

GOURAB MUKHERJEE

EMPLOYMENT:

Marshall School of Business, University of Southern California

April 2022 - Present Associate Professor (with tenure) of Data Sciences and Operations.

Aug 2013 - April 2022 Assistant Professor of Data Sciences and Operations.

EDUCATION:

2008 - 2013: **Ph.D. in Statistics**, Stanford University.

Dissertation Title: Sparsity and shrinkage in predictive density estimation.

Advisor: Iain M. Johnstone.

2006 - 2008: **Master in Statistics**, Indian Statistical Institute, Kolkata.

Specialization: Mathematical Statistics and Probability.

2003 - 2006: **Bachelor in Statistics**, Indian Statistical Institute, Kolkata.

CONTACT INFORMATION:

3670 Trousdale Parkway, Bridge Hall 300 B,

University of Southern California, Los Angeles, California 90089-0809.

• **Tel:** (213) 740-6612 • **Email:** gourab@usc.edu

• **Webpage:** <https://gmukherjee.github.io/>

RESEARCH INTERESTS:

A) STATISTICS THEORY AND METHODS

- Longitudinal Data Analysis, Predictive Inference, Empirical Bayes and Shrinkage Methods;
- Predictive Density Estimation, Sparse Estimation and Penalized Regression methods;
- Spatial Econometrics, Spatio-temporal Data Analysis;

B) APPLICATIONS AND STATISTICAL METHODS

- Quantitative Marketing: Mixed-effects Models for customer level analytics;
- Statistical Methods in Inventory planning: Quantile and Asymmetric losses;
- Healthcare & Prescription Drug Consumption: Clustering, Non-parametric two-sample tests;
- Business Statistics: Pricing of Digital Goods, Freemium models;

AWARDS AND HONORS:

- Award in Teaching Excellence from Marshall School of Business, 2025.
- 2025 Greif Entrepreneurship Research Award.
- Selected by Poets & Quants Best 40-Under-40 MBA Professors list, 2024.
- International Indian Statistical Association (IISA) Early Career Award in Statistics and Data Sciences (ECASDS), December 2022.
- Institute of Mathematical Statistics Laha Travel Award, 2012.
- NSF travel award to present at the 8th World Congress on Probability and Statistics in Istanbul, 2012.
- Centennial Teaching Assistant Award from the school of Humanities and Sciences, Stanford University, 2011.
- M. R. Iyer memorial medal for outstanding performance in Bachelor of Statistics (rank 1 in the program) at the Indian Statistical Institute, 2006.

GRANTS:

- USC Office of Research and Innovation (OORI) USC & Indian Institute of Science (USC-IISc) Collaborative Network Award, Efficient Score Estimation: Bridging Classical Statistical Theory with New-Age AI Applications, 2024-2025.
- Institute for Outlier Research in Business, The Origins of Emerging Fields: Mapping the Scientific and Geographic Roots of AI Pioneers, joint with Shantanu Dutta and Hyo Kang; 2024-2025.
- Institute for Outlier Research in Business, Gaming Analytics: Personalization, Sequencing, Sustainability, and Lifetime Value in Freemium Games, joint with Shantanu Dutta and Sunghan Ryu; Amount: \$ 10000 for 2023-2024.
- Adobe's Data Science Research Award, Statistical Joint Modeling for Integrated Marketing Flows: Personalized Promotions Optimized Over Journey and Lifespan, joint with Adel Javanmard; 2022-2023.
- Institute for Outlier Research in Business, Role of Digital Platforms in Increasing Access to Affordable Medication, joint with Shantanu Dutta and Dinesh Puranam; 2021-2022.
- NSF-DMS-1811866 Grant: Optimal Shrinkage and Empirical Bayes Prediction under Asymmetry, Censoring and Nonexchangeability. *Single PI*, 1st June 2018 - 31st Dec 2021, NSF Division of Mathematical Sciences Statistics program;

- USC Zumberge Individual Award, Project Title: High-dimensional Prediction under Asymmetric Check Loss, 2016-2017.

