

Education

- 2021–? **PhD Statistics** *Carnegie Mellon University*. Pittsburgh, PA
Advised by Aaditya Ramdas
- 2019–2020 **MS Machine Learning** GPA: 4.0/4.3
Advised by David P. Woodruff, Justin Khim and Pradeep Ravikumar on various projects.
- 2015–2019 **BS Computer Science** GPA: 3.88/4.0

Publications

1. Active, anytime-valid risk controlling prediction sets
Z. Xu, N. Karampatziakis, and P. Mineiro NeurIPS 2024
2. Post-selection inference for e-value based confidence intervals
Z. Xu, R. Wang, and A. Ramdas Electronic Journal of Statistics 2024
Runner-up Poster Prize @ MCP 2022
3. Foundations of testing for finite-sample causal discovery
T. Yan, **Z. Xu**, and Z. C. Lipton ICML 2024
4. Online multiple testing with e-values
Z. Xu and A. Ramdas AISTATS 2024
5. Risk-limiting financial audits via weighted sampling without replacement
S. Shekhar, **Z. Xu**, Z. C. Lipton, P. J. Liang, and A. Ramdas UAI 2023
6. Memory bounds for the experts problem
V. Srinivas, D. P. Woodruff, **Z. Xu**, and S. Zhou STOC 2022
7. A unified framework for bandit multiple testing
Z. Xu, R. Wang, and A. Ramdas NeurIPS 2021
8. Dynamic algorithms for online multiple testing
Z. Xu and A. Ramdas Math. and Sci. ML 2021
9. Class-weighted classification: Trade-offs and robust approaches
Z. Xu, C. Dan, J. Khim, and P. Ravikumar ICML 2020
10. Strategy and policy learning for non-task-oriented conversational systems
Z. Yu, **Z. Xu**, A. W. Black, and A. Rudnicky SIGDIAL 2016
11. Chatbot evaluation and database expansion via crowdsourcing
Z. Yu, **Z. Xu**, A. W. Black, and A. Rudnicky RE-WOCHAT workshop of LREC 2016

Preprints

12. Bringing closure to FDR control: beating the e-Benjamini-Hochberg procedure
Z. Xu and A. Ramdas 2025

13. Active multiple testing with proxy p-values and e-values
Z. Xu, C. Wang, L. Wasserman, K. Roeder, and A. Ramdas 2025
14. An online generalization of the (e-)Benjamini-Hochberg procedure
L. Fischer, **Z. Xu**, and A. Ramdas 2024
15. More powerful multiple testing under dependence via randomization
Z. Xu and A. Ramdas 2023

Projects

- Oct. **Real Estate Auditing**, *Carnegie Mellon University*
- 2021-Dec. I am providing statistical help (e.g. data analysis, writing expert reports, etc.) for a lawsuit against Allegheny County concerning their practices for computing the assessed values (and consequently property taxes) of newly purchased homes. This was in collaboration with Barbara Stern, John Silvestri, Esq., and Prof. Aaditya Ramdas. Recent news coverage of the case is linked [here](#).

Industry

- Feb-May. **ML Research Intern**, *Netflix*, New York, NY
- 2025 Team: *Machine Learning Inference Research*. Mentor: Michael Lindon. Manager: Nathan Kallus. Applying multiple testing and e-values to improving the A/B testing engine.
- May-Aug. **Research Intern**, *Microsoft Research*, Redmond, WA
- 2023 Team: *Reinforcement Learning*. Mentor: Paul Mineiro. I developed an anytime-valid method that uses active learning for calibrating the risk of black-box machine learning models [1].
- Mar.-May. **Consultant**, *Growthbook*, Remote
- 2023 I consulted on a project for implementing safe anytime-valid inference (SAVI) methods into Growthbook's A/B testing engine. Documentation is linked [here](#) and the open source implementation is [here](#).
- June-Aug. **Engineering Intern**, *Twitter*, Remote
- 2022 Team: *Experimentation Data Science*. Mentors: Luke Sonnet, Umashanthi Pavalanathan. Manager: Brent Cohn. I analyzed use of safe-anytime valid inference (SAVI) methods for A/B testing.
- May-Aug. **Science Intern**, *CTRL-labs (now part of Facebook Reality Labs)*, New York
- 2018 I developed state-of-the-art LSTM ensemble model that models hand movement from electromyography (EMG) signals in **TensorFlow**.
Built parser for constructing acyclic graph pipeline for preprocessing real time EMG signals.
- May-Aug. **Software Engineering Intern**, *Bloomberg*, New York
- 2017 I worked on the Message Infrastructure team, where I imported RapidCheck, a Haskell QuickCheck inspired testing framework, into the Bloomberg **C++** environment.
- May-Aug. **Software Engineering Intern**, *PicMonkey*, Seattle
- 2016 I helped build the user interface and photo editing features for the launch of the mobile photo editor app in both **Android** and **iOS**.

Talks

- | | | |
|------------|---|--|
| July. 2022 | Valid inference under S^3 bias for A/B testing | <i>Twitter ML Modeling Seminar</i> |
| Jun. 2022 | Post-selection inference for e-value based confidence intervals | <i>Safe, Anytime-Valid Inference (SAVI) and Game-theoretic Statistics Workshop</i> |
| Mar. 2022 | Post-selection inference for e-value based confidence intervals | <i>International Seminar on Selective Inference</i> |
| Nov. 2021 | A unified framework for bandit multiple testing | <i>Waterloo Student Conference in Statistics, Actuarial Science and Finance</i> |

Sep. 2021 Dynamic algorithms for online multiple testing

Workshop on current and future trends in multiple hypothesis testing (MRC Cambridge)

Teaching

Teaching Assistant

36–402: Advanced Methods for Data Analysis (Spring 2023, 2024)

36–750: Statistical Computing (Fall 2023)

36–650: Statistical Computing (Fall 2021)

15–251: Great Theoretical Ideas in Computer Science (Fall 2017, Spring 2018, Fall 2018)

15–150: Introduction to Functional Programming (Fall 2016, Spring 2017)

Service

2024–2025 Mentorship Program in Stat&DS organizer + mentor

2022–2023 CMU StatML Reading Group (SMLRG) organizer

2020 SCS Master's Advisory Committee

2020 MLD Master's Admissions Committee

Reviewing

2025 AISTATS, ICML, Statistical Methods in Medical Research

2024 NeurIPS, Biometrika, Statistica Neerlandica

2023 STOC, Electronic Journal of Statistics, New England Journal of Statistics in Data Science

2022 Mathematical and Scientific Machine Learning

2021 AISTATS