Haofeng Zhang

hz2553@columbia.edu | https://hz0000.github.io/ | Update: 02/07/2024

Research Interests

My primary research interests are in data-driven decision-making and prediction at the intersection of operations research, data science, and machine learning. I am particularly interested in the following fields:

- **Methodologies and theory**
  - **Uncertainty quantification and reduction**, including simulation-based approaches, generative models, model calibration, distributional robustness and shift.
  - **Data-driven decision-making under uncertainty**, including data-driven contextual optimization, sequential decision-making under uncertainty (e.g., bandits and causality), and stochastic optimization under constraints.
- **Expanding to applications in real-world problems**, including recommendation systems and generative artificial intelligence.

Education

Columbia University, New York City, New York
Department of Industrial Engineering and Operations Research

- Doctor of Philosophy (Ph.D.) in Operations Research — 2024
- Master of Science (M.S.) in Operations Research — 2019

University of Science and Technology of China, Hefei, China
Department of Mathematics

- Bachelor of Science (B.S.) in Mathematics and Applied Mathematics — 2017

Publications & Working Papers

Author ordering in most papers is alphabetical as is convention in OR/IE/MS, while exceptions are marked by *.

Published papers

- Ziyi Huang, Henry Lam, and **Haofeng Zhang**. Efficient Uncertainty Quantification and Reduction for Over-Parameterized Neural Networks. *Advances in Neural Information Processing Systems (NeurIPS)*, 2023.
- Ziyi Huang, Henry Lam, Amirhossein Meisami, and **Haofeng Zhang**. Optimal Regret Is Achievable with Bounded Approximate Inference Error: An Enhanced Bayesian Upper Confidence Bound Framework. *Advances in Neural Information Processing Systems (NeurIPS)*, 2023.
- Henry Lam and **Haofeng Zhang**. Doubly Robust Stein-Kernelized Monte Carlo Estimator: Simultaneous Bias-Variance Reduction and Supercanonical Convergence. *Journal of Machine Learning Research (JMLR)*, 24(85):1-58, 2023.

  New England Statistics Symposium Student Paper Award 2022
- Ziyi Huang, Yu Gan, Theresa Lye, Yanchen Liu, **Haofeng Zhang**, Andrew Laine, Elsa Angelini, and Christine Hendon.* Cardiac Adipose Tissue Segmentation via Image-Level Annotations. *IEEE Journal of Biomedical and Health Informatics (JBHI)*, 2023.
- Ziyi Huang, Henry Lam, and **Haofeng Zhang**. Conditional Coverage Estimation for High-Quality Prediction Intervals. *Journal of Systems Science and Systems Engineering* (Invited Paper to the Special Issue on Simulation and AI), 1-31, 2023.
- Haoxian Chen, Ziyi Huang, Henry Lam, Huajie Qian, and **Haofeng Zhang**. Learning Prediction Intervals for Regression: Generalization and Calibration. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, pp. 820-828. PMLR, 2021.

- Henry Lam and **Haofeng Zhang**. Neural Predictive Intervals for Simulation Metamodeling. In *Winter Simulation Conference (WSC)*, pp. 1-12. IEEE, 2021.
- Ziyi Huang, **Haofeng Zhang**, Andrew Laine, Elsa Angelini, Christine Hendon, and Yu Gan.* Co-Seg: An Image Segmentation Framework against Label Corruption. In *IEEE International Symposium on Biomedical Imaging (ISBI)*, pp. 550-553. IEEE, 2021.
- Ziyi Huang, Yu Gan, Theresa Lye, **Haofeng Zhang**, Andrew Laine, Elsa Angelini, and Christine Hendon.* Heterogeneity Measurement of Cardiac Tissues Leveraging Uncertainty Information from Image Segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 782-791. Springer, Cham, 2020.
- Henry Lam and **Haofeng Zhang**. On the Stability of Kernelized Control Functionals on Partial and Biased Stochastic Inputs. In *Winter Simulation Conference (WSC)*, pp. 344-355. IEEE, 2019.

