

Soham Jana

Sherrerd Hall, Room 213
Princeton University
Princeton, NJ, USA

Updated on: April 9, 2024
Website: <https://janasoham.github.io>
Email: soham.jana@princeton.edu
Phone: +1 917-520-7112

Research Interests

Theoretical and methodological aspects of high-dimensional statistics, robust estimation, Markov decision process, non-parametric estimation.

Education

- PhD. in Statistics and Data Science** May 2022
Yale University, New Haven, CT, USA
Thesis: Learning non-parametric and high-dimensional distributions via information-theoretic methods
Advisor: Prof. Yihong Wu
- Master of Statistics (Hons.)** (First class with distinction) May 2017
Indian Statistical Institute, Kolkata, West Bengal, India
Specialization: Theoretical Statistics
Dissertation: Characterization of single-integral non-kernel divergences
Advisor: Prof. Ayanendranath Basu
- Bachelor of Statistics (Hons.)** (First class with Distinction) May 2015
Indian Statistical Institute, Kolkata, West Bengal, India

Work experiences

- University of Notre Dame, Notre Dame, IN, USA**
Assistant Professor, Department of Applied and Computational Mathematics and Statistics. Starting August 2024
- Department of Operations Research and Financial Engineering, Princeton University, Princeton, NJ, USA**
Postdoc June 2022 – July 2024
Hosts: Prof. Sanjeev Kulkarni and Prof. Jianqing Fan
- Researcher, The First Republic Bank Research and Lifelong Learning Program June 2022 – May 2023
- Lecturer Spring 2023 and Fall 2023

Preprints (“*”: Authors list not in alphabetical order)

1. Soham Jana, Jianqing Fan, Sanjeev Kulkarni*. [A general theory for robust clustering via trimmed mean](#). arXiv preprint arXiv:2401.05574 (2024). (Submitted to the Annals of Statistics)
2. Soham Jana, Kun Yang, and Sanjeev Kulkarni*. [Adversarially robust clustering with optimality guarantees](#). arXiv preprint arXiv:2306.09977 (2023).
3. Soham Jana, Yury Polyanskiy, and Yihong Wu. [Optimal empirical Bayes estimation for the Poisson model via minimum-distance methods](#). arXiv preprint arXiv:2209.01328 (2022).

Journal publications (“*”: Authors list not in alphabetical order)

1. Soham Jana, Henry Li, Yutaro Yamada, and Ofir Lindenbaum. [Support recovery with Stochastic Gates: theory and application for linear models](#). Elsevier Signal Processing (2023), 213, p.109193.
2. Yanjun Han, Soham Jana and Yihong Wu, [Optimal Prediction of Markov Chains With and Without Spectral Gap](#), in IEEE Transactions on Information Theory, vol. 69, no. 6, pp. 3920-3959, June 2023, doi: 10.1109/TIT.2023.3239508. (**Extended from the NeurIPS version with analysis of higher-order Markov chains and different loss functions**)
3. Soham Jana and Ayanendranath Basu.* [A characterization of all single-integral, non-kernel divergence estimators](#). IEEE Transactions on Information Theory 65.12 (2019): 7976-7984.

Conference publications (“*”: Authors list not in alphabetical order)

1. Soham Jana, Yury Polyanskiy, Anzo Teh, and Yihong Wu. [Empirical Bayes via ERM and Rademacher complexities: the Poisson model](#). In Conference on Learning Theory 2023 Jul 15, PMLR 195:5199-5235.
2. Yanjun Han, Soham Jana, and Yihong Wu. [Optimal prediction of Markov chains with and without spectral gap](#). NeurIPS 2021.
3. Soham Jana, Yury Polyanskiy, and Yihong Wu. [Extrapolating the profile of a finite population](#). In Conference on Learning Theory 2020 Jul 15 (pp. 2011-2033). PMLR.

Talks

| | |
|---|----------------------|
| Neural information processing systems (NeurIPS) | Dec 2021 |
| Conference on learning theory (COLT) | July 2020, July 2023 |

Teaching

Princeton University

| | |
|--|-------------|
| Probability and stochastic systems (ORF 309/ENG 309/MAT 380) | Spring 2023 |
| Statistical Machine Learning (ORF 570) | Fall 2023 |

Honors and awards

| | |
|---|-----------|
| INSPIRE Scholarship, Govt. of India | 2012-2017 |
| Indian National Mathematical Olympiad (INMO) merit certificate (For being among top 75 in the country) | 2012 |

Services

Paper reviewer

Annals of Statistics
IEEE Transactions on Information Theory
IEEE International Symposium on Information Theory (ISIT)
Stat - an ISI Journal

Organizational duties at conferences

| | |
|---|-------------|
| <i>Joint Statistical Meetings</i> | August 2024 |
| <i>Session chair: New Advances in Nonparametric Hypothesis Testing - Part I</i> | |
| <i>IEEE Conference on Information Sciences and Systems</i> | March 2024 |
| <i>Session chair: Machine learning and statistical inference</i> | |

Yale S&DS M.A. admission committee

2021
Reviewer: one of the committee members handling over 150 applications and making admission recommendations

Yale S&DS graduate reading group

2020
Co-organizer Scheduled talks and lead discussion sessions

Yale Women in Data Science (WiDS) workshop

2020
Served as a mentor for Yale undergrad students participating in the WiDS Datathon Challenge 2020

South Asian Graduate and Professional Association at Yale (SAGA)

2018 – 2021
Treasurer, core committee member and cultural committee head
Objective: organizing socio-cultural events to promote diversity and inclusion at Yale