HAOLIN ZOU

+1-929-216-9678 | kyrinzou@gmail.com | Linkedin | Google Scholar

New York, NY 10025, United States

Research Interest

I have a broad interest in the theoretical foundations of machine learning and artificial intelligence, with a primary focus on proportional high dimensional asymptotics. My current research focuses include: risk estimation and cross validation methods, machine unlearning, data evaluation and interpretability.

Education

- **Columbia University** *Sep 2020 - May 2026*
  Ph.D in Statistics, GPA: 4.0/4.0 New York, USA
  - Selected courses: Prob and Stats (theoretical and computational), Topics (high-dim stats and decoupling).
- **Columbia University** *Sep 2018 - Dec 2019*
  Master in Actuarial Science, GPA: 4.0/4.0 New York, USA
  - Selected courses: Actuarial Methods, Global Capital and Investment, Quant Methods for Finance.
- **Peking University** *Sep 2014 - Jul 2018*
  Bachelor of Applied Mathematics, GPA: 3.6/4.0; Bachelor of Economics (double major), GPA: 3.8/4.0 Beijing, China
  - Selected courses: Calculus (basic, complex, real and functional), Algebra (advanced and abstract), Computation(data structure and algorithms), Stats (time series and stochastic processes), Economics (micro and macro), Finance.

Selected Papers

J=Journal, C=Conference, S=In Submission, A=On Arxiv

[1] Zou, H. et al. (2025). Approximate Certified Data Removal under High Dimensional Regime. Under submission.

[2] Zou, H. et al. (2024). A Complete Error Analysis of the K-fold CV for R-ERM in High Dimensions. Under submission for publication in *IEEE Trans.Inf.Theory*.

[3] Zou, H. et al. (2024). A Scalable Formula for the Moments of a Family of Self-Normalized Statistics. Submitted to in *J. Appl. Probab*.

[4] Baydil, A., de la Peña, V., Zou, H. and Yao, H. (2025). Unbiased estimation of the Gini coefficient. *Stat. Prob. Letters*, 222: 110376.

[5] Zou, H., et al. (2025). Leave-one-out Cross Validation in High Dimensional Settings. *AISTATS 2025*.

[6] Auddy, A., Zou, H., Rahnama Rad, K. and Maleki, A. (2024). Approximate Leave-one-out CV for Regression with L1 Regularizers. *IEEE Trans. Inf. Theory*.

[7] Auddy, A., Zou, H., Rahnama Rad, K. and Maleki, A. (2024). Approximate Leave-one-out CV for Regression with L1 Regularizers. *Proceedings of The 27th International Conference on AISTATS*,238:2377-2385. **Selected for oral presentation**.

[8] de la Peña, V., Gzyl, H., Mayoral, S., Zou, H., and Alemayehu, D. (2024). Prediction and estimation of random variables with infinite mean or variance. *Commun. Stat-Theory and Methods*, 1-15.

Presentations

- **Minghui Yu Memorial Conference 2025** *Apr 2025*
  Volunteer presentation on approximate data removal
- **Columbia Statistics Seminar** *Nov 2024*
  Volunteer presentation on high dimensional statistics
- **AISTATS 2024** *May 2024*
  Paper S.1 Selected for oral presentation in Oral Session 9 ("Statistics")
- **Minghui Yu Memorial Conference 2024** *Apr 2024*
  Volunteer presentation on approximate leave-one-out cross validation
- **INFORMS 2022** *Oct 2022*
  Session chair ("Heavi-tailedness, Dependence and Robustness"), presentation on the bias of Gini coefficient

Teaching Experience

- **Co-instructor** *Apr 2024*
  Short Course on Decoupling and Self-normalized Inequalities, Georgia Institute of Technology
  - Co-instructed with Victor de la Peña on the application of decoupling and self-normalization, including bandit and sorting problems.
- **Instructor: Recitation** *Jan 2024 - May 2024*
  Columbia University
  - Weekly recitation for Stat Inference and Modeling.
- **Teaching Associate** *Sep 2020 - Nov 2024*
  Columbia University
  - Introduction to Statistics (2020 Fall, 2023 Summer and 2023 Fall)
  - Statistical Inference (2021 Spring, 2021 Fall, 2022 Spring and 2024 Fall)
  - Probability (2022 Fall and 2024 Fall)
  - Generalized Linear Models (2023 Spring)
  - Statistical Inference and Modeling (2024 Spring)

Services

- **Peer Reviewing**
  - Annals of Applied Statistics
  - Journal of the Royal Statistical Society
  - IEEE Transactions on Information Theory
  - Conference on Neural Information Processing Systems (NeurIPS)

Awards and Distinctions

- **Second Prize** *Sep 2017*
  Contemporary Undergraduate Mathematical Contest in Modeling (CUMCM)
- **First Class Scholarship** *Sep 2017*
  Yizheng Alumni Scholarship, top 25%
- **Honorable Mention** *Feb 2017*
  Mathematical Contest in Modeling (MCM/ICM)
- **First Prize** *Oct 2014*
  China Undergraduate Contest in Physics

Additional Information

Coding and Computing: Python (proficient), R, MS Office, STATA, STAN (advanced)
Languages: English (proficient), Chinese (native), Japanese (basic), Latin (beginner).