants on phenotypes using self-selected biobank data.

Assessing factors of food vulnerability in the US

Collaborator: Prof. Chaitra Gopalappa

Develop statistical models to predict food insecurity in the US, quantify uncertainty of the estimates, identify social determinants of food insecurity.

Postdoctoral researcher, Stanford University <i>Selecting genetic variants under sampling bias</i> Advisor: Prof. Chiara Sabatti, Collaborator: Susan Service	2022-2024
--	-----------

Developed a tilted knockoff method to identify genetic variants associated with endophenotypes to understand distinct and common genetic variants that affect severe mental disorders.

Graduate researcher, Stanford University

Inferring coefficients of a high-dimensional Generalized Linear Model (GLM) 2017-2022

Advisor: Prof. Emmanuel Candès

Developed theories and methods to estimate the distribution of the maximum likelihood estimates (MLE) of a high-dimensional GLM as the number of variables grows with the number of observations.

Constructing a moderate- to high-dimensional histogram 2017-2023

Advisor: Prof. Guenther Walther

We developed a Beta-tree histogram for moderate- to high-dimensional data and provided confidence intervals for the average density in each region.

Technical Mentor, Data Science for Social Good, Stanford University

Identifying behavioral health conditions from police records Summer 2022

Advisor: Dr. Balasubramanian Narasimhan

We developed a data analysis pipeline for journalists from the California Reporting Project to identify police records that mention behavioral health conditions and output relevant information.

Operationalizing equity tiebreaker in San Francisco student school assignment Summer 2021

Advisor: Dr. Balasubramanian Narasimhan and Prof. Irene Lo

We developed an equity tiebreaker and evaluated its effect on enhancing equity in elementary student assignment. We presented our work at the Ad hoc committee meeting on student assignment at San Francisco Unified School District.

Fellow, Data Science for Social Good, Stanford University

Forecasting platelet blood bag demand to reduce inventory wastage at the Stanford Blood Center Summer 2019

Advisor: Prof. Chiara Sabatti and Dr. Balasubramanian Narasimhan

We predicted platelet blood bag demand at the Stanford Blood Center by combining surgery information and recurring transfusion data.

PUBLICATIONS

Manuscript and preprints

1. Sabatti, C. & Zhao, Q. (2022 preprint) Near-peer mentoring in data science: Two experiences at Stanford University, [arXiv 2211.08637](https://arxiv.org/abs/2211.08637)

2. **Zhao, Q.** (2022) Growing by Mentoring: A guide for Data Science for Social Good mentors
3. Walther, G. & **Zhao, Q.** (2022 *preprint*) Beta-trees: Multivariate histograms with confidence statements, [arXiv 2308.00950](https://arxiv.org/abs/2308.00950)

Peer reviewed articles

4. **Zhao, Q.** & Candès, E. (2023). An adaptively resized parametric bootstrap for inference in high-dimensional generalized linear models, *Statistics Sinica*.
5. **Zhao, Q.**, Sur, P. & Candès, E. (2022). The asymptotic distribution of the MLE in high-dimensional logistic models: Arbitrary covariance. *Bernoulli* 28 (3)
6. Zhu, J., **Zhao, Q.**, Katsevich, E. & Sabatti, C. (2019). Exploratory gene ontology analysis with interactive visualization. *Scientific Reports* 9, 7793
7. Orlova, D.Y., Meehan, S., Parks, D., **Zhao, Q.** et al. (2018) QFMatch: Multidimensional flow and mass cytometry samples alignment. *Scientific Reports* 8, 3291

PRESENTATIONS

1. “A resized parametric bootstrap method for inference of a high-dimensional generalized linear model” (Virtual presentation) *7th International Conference on Econometrics and Statistics*, Beijing, China, 2024
2. “Beta-Trees: Multivariate Histograms With Confidence Statements” (Presentation) *The Western North American Region of The International Biometrics Society (WNAR) Annual Meeting*, Fort Collins, CO, 2024; *The 37th New England Statistics Symposium*, Storrs, CT, 2024
3. “Near-peer mentoring in data science --- two experiences at Stanford University” (Presentation) *WNAR Annual Meeting*, Fort Collins, CO, 2024
4. “Controlled variable selection with a biased sample using tilted knockoff” (Presentation) *WNAR Annual Meeting*, Anchorage, AK, 2023
5. “Inferring model parameters of a high-dimensional Generalized Linear Model, with application to genetic studies” (Presentation) *The 36th New England Statistics Symposium*, Boston, MA, 2023
6. “Co-creating learning experience with students in group data science projects” (Poster) *US Conference on Teaching Statistics*, State College, PA, 2023
7. “Selecting genetic variants using knockoffs under collider bias.” (Presentation) *Stanford Causal Science Conference*, Stanford University, Stanford, CA, 2022
8. “Writing a reproducible manuscript in R.” (Lightning talk) *Gear-up for Science Data Symposium*, Stanford University, Stanford, CA, 2022
9. “Forecasting platelet blood bag demand to reduce inventory wastage at the Stanford Blood Center.” (Presentation joint with Guthrie, E. & King, C.) *RStudio Conference*, San Francisco, CA, 2020

TEACHING & MENTORING

Instructor, University of Massachusetts, Amherst

2024-Present

Undergraduate level courses

“Fundamental concepts of statistics”

Spring, 2024

Introductory course in statistical concepts (data visualization, sampling, hypothesis testing, confidence intervals) with hands-on demonstration and activities.

