Ruchira Ray

ruchira.ray@columbia.edu Website Google Scholar LinkedIn

Education

2021/09 – Present **Ph.D. in Statistics**,Columbia University
Thesis topics: *Differential Privacy, uncertainty quantification in Bayesian ML*
Advisors: Cynthia Rush and Marco Avella Medina
Committee: Rachel Cummings, David Blei, Sumit Mukherjee

2024/05 **M.Phil. in Statistics**, Columbia University

2017/08 – 2021/05 **M.A. in Statistics & Data Science**, Yale University

B.S. in Statistics & Data Science, Yale University
With distinction
Thesis: *A Bayesian Mixed Model of Longitudinal Development of Attention in Infants*
Advisor: Dr. Joseph Chang

Research Experience

2021/09 – Present **Graduate Research Assistant**, Columbia University, Department of Statistics
Asymptotics for post-Bayesian methods, differentially private M-estimation, optimal privacy budget allocation with heterogeneous data

2020/05 – 2021/08 **Research Assistant**, Yale University, Autism Center of Excellence
Statistical analysis of Functional Magnetic Resonance Imaging (fMRI) data

2019/05 – 2021/05 **Research Assistant**, Yale University, Department of Statistics & Data Science
High-dimensional Gaussian mixture estimation

2018/05 – 2019/12 **Research Assistant**, Yale University, Nanobiology Institute, Berro Lab
Differential equation-based modeling of endocytosis dynamics in yeast cells

Research Interests

- Differential Privacy, algorithmic fairness
- High-dimensional statistics
- Probabilistic machine learning, Bayesian asymptotics, approximate inference
- Applications in clinical trial design, public health, and surveys

Preprints & Publications

Preprints

1. R. Ray, P. Liu, and R. Cummings, "Optimal privacy budget allocation for differentially private mean estimation," *Draft available upon request*, 2026+.
2. R. Ray, M. A. Medina, and C. Rush, "Differentially private m-estimation via exponential mechanism," *Draft available upon request*, 2026+.
3. R. Ray, M. A. Medina, and C. Rush, "Statistical guarantees for data-driven posterior tempering," 2026+. URL: https://arxiv.org/abs/2601.09122

Publications

1. R. Ray, M. A. Medina, and C. Rush, "Asymptotics for power posterior mean estimation," in *2023 59th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2023, pp. 1–8. URL: https://ieeexplore.ieee.org/document/10313460

In preparation

1. R. Ray, "Bias reduction of wasserstein projection estimators."
2. R. Ray, "Scalable sampling from the exponential mechanism via variational inference."

Presentations

2026/04 Minghui Yu Memorial Conference, Columbia University Department of Statistics, *"Differentially private M-estimation via the exponential mechanism,"* **contributed talk**

2026/02 Responsible Computing: Bringing Together Legal and Technical Perspectives, Columbia University, *"Towards optimal estimators under individual differential privacy constraints,"* **lightning talk**

2025/05 Workshop on Advances in post-Bayesian methods, University College London, *"Statistical Guarantees for Data-driven Posterior Tempering,"* **contributed talk**

2025/04 Optimization and Statistical Learning Workshop, Columbia University, *"Statistical Guarantees for Data-driven Posterior Tempering,"* **poster**

2025/03 Student Seminar, Columbia University Department of Statistics, *"Statistical Guarantees for Data-driven Posterior Tempering,"* **invited talk**

2024/07 International Society for Bayesian Analysis (ISBA) World Meeting, *"Statistical Guarantees for Data-driven Posterior Tempering,"* **poster**

2024/04 Forty Years at the Interplay of Information Theory, Probability and Statistical Learning, Yale Institute for Foundations of Data Science, *"Statistical Guarantees for Data-driven Posterior Tempering,"* **poster**

Minghui Yu Memorial Conference, Columbia University Department of Statistics, *"Asymptotics for power posterior mean estimation,"* **contributed talk**

2024/03 Columbia University Data Science Day, *"Statistical Guarantees for Data-driven Posterior Tempering,"* **poster**

2023/09 Allerton Conference on Communication, Control, and Computing, *"Asymptotics for power posterior mean estimation,"* **invited talk**

2023/05 Workshop on Information-Theoretic Methods for Trustworthy Machine Learning, Simons Institute for the Theory of Computing, *"Statistical Guarantees for Data-driven Posterior Tempering,"* **invited talk**

2023/04 Minghui Yu Memorial Conference, Columbia University Department of Statistics, *"Statistical Guarantees for Data-driven Posterior Tempering,"* **contributed talk**