PUBLISHED:

1. Luo J, Banerjee T, **Mukherjee G**, and Sun W. *Empirical Bayes Estimation with Side Information: A Nonparametric Integrative Tweedie Approach*. <https://arxiv.org/abs/2308.05883>, accepted in *Statistica Sinica*, 2025.
2. Karmakar B, **Mukherjee G**, and Kar W. *Using Penalized Synthetic Controls on Truncated data: A Case Study on Effect of Marijuana Legalization on Physicians' Payments by Opioid Manufacturers*, <https://doi.org/10.1080/01621459.2024.2406583>, (forthcoming) **Journal of American Statistical Association**, 2024.
3. Bien J and **Mukherjee G**. *Generative AI for Data Science 101: Coding Without Learning To Code*, <https://arxiv.org/abs/2401.17647>, **Journal of Statistics and Data Science Education**, 2024.
4. Banerjee T, Bhattacharya B and **Mukherjee G**. *Bootstrapped Edge Count Tests for Nonparametric Two-Sample Inference Under Heterogeneity*, <https://arxiv.org/abs/2304.13848>, **Journal of Computational and Graphical Statistics** (forthcoming), 2024.
5. Bhuyan R, Javanmard A, Kim S, **Mukherjee G**, Rossi R, Yu T, Zhao H. Structured Dynamic Pricing: Optimal Regret in a Global Shrinkage Model, <https://arxiv.org/pdf/2303.15652.pdf>, **Journal of Machine Learning Research**, 2024.
6. Banerjee T, Liu P, **Mukherjee G**, Dutta S, and Che H. *A Crossed Random Effects Joint Modeling Framework for Improved Prediction of Player Responses in Massively Multiplayer Online Role Playing Games*, **Annals of Applied Statistics**, 17(3): 2533-2554 (2023).
7. Mukhopadhyay S, Kar W, and **Mukherjee G**. *Estimating Promotion Effectiveness in Email Marketing: A high-dimensional Bayesian Joint Model for Nested Imbalanced Data*, **Annals of Applied Statistics**, 17(1): 476-497 (2023).
8. **Mukherjee G**, and Johnstone IM. *On Minimax Optimality of Sparse Bayes Predictive Density Estimates*, **Annals of Statistics**, 2022, Vol. 50, No. 1, 81-106.
9. Karmakar B, Liu P, **Mukherjee G**, Che H, and Dutta S. *Improved Retention Analysis in Freemium Role-Playing Games by Jointly Modeling Players' Motivations, Progressions and Churns*, **Journal of Royal Statistical Society, Series A**, 2022, 185:102–133.
10. George E, **Mukherjee G**, and Yano K. *Optimal Shrinkage Estimation of Predictive Densities under α -divergences*. **Bayesian Analysis**, (2021) 16, Number 4, pp. 1139–1155.

11. Banerjee T, **Mukherjee G**, and Paul D. *Improved Shrinkage Prediction under a Spiked Covariance Structure*. **Journal of Machine Learning Research**, (2021) 22(180):1-40.
12. Gangopadhyay U, and **Mukherjee G**. *On Discrete Priors and Sparse Minimax Optimal Predictive Densities*, **Electronic Journal of Statistics** (2021), Vol. 15, No. 1, 1636-1660.
13. Banerjee T, Liu Q, **Mukherjee G**, and Sun W. A General Framework for Empirical Bayes Estimation in the Discrete Linear Exponential Family. **Journal of Machine Learning Research**, (2021) 22(67):1-46.
14. Gang B, Mukherjee G, and Sun W. *Large-Scale Shrinkage Estimation under Markovian Dependence*. **Book chapter for the IISA Series on Statistics and Data Science**, (2021), arXiv:2003.01873 [stat.ME] .
15. Sen N, **Mukherjee G**, and Arvin A. *The Use of Single Cell Mass Cytometry to Define the Molecular Mechanisms of Varicella-Zoster Virus Lymphotropism*. **Frontiers in Microbiology**, (2020), Vol 11, Article 1224.
16. Banerjee T, **Mukherjee G**, and Sun W. *Adaptive Sparse Estimation with Side Information*. **Journal of American Statistical Association**, (2020), Vol. 115, No. 532, Pages 2053-2067.
17. Banerjee T, Bhattacharya B, and **Mukherjee G**. *A nearest-neighbor based nonparametric test for viral remodeling in heterogeneous single-cell proteomic data*. **Annals of Applied Statistics**, (2020) 14(4): 1777-1805.
18. Ma T, Luo X, George A , **Mukherjee G**, Sen N, Spitzer T, Giudice L, Greene WC, and Roan N. *HIV efficiently infects T cells from the endometrium and remodels them to promote systemic viral spread*. **Elife** 9 (2020): e55487.
19. Banerjee T, **Mukherjee G**, Dutta S, and Ghosh P. *A Large-scale Constrained Joint Modeling Approach For Predicting User Activity, Engagement And Churn With Application To Freemium Mobile Games*. **Journal of American Statistical Association**, (2020), Vol. 115, No. 530, Pages 538-554.
20. George EI, Marchand E, **Mukherjee G**, and Paul D. New and Evolving Roles of Shrinkage in Large-Scale Prediction and Inference. **BIRS Workshop Report**, Pages 1-14, (2019).
21. Banerjee T and **Mukherjee G**. Discussion of CARS: covariate assisted ranking and screening for large-scale two-sample inference by Cai, Sun and Wang. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)** (2019).
22. Brown LB, **Mukherjee G**, and Weinstein A. *Empirical Bayes Estimates for a 2-Way Cross-Classified Additive Model*. **Annals of Statistics**, Vol. 46, No. 4 (2018), Pages 1693-1720.

