Soham Jana

Postdoctoral Research Associate
Operations Research and Financial Engineering
Princeton University

Updated on: October 18, 2023
Website: https://janasoham.github.io
Email: soham.jana@princeton.edu

Research Interests

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

Education

PhD. in Statistics and Data Science 2017–2022
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) 2015–2017
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) 2012–2015
Indian Statistical Institute, Kolkata, West Bengal, India

Work experiences

Post-doctoral Research Associate 2022–current
Princeton University, Princeton, New Jersey, USA
Research area: Robust clustering, data depth
Advisors: Prof. Sanjeev Kulkarni and Prof. Jianqing Fan

The First Republic Bank Research and Lifelong Learning Program 2022–2023
Princeton University, Princeton, New Jersey, USA
Advisors: Prof. Sanjeev Kulkarni, Prof. Ronnie Sircar, and Prof. Mete Soner
Research area: Capital call line of credit, resource planning

Preprints ("∗": Authors list not in alphabetical order)

1. Soham Jana, Kun Yang, and Sanjeev Kulkarni*. **Adversarially robust clustering with optimality guarantees**. arXiv preprint arXiv:2306.09977 (2023). (Submitted to the Journal of the American Statistical Association)

2. 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). To be submitted to Information and Inference: A Journal of the IMA.

3. Jianqing Fan, Soham Jana, Sanjeev Kulkarni. **A general theory for robust clustering via trimmed mean**. Manuscript under preparation, to be submitted to the Annals of Statistics.

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) 2021

Conference on learning theory (COLT) 2020, 2023

Course Instructor (at Princeton University)

Probability and stochastic systems Spring 2023
ORF 309/ENG 309/MAT 380

Statistical Machine Learning Fall 2023
ORF 570 (Assistant Instructor under Prof. Jianqing Fan)

Graduate teaching assistant (at Yale University)

Stochastic processes Spring 2021
S&DS 351–551. Instructor: Prof. Joseph Chang

Information theory Fall 2020
S&DS 364–664. Instructor: Prof. Andrew Barron

Probability theory Fall 2019
S&DS 241–541. Instructor: Prof. Winston Lin

Advanced probability Spring 2019
S&DS 400–600. Instructor: Prof. Sekhar Tatikonda

Statistical inference Fall 2018
S&DS 410–610. Instructor: Prof. Zhou Fan

Stochastic Process Spring 2018
S&DS 251–551. Instructor: Prof. Sahand Negahban

Honors and awards

INSPIRE Scholarship, Govt. of India 2012-2017

Indian National Mathematical Olympiad (INMO) merit certificate 2012
(For being among top 75 in the country)

Services

Paper reviewer
IEEE Transactions on Information Theory, Stat - an ISI Journal

Yale S&DS M.A. admisssion 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

Yale South Asian Graduate and Professional Association (SAGA) 2018- 2021
Treasurer, core committee member and cultural committee head
Objective: organizing socio-cultural events to promote diversity and inclusion at Yale

References

Sanjeev Kulkarni
William R. Kenan Jr. Professor
Electrical Engineering
Princeton University
Princeton, NJ, USA

Jianqing Fan
Frederick L. Moore '18 Professor of Finance
Operations Research and Financial Engineering
Princeton University
Princeton, NJ, USA

Yihong Wu
Professor
Statistics and Data Science
Yale University
New Haven, CT, USA

Yury Polyanskiy
Professor
Electrical Engineering and Computer Science
Massachusetts Institute of Technology
Cambridge, MA, USA