2022/08 American Statistical Association Joint Statistical Meeting, *"The Third Moment Tensor Method with Principal Components and Basis Expansion,"* **contributed talk**

2022/07 International Congress of Infant Studies, *"Predictive Links between Salience Network Connectivity and Attention to Social Partners in Neonates with Familial History of Autism,"*, Available online (pp. 200 - 201), **poster**

2022/05 International Society for Autism Research, Annual Meeting, *"Predictive Links between Salience Network Connectivity and Attention to Social Partners in Neonates with Familial History of Autism,"*, Available online (pp. 274), **poster**

2019/12 American Society for Cell Biology Annual Meeting, *"A New Biomechanical Mathematical Model of Force Effects on Actin Dynamics During Clathrin-mediated Endocytosis,"*, Available online (Abstract P89/B91), **poster**

2019/10 Annual Conference on Quantitative Biology, Northwestern University, *"A New Biomechanical Mathematical Model of Force Effects on Actin Dynamics During Clathrin-mediated Endocytosis,"* Awarded first prize for best undergraduate research, **poster**

Teaching

Instructor

2026 Fall — **Probability Theory (STAT GR5203)**, Master's level
Instructor, Columbia University
Course website

Recitation leader

2025 Spring — **Statistical Inference and Modeling (STAT GR5703)**, Master's level
RL, Columbia University
Course description

2024 Fall — **Probability and Statistics for Data Science (STAT GR5701)**, Master's level
RL, Columbia University
Course description

2021 Fall — **Probability Theory (STAT GU4203)**, Upper-undergraduate level
RL, Columbia University
Course description

Teaching Assistant

2026 Spring — **Applied Categorical Data Analysis (STAT UN2104)**, Undergraduate level
TA, Columbia University
Course description

2025 Fall — **Applied Linear Regression Analysis (STAT UN2103)**, Undergraduate level
TA, Columbia University
Course description

2024 Spring — **Machine Learning (STAT GR5241)**, Master's level
TA, Columbia University
Course description

2023/2022 Fall — **Exploratory Data Analysis and Visualization (STAT GR5702)**, Master's level
TA, Columbia University
Course description

2023 Spring — **Interpretable Machine Learning (STAT GR5293)**, Master's level
TA, Columbia University
Course description

2022 Spring — **Machine Learning (STAT GU4241)**, Upper-undergraduate level
TA, Columbia University
Course description

2020/2019 Spring — **Theory of Statistics**, Upper-undergraduate/Master's level
Undergraduate TA, Yale University
Course description

2019 Fall — **Probability and Bayesian Statistics**, Upper-undergraduate/Master's level
Undergraduate TA, Yale University
Course description

Awards

2021/09 – Present — Columbia University Graduate Fellowship

2024/07 — International Society of Bayesian Analysis (ISBA) Student Travel Award

2024/04 — Yale Institute for Foundations of Data Science, Travel Award

Awards (continued)

	Columbia University Graduate School of Arts and Sciences Student Travel Award
2023/06	Columbia University Arts and Sciences Graduate Council Student Travel Award
2023/05	Simons Institute, Travel Award
2020/05 – 2020/08	Yale University Dean's Research Fellowship
	Yale University Trumbull College Richter Fellowship
2019/10	First prize for undergraduate research, Annual Conference on Quantitative Biology, Northwestern University
2018/05 – 2018/08	Yale University First-Year Summer Research Fellowship
2016/08 – 2017/05	National Honor Society member (Warren Township High School chapter)
2013/08 – 2017/05	National Merit Scholarship

Service

Reviewer
- Journal of Machine Learning Research (JMLR)
- Symposium on Foundations of Responsible Computing (FORC)

Organizer
- Statistics department organizer, graduate student union (2022/06 – 2025/09)
- Minghui Yu Memorial Conference (Lightining talks organizer, 2025/04)
- Columbia Statistics PhD Student Seminar (2024/05 – 2025/05)

Panelist
- GROW Columbia (2026/10, 2025/09, 2024/09)
- Andover Girls STEM Conference (2025/04)

Mentor
- Statistics PhD peer mentor (2023/09 – 2025/05)
- Women in Science at Columbia (WISC) (2025/01 – 2025/05)

Volunteer
- Optimization and Statistical Learning Workshop (2025/04)
- StatFest 2024 (Exhibitor room coordinator, 2024/09)

Skills

Coding	R, Python, Stan, MATLAB, basic command line experience in Linux and SLURM scheduler for high performance computing
Misc.	Academic research, teaching, LaTeX typesetting and publishing

Extracurricular

2018/05 – 2020/05 — **President**, Berkeley College Orchestra (BCO), Yale University