23. Banerjee T, **Mukherjee G**, and Radchenko P. *Feature screening in Large Scale Cluster Analysis*, **Journal of Multivariate Analysis**, (2017), Vol 161, Pages 191-212.
24. Cavois M, Banerjee T, **Mukherjee G**, Raman N, Hussien R, Rodriguez BA, Vasquez J, Spitzer MH, Lazarus NH, Jones JJ, Ochsenbauer C, McCune JM, Butcher EC, Arvin AM, Sen N, Greene WC, and Roan NR. *Mass cytometric analysis of HIV entry, replication, and remodeling in tissue CD4+ T cells*. **Cell Reports**, (2017), Volume 20 , Issue 4 , 984-998.
25. Radchenko P and **Mukherjee G**. *Convex clustering via L_1 fusion penalization*. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, (2017) Vol. 79, No. 5 , Pages 1527-1546.
26. **Mukherjee G**. Book Review of: *Robust Methods for Data Reduction* by A. Farcomeni and L. Greco in **Journal of the American Statistical Association**, 917, Vol 111, (2016).
27. Sen N, **Mukherjee G**, and Arvin AM. *Single Cell Mass Cytometry Reveals Remodeling of Human T Cell Phenotypes by Varicella Zoster Virus*. **Methods**, (2015), Volume 90, Pages 85-94.
28. **Mukherjee G**, and Johnstone IM. *Exact minimax estimation of the predictive density in sparse Gaussian models*. **Annals of Statistics**, (2015), Vol. 43, No. 3, Pages 937-961.
29. Sen N, **Mukherjee G**, Sen A, Bendall S, Sung P, Nolan G, and Arvin AM. *Single Cell Analysis of the Remodeling of Human Tonsil T Cells by Varicella-Zoster Virus*. **Cell Reports**, (2014), Volume 8, Issue 2, 633-645.
30. Sen A, Rott L, Phan N, **Mukherjee G**, and Greenberg HB. *The rotavirus NSP1 protein inhibits IFN-mediated STAT1 activation*. **Journal of Virology**, (2014), 88.1 : 41-53.
31. Gustin SL, **Mukherjee G**, Baker VL, Westphal LM, Milki AA, and Lathi RB. *Early pregnancy testosterone after ovarian stimulation and pregnancy outcome*. **Fertility and Sterility**, (2012), Volume 97, Issue 1.
32. Sen A, Rothenberg ME, **Mukherjee G**, Feng N, Nair N, Kalisky T, Johnstone IM, Clarke MF, and Greenberg HB. *Innate immune response to homologous rotavirus infection in the small intestinal villous epithelium at single-cell resolution*. **Proceedings of the National Academy of Sciences (PNAS)** 109.50 (2012): 20667-20672.
33. Gustin SL, **Mukherjee G**, Baker VL, Westphal LM, Milki AA , Ma C, and Lathi RB. *First Trimester Testosterone After Ovarian Stimulation and Its Effect on Pregnancy Outcomes*. **Fertility and Sterility**, (2011), Volume 95, Issue 4.

SERVICE

1. 2025 Chair Elect for ASA's Statistics in Marketing Section
2. **Editorial And Referee Works:**
 - ☐ Associate Editor of Journal of Business and Economic Statistics, 2024-Current.
 - ☐ Associate Editor of Annals of Applied Statistics, 2023-Current.
 - ☐ Associate Editor of Statistica Sinica, 2023-Current.
 - ☐ Associate Editor of Sankhya Series A, 2019-Current.
3. **Conference & Workshop Organizer:**
 - ☐ Co-organizer of the workshop on *New and Evolving Roles of Shrinkage in Large-Scale Prediction and Inference* held in the Banff International Research Station from Apr 7- 12, 2019. The workshop had 26 talks and was attended by 36 participants.
 - ☐ Local organizing committee member of the 2024 Tenure Project Conference that discussed important issues affecting Black, Latinx, and Native junior faculty obtaining tenure in US business schools.
 - ☐ Scientific Program Committee member of 14th International Conference of the ERCIM WG on Computational and Methodological Statistics (CMStatistics 2021), London, December 2021.
 - ☐ Co-organizer of the Statistics Empowering Data Science conference at USC, Jan 2024.
4. **University Service at USC:**
 - ☐ Statistics PhD coordinator (Sep 2