Jing Wang

Education

2021.9–present **Ph.D. candidate, Statistics**, *University of Connecticut*, Storrs, Connecticut
2018.9–2021.6 **M.S., Statistics**, *AMSS, Chinese Academy of Sciences*, Beijing, China
2014.9–2018.6 **B.S., Mathematics and Applied Mathematics**, *Central South University*, Changsha, China

Position

2022.1–present **Research Assistant**, *Department of Statistics, University of Connecticut*, Storrs, Connecticut
2021.9–2023.5 **Teaching Assistant**, *Department of Statistics, University of Connecticut*, Storrs, Connecticut

Research Experience

- **Robust data fusion via subsampling**
  - Proposed a subsampling approach to overcome negative transfer in data fusion/transfer learning;
  - Developed a theoretical framework for a prototype transfer learning problem;
  - Conducted a non-asymptotic analysis of a random subsampling procedure for variance reduction;
  - Conducted a non-asymptotic analysis of a target-guided procedure for bias reduction;
  - Designed a select-then-combine approach for better bias-variance trade-off;
  - Demonstrated that subsampling can prevent negative transfer by outperforming full data estimators.
- **Statistical inference for streamed longitudinal data**
  - A discussion paper of "Statistical inference for streamed longitudinal data";
  - Showed potential advantages of subsampling methods when applied to correlated data;
  - Provided a bias-variance trade-off perspective on the existing method;
  - The paper has been published on *Biometrika* [1].
- **Rare-events data with sparse models**
  - Conducted theoretical analysis on subsampling for highly imbalanced binary data with sparse models;
  - Developed an algorithm that combines subsampling with adaptive lasso for rare-events data with sparse models;
  - Identified a scaling issue for existing subsampling methods;
  - Proposed a scale-invariant optimal subsampling probability based on prediction to address the issue;
  - Theoretically justified that subsampling estimators can be as efficient as full data estimators;
  - The paper has been published on *NeurIPS 2024* [2].
- **Sampling with replacement v.s. Poisson sampling**
  - Conducted a rigorous comparative study of two popular subsampling techniques;
  - Provided detailed theoretical analysis on unconditional and conditional asymptotic normality;
  - Proposed a new approach for unconditional asymptotic normality based on characteristic functions;
  - The paper has been published on *IEEE Transactions on Information Theory* [3].
- **Subsampling algorithms under measurement constraints**
  - Conducted theoretical analysis on subsampling for generalized linear models using a martingale framework;
  - Provided more detailed unconditional asymptotic theorems compared to existing literature;
  - Proposed an unweighted subsampling over existing weighted one for massive data with efficiency analysis;
  - The paper has been published on *Canadian Journal of Statistics* [4].

Ongoing project

- **Zero reduction sampling for rare-events data**
  - Extend rare-events data analysis with binary responses to data with excess zeros;
  - Conduct theoretical analysis of subsampling techniques for data with excess zeros;
  - Develop a goodness-of-fit test and an estimator based on the test.

+1(959)-929-9602 • jing.7.wang@uconn.edu • maayawang.github.io/mysite/
MaayaWANG

- **Survival analysis with excess censorings**
  - Establish a connection between Cox regression and logistic regression under the rare-events framework;
  - Conduct theoretical analysis for Cox regression with excess censorings;
  - Propose a subsampling and estimation procedure for survival data with rare-events.

Teaching Experience

Spring 2023 STAT 1100Q. Elements of Statistics (018D, 032D)

Fall 2022 STAT 1000Q. Introduction to Statistics I (027D-029D)

Honors & Awards

Summer Fellowship, Department of Statistics, University of Connecticut *May, 2024*

Pre-doctoral Fellowship, Department of Statistics, University of Connecticut *Jan., 2024*

Best Performance in Probability, Department of Statistics, University of Connecticut *Oct., 2023*

Best Performance in Inference, Department of Statistics, University of Connecticut *Oct., 2023*

MassMutual Student Paper Award. 36th New England Statistics Symposium *Jun., 2023*

Summer Fellowship, Department of Statistics, University of Connecticut *Jun., 2023*

Fairfield & Dolores Smith Award, Department of Statistics, University of Connecticut *Oct., 2022*

Gottfried Noether Award, Department of Statistics, University of Connecticut *Oct., 2022*

Munich RE/HSB Student Poster Award, 34th New England Statistics Symposium *Sep., 2021*

Best student paper(2nd class), 15th seminar of Uniform design Profession Committee of CMS *Jun., 2021*

Distinguished Graduates, Central South University *Jun., 2021*

2016 S.T.Yau College Mathematics Contests Honorable Mention in Probability and Statistics *Jul., 2016*

Publications

[1] Jing Wang, HaiYing Wang, and Kun Chen. Discussion of Statistical inference for streamed longitudinal data. *Biometrika*, 110(4):863–866, 11 2023.

[2] Jing Wang, HaiYing Wang, and Hao Zhang. Scale-invariant optimal sampling for rare-events data and sparse models. *Advances in Neural Information Processing Systems*, 37:98384–98418, 2024.

[3] Jing Wang, Jiahui Zou, and HaiYing Wang. Sampling with replacement vs poisson sampling: a comparative study in optimal subsampling. *IEEE Transactions on Information Theory*, 68(10):6605–6630, 2022.

[4] Jing Wang, HaiYing Wang, and Shifeng Xiong. Unweighted estimation based on optimal sample under measurement constraints. *Canadian Journal of Statistics*, 52(1):291–309, 2024.

Conference Presentations

(1) Subsampling for transfer learning (invited talk), 18th International Joint Conference on Computational and Financial Econometrics (Virtual), King's College London, London, United Kingdom, Dec., 2024

(2) Scale-invariant optimal sampling and variable selection with rare-events data (poster), The Thirty-Eighth Annual Conference on Neural Information Processing Systems, Vancouver, Canada, Dec., 2024

(3) Unweighted estimation based on optimal sample under measurement constraints (poster), Design and Analysis of Experiments, Virginia Tech, Blacksburg, Virginia, May, 2024

(4) Subsampling for transfer learning (invited talk), 37th New England statistics symposium, University of Connecticut, Storrs, Connecticut, May, 2024

(5) Scale-invariant optimal sampling and variable selection with rare-events data (invited talk), 2023 seminar on design of experiments and uncertainty quantification, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China, Jun., 2023

(6) Scale-invariant optimal sampling and variable selection with rare-events data (invited talk), 6th international conference on design of experiments, University of Memphis, Memphis, Tennessee, May, 2023

(7) Unweighted estimation based on optimal sample under measurement constraints (invited talk), 2022 joint statistical meetings, Washington D.C., Aug., 2022

(8) Unweighted estimation based on optimal sample under measurement constraints (contributed talk), 15th seminar of Uniform design Profession Committee of CMS, Zhongnan University of Economics and Law, Wuhan, China, Jun., 2021

Skills

Programming languages: Julia, R, Python, Matlab

Editorial tools: LaTeX, Emacs, VS code

Other tools: Linux, Git, Github