

# Soham Jana

Postdoctoral Research Associate,  
Operations Research and Financial Engineering,  
Princeton University

Updated on: January 12, 2024  
Website: <https://janasoham.github.io>  
Email: [soham.jana@princeton.edu](mailto: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

- |  |           |
|--|-----------|
| <b>PhD. in Statistics and Data Science</b><br>Yale University, New Haven, CT, USA<br>Thesis: Learning non-parametric and high-dimensional distributions via information-theoretic methods<br>Advisor: Yihong Wu  | 2017–2022 |
| <b>Master of Statistics (Hons.)</b> (First class with distinction)<br>Indian Statistical Institute, Kolkata, West Bengal, India<br>Specialization: Theoretical Statistics<br>Dissertation: Characterization of single-integral non-kernel divergences<br>Advisor: Ayanendranath Basu | 2015–2017 |
| <b>Bachelor of Statistics (Hons.)</b> (First class with Distinction)<br>Indian Statistical Institute, Kolkata, West Bengal, India  | 2012–2015 |

## Work experience

- |   |              |
|---|--------------|
| <b>Post-doctoral research associate</b><br>Princeton University, Princeton, New Jersey, USA<br>Research area: Robust clustering, data depth<br>Advisor: I am currently working jointly with Sanjeev Kulkarni and Jianqing Fan                     | 2022–current |
| <b>The First Republic Bank Research and Lifelong Learning Program</b><br>Princeton University, Princeton, New Jersey, USA<br>Advisors: Sanjeev Kulkarni, Roni Sircar, Mete Soner<br>Research area: Capital call line of credit, resource planning | 2022–2023    |

**Preprints**(Authors lists that are not in alphabetical order denoted by “\*”)

1. Soham Jana, Jianqing Fan, Sanjeev Kulkarni\*. [A general theory for robust clustering via trimmed mean](#). arXiv preprint arXiv:2401.05574 (2024)
2. Soham Jana, Kun Yang, and Sanjeev Kulkarni\*. [Adversarially robust clustering with optimality guarantees](#). arXiv preprint arXiv:2306.09977 (2023). (Submitted)

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 lists that are not in alphabetical order denoted by “\*”)

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 lists that are not in alphabetical order denoted by “\*”)

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

<b>Probability and stochastic systems</b> ORF 309/ENG 309/MAT 380	Spring 2023 Princeton University
--	-------------------------------------

#### Graduate teaching assistance

<b>Stochastic processes</b> S&DS 351–551/EENG 434/ENAS 502 Instructor: Joseph Chang	Spring 2021 Yale University
<b>Information theory</b> S&DS 364–664/EENG 454 Instructor: Andrew Barron	Fall 2020 Yale University

**Probability theory**  
S&DS 241–541  
Instructor: Winston Lin

Fall 2019  
Yale University

**Advanced probability**  
S&DS 400–600/Math 600  
Instructor: Sekhar Tatikonda

Spring 2019  
Yale University

**Statistical inference**  
S&DS 410–610  
Instructor: Zhou Fan

Fall 2018  
Yale University

## 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**  
IEEE Transactions on Information Theory  
Stat - an ISI Journal

**Yale S&DS M.A. admission committee**  
Reviewer: one of the committee members handling over  
150 applications and making admission recommendations

2021

**Yale S&DS graduate reading group**  
Co-organizer  
Scheduled talks and lead discussion sessions

2020

**Yale Women in Data Science (WiDS) workshop**  
Served as a mentor for Yale undergrad students participating  
in the [WiDS Datathon Challenge 2020](#)

2020

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

2018- 2021