DANA YANG

Assistant Professor
Cornell University, Department of Statistics and Data Science

Research interests: The broad area of high-dimensional statistics and machine learning, including random network analysis, optimality analysis, Bayesian analysis, oracle inequalities, nonparametric estimation, convergence analysis for algorithms, rapidly mixing Markov chains, ethics and safety in machine learning.

EDUCATION

Yale University 2014 - 2019
Ph.D. in Statistics & Data Science

Advisors: Prof. David Pollard, Prof. John Lafferty, Prof. Yihong Wu

Thesis: "A few topics in statistics"

Yale University 2013-2014
M.A. in Statistics

Tsinghua University 2009-2013
B.S. in Mathematics

University of Washington Spring 2012
Exchange student in Mathematics & Statistics

AWARDS AND HONORS

Simons-Berkeley Research Fellowship
The Simons Institute for the Theory of Computing, UC Berkeley Aug 2021

Francis J. Anscombe Award for Academic Excellence
Department of Statistics and Data Science, Yale University May 2019

PUBLICATIONS AND PREPRINTS

1. Jiaming Xu, Kuang Xu and Dana Yang, "Learner-Private Convex Optimization", *ICML*, 2021.
2. Jiaming Xu, Kuang Xu and Dana Yang, "Optimal query complexity for private sequential learning against eavesdropping", *AISTATS*, 2021.
3. Jian Ding, Yihong Wu, Jiaming Xu and Dana Yang, "The planted matching problem: Sharp threshold and infinite-order phase transition", *arXiv preprint arXiv: 2103.09383*, 2021.
4. Victor-Emmanuel Brunel, Jason M. Klusowski and Dana Yang, "Estimation of convex supports from noisy measurements", *Bernoulli*, 2020.
5. Pierre C. Bellec and Dana Yang, "The cost-free nature of optimally tuning Tikhonov regularizers and other ordered smoothers", *ICML*, 2020.
6. Jian Ding, Yihong Wu, Jiaming Xu and Dana Yang, "Consistent recovery threshold of hidden nearest neighbor graphs", *IEEE Transactions on Information Theory*, 2021. (Shorter version in *COLT*, 2020)
7. Dana Yang, "Posterior asymptotic normality for an individual coordinate in high-dimensional linear regression", *Electronic Journal of Statistics*, 13(2), pages 3082-3094, 2019.

8. Jason M. Klusowski, Dana Yang and W.D. Brinda, "Estimating the coefficients of a mixture of two linear regressions by expectation maximization", *IEEE Transactions on Information Theory*, 65(6), pages 3515-3524, 2019.

9. W.D. Brinda, Jason M. Klusowski and Dana Yang, "Hölder's identity", *Statistics & Probability Letters*, Volume 148, Pages 150-154, 2019.

10. David Pollard and Dana Yang, "Rapid mixing of a Markov chain for the exponentially weighted aggregation estimator", *arXiv preprint arXiv: 1909.11773*, 2019.

11. Dana Yang, John Lafferty and David Pollard "Fair quantile regression", *arXiv preprint arXiv: 1907.08646*, 2019.

12. Kun Tian, Xiaoqian Yang, Qin Kong, Changchuan Yin, Rong L. He and Stephen S-T Yau, "Two dimensional Yau-Hausdorff distance with applications on comparison of DNA and protein sequences", *PloS one, 10(9)*, 2015.

13. Sören R. Künzel, David Pollard and Dana Yang, "Remarks on Kneip's linear smoothers", *arXiv preprint arXiv: 1405.1744*, 2014.

RESEARCH EXPERIENCE

The Simons Institute for the Theory of Computing

Research Fellow Aug 2021-Dec 2021

Program: Computational Complexity of Statistical Inference

The Fuqua School of Business, Duke University

Postdoctoral Associate Aug 2019-Jul 2021

Supervisor: Prof. Jiaming Xu, Assistant Professor at the Fuqua School of Business, Duke University

The Simons Institute for the Theory of Computing

Visiting Postdoc Aug 2020-Dec 2020

Program: Probability, Geometry, and Computation in High Dimensions

Emonet Lab, Yale University

Research Assistant Mar 2016-Mar 2017

Supervisor: Prof. Thierry Emonet, Associate Professor of Molecular, Cellular & Developmental Biology and Physics, Yale University

Project: Detection of behavioral patterns of Escherichia coli (E. coli) bacteria.