Technical Mentor, Data Science for Social Good, Stanford University

Summer 2021, 2022

- Stanford Data Science for Social Good is an eight-week summer program where fellows tackle data science projects with positive social impact.
- Collaborated with community partners to formulate project goals, approaches, and milestones.
- Facilitated daily meetings where I guided fellows to brainstorm ideas, discuss different approaches, share progress, and provide feedback on each other’s work.
- Advised project work, presentations, and final report, where I asked for clarifying information, answered questions, suggested related research or approaches.
- Held working session with fellows and invited graduate students to speak on related data science topics.
- Met with fellows to discuss individual goals and provide relevant resources.
- Taught training sessions on multiple hypotheses testing, using Git for collaboration, and topic modelling.

Teaching Assistant, Stanford University

2016-2021

Undergraduate level courses

“Biostatistics”

Winter 2021, Fall 2019

Introductory course in statistical methods for biological data (t-test, categorical data analysis, linear regression).

“Statistical methods in engineering and the physical sciences”

Fall 2020, Spring 2018

Introductory course in probability and statistical methods for undergraduates majoring in physical science and engineering.

“Introduction to statistical methods”

Summer 2020

Introductory course in statistics for undergraduate and high school students (data summary and visualization, sampling, hypothesis testing, modelling continuous and categorical relationships).

“Data science 101”

Spring 2020

Undergraduate course in statistics (data visualization, sampling and resampling, linear models, hypothesis tests).

“Riding the data wave”

Fall 2019

Freshman seminar on basic statistical concepts (mean, variation, association) and exploratory data analysis.

Graduate level courses

- “Introduction to statistical learning” Winter 2019
Upper-division undergraduate and masters level course on introductory machine learning methods (regression, clustering, splines).
- “Data mining and analysis” Fall 2016, Summer 2017
Upper-division undergraduate and masters level course applied machine learning methods (regression, clustering, splines, and semiparametric methods).
- “Applied multivariate analysis” Winter 2020
Upper-division undergraduate and masters level course in applied multivariate statistical methods (PCA, clustering, mixture models, EM algorithms).
- “Advanced statistical theory” Spring 2019
Topic class in advanced statistical theory (concentration, high-dimensional PCA, nonparametric methods, compressed sensing, graphical models and message passing algorithm).
- “Modern applied statistics: Learning” Winter 2019
Advanced course in applied machine learning methods (regression, clustering, splines and reproducing kernel Hilbert space).
- “Theory of statistics” Spring 2018
Graduate level course in advanced statistical theory (multiple testing, knockoff, selective inference, estimating a multivariate Normal mean).

AWARDS & FELLOWSHIPS

- ADVANCE Faculty Fellow** 2024-2025
University of Massachusetts, Amherst
- Departmental Teaching Assistant Award** June 2020
Department of Statistics, Stanford University

GRANTS

- Teaching Advancement Grant** Dec 2022
Center for Teaching and Learning, Stanford University

ACADEMIC SERVICE

- Journal referee**
Journal of Computational and Graphical Statistics
- Student paper committee**
New England Statistics Symposium, The Western North American Region of The International Biometric Society conference

SERVICE & OUTREACH

Justice, Equity, Diversity and Inclusion (JEDI) committee Postdoc representative

Department of Biomedical Data Science, Stanford University Fall 2022-2023

- Identify students' and faculties' perspectives on specific issues and suggestions.
- Design and coordinate activities to create a more diverse and inclusive environment in the department.

Stanford Future Advancers of Science and Technology (FAST) Mentor & Outcomes officer

Stanford University 2020-2023

- FAST is a program where Stanford graduate students mentor local high schoolers on science projects and share enthusiasm for science and research. High school students present their work in local science fairs, state science fairs, and an annual Symposium at Stanford University.
- Meet with students twice a month to brainstorm project ideas, carry out experiments and analyze data.
- Design and analyze program surveys and interviews to assess the effect of the FAST on students' confidence and ability to conduct scientific inquiries, and their attitude towards STEM fields, as well as mentors' confidence in mentoring high schoolers. Recommended mentor training topics based on survey results.

Inclusive Mentoring in Data Science Workshop Mentor

Stanford University Winter 2022, 2023

- Mentored an underrepresented minority student from non-R1 institution to plan academic and career trajectory, prepare resume, and discuss data science topics.

Stanford Women in Math Mentoring (SWIMM) Mentor

Stanford University 2017-2020

- SWIMM is a mentoring program to encourage undergraduate women to pursue advanced study in mathematics by pairing them with graduate student mentors, sharing resources about events, courses, and fellowship applications, and introducing them to a network of other undergraduate students.
- Monthly meetup and check-in with undergraduate mentee to discuss course selection, campus life and career development.

Ph.D. student social coordinator

Department of Statistics, Stanford University 2017

- Organized department social events, e.g., tea, happy hours and Chinese New Year potluck.
- Applied for Stanford SPICE grant (Student Projects for Intellectual Community Enhancement) (\$5000) to fund department retreats where Ph.D. students present their research in a casual environment, and socialize with fellow students and faculty members (more than 75% Ph.D. students and three faculty members attend each year).