  ACM SIGSIM WSC 2019 Student Award 2019

Papers under review/revision

- Adam N. Elmachtoub, Henry Lam, **Haofeng Zhang**, and Yunfan Zhao. Estimate-Then-Optimize versus Integrated-Estimation-Optimization versus Sample Average Approximation: A Stochastic Dominance Perspective. Under revision in *Operations Research.*

  Finalist, INFORMS George Nicholson Student Paper Competition 2023

- Henry Lam and **Haofeng Zhang**. Prediction Intervals for Simulation Metamodeling. Under revision in *ACM Transactions on Modeling and Computer Simulation (TOMACS).*

Working papers

- Ziyi Huang and **Haofeng Zhang**. Bayesian Bandit Algorithms with Approximate Inference in Stochastic Linear Bandits.
- Ziyi Huang, Henry Lam, Amirhossein Meisami, **Haofeng Zhang**, Jie Zhang, and Yunfan Zhao. Calibration of Deep Bayesian Bandits via Off-Policy Evaluations.
- Ziyi Huang, Henry Lam, and **Haofeng Zhang**. Validating Stochastic Simulation Models via Maximum Mean Discrepancy.
- Ziyi Huang, Hongshan Liu, **Haofeng Zhang**, Fuyong Xing, Andrew Laine, Elsa Angelini, Christine Hendon, and Yu Gan.* Push the Boundary of SAM: A Pseudo-label Correction Framework for Medical Segmentation.

Funding & Selected Awards

Supports from the following funding sources are gratefully acknowledged:

- Cheung-Kong Innovation Doctoral Fellowship 2021-2023 (covering tuition and stipend, approximately $160K in total)

The following awards are gratefully acknowledged:

- Finalist, INFORMS George Nicholson Student Paper Competition 2023
- NeurIPS Scholar Award 2023
- New England Statistics Symposium Student Paper Award 2022
- ACM SIGSIM WSC 2019 Student Award 2019
- Graduation Thesis Award 2017 (top 3%)
- Undergraduate Research Award 2016 (top 2%)
- National Scholarship 2015 (top 1%)

Professional Services

- **Co-Session Chair**: WSC 2022, INFORMS Annual Meeting 2021

- **Reviewer or contribute to reviewing**:

  - Journals: Applied Probability Journals, INFORMS Journal on Computing, SIAM Journal on Optimization, Operations Research, Management Science.

  - Conferences: AISTATS, IJCAI, ISBI, NeurIPS.

Selected Invited Talks

Estimate-Then-Optimize versus Integrated-Estimation-Optimization versus Sample Average Approximation: A Stochastic Dominance Perspective

- INFORMS Optimization Society Conference (Houston TX, 03/2024), INFORMS Annual Meeting (Phoenix AZ, 10/2023), Purdue Research Symposium on Operations (West Lafayette IN, 09/2023), International Conference Stochastic Programming (Davis CA, 07/2023), SIAM Conference on Optimization (Seattle WA, 06/2023).

Doubly Robust Stein-Kernelized Monte Carlo Estimator: Simultaneous Bias-Variance Reduction and Supercanonical Convergence

- New England Statistics Symposium (Storrs CT, 05/2022), INFORMS Annual Meeting (Anaheim CA, 10/2021).

Validating Stochastic Simulation Models via Maximum Mean Discrepancy

- INFORMS Annual Meeting (Indianapolis IN, 10/2022).

Learning Prediction Intervals for Regression: Generalization and Calibration

- ICML Workshop on Distribution-Free Uncertainty Quantification (Baltimore MD, 07/2022), INFORMS Annual Meeting (Virtual, 11/2020).

On the Stability of Kernelized Control Functionals on Partial and Biased Stochastic Inputs

- INFORMS Annual Meeting (Seattle WA, 10/2019).

Other Experience

Data Science Intern at Adobe Inc. *05/2022 - 08/2022*

- Improve existing algorithms for a personalized recommendation system. Build new modules for model and hyperparameter calibration, introducing off-policy evaluation approaches into deep-learning-based Bayesian bandit algorithms. Investigate algorithm performance in the presence of practical issues, including low click-through rates, approximate Bayesian inference, changing action sets, and delayed feedback.

Data Science Intern at Adobe Inc. *05/2021 - 08/2021*

- Implement and compare multiple contextual bandit algorithms for building a new personalized recommendation system. Propose robust deep-learning-based Bayesian bandit algorithms that are effective and computationally efficient in practice.

Research Assistant & Teaching Assistant at Columbia University *2018 - 2024*

- As a teaching assistant on the following graduate-level courses: Introduction to Financial Engineering (Fall 2020); Quantitative Risk Management (Spring 2020); Credit Risks & Credit Derivatives (Fall 2019); Applications Programming for Financial Engineering (Spring 2019, Fall 2018)

Skills

- **Programming Languages & Software**: Python (including PyTorch, TensorFlow, Keras, SciPy, scikit-learn, Numpy, Matplotlib); R; C; MATLAB; SQL; LaTeX