TEACHING EXPERIENCE

Primary Instructor: responsible for designing and teaching a 25-lecture course.

- S&DS S107: Introduction to Statistics, Yale University, Summer 2017.

Teaching Fellow: responsible for holding recitations/office hours, grading, and occasionally giving lectures.

- S&DS 625: Statistical Case Studies, Yale University, Fall 2016.
- S&DS 610: Statistical Inference, Yale University, Fall 2016.
- S&DS 551: Stochastic Processes, Yale University, Spring 2016.
- S&DS 600: Advanced Probability, Yale University, Spring 2016.

- S&DS 538: Probability and Statistics, Yale University, Fall 2015.
- S&DS 661: Data Analysis, Yale University, Fall 2015.
- S&DS 600: Advanced Probability, Yale University, Spring 2015.
- S&DS 610: Statistical Inference, Yale University, Fall 2014.
- S&DS 103: Introduction to Statistics: Social Sciences, Yale University, Fall 2013.

Statistical Consultant: responsible for the computational interactions in weekly meetings mostly with Yale researchers

- S&DS 627: Statistical Consulting, Yale University, 2016-2019

TALKS AND PRESENTATIONS

Cornell ORIE colloquium	2022
Cornell SGS Seminar	2022
INFORMS Annual Meeting	2021
The Simons Institute for the Theory of Computing, Joint IFML/CCSI Symposium	2021
University of California, Riverside, Department of Statistics	2021
Harvard University, Probabilitas Seminar	2021
International Conference on Learning Theory (ICML)	2021
International Conference on Artificial Intelligence and Statistics (AISTATS)	2021
ETH Zürich, Young Data Science Researcher Seminar	2021
Cornell University, Department of Statistics and Data Science	2021
Texas A&M University, Department of Statistics	2021
Johns Hopkins University, Mathematical Institute for Data Science	2021
University of California, Berkeley, Department of IEOR	2021
University of Washington, Department of Statistics	2021
George Mason University, Department of Statistics	2021
Purdue University, School of Industrial Engineering	2021
University of Toronto, Department of Statistical Sciences	2021
University of California, Irvine, Department of Statistics	2021
University of Notre Dame, Department of Applied and Computational Math and Statistics	2021
INFORMS Annual Meeting	2020
International Conference on Learning Theory (ICML)	2020
Conference on Learning Theory (COLT)	2020
NeurIPS, Privacy in Machine Learning workshop (spotlight)	2019
Probability Seminar Series, Department of Mathematics, Duke University	2019
Decision Science Seminar Series, The Fuqua School of Business, Duke University	2019
SMIL@Y Research Meeting, Yale University	2018

Joint Statistical Meetings (JSM) 2016

Yale Probability Network Group 2015,2014

PROGRAMMING SKILLS

R, Python

· Top 3000 competitors, Google Code Jam.

· Second prize in the National Computing Olympiad, China.

SERVICE

Reviewer for Annals of Statistics 2022

Reviewer for Innovations in Theoretical Computer Science (ITCS) 2021

Reviewer for ACM Conference on Economics and Computation (EC'20) 2020

Reviewer for Artificial Intelligence and Statistics (AISTATS) 2020

Yale S&DS M.A. admissions committee 2015-2019
Reviewer: one of four committee members handling over 300 applications each year and making admission recommendations.

Yale YHack competition 2018
Judge: reviewing around 50 submitted computational projects.

Yale S&DS Graduate Student Seminar Series 2018
Organizer: scheduling talks and leading discussions.

Reviewer for Bernoulli 2015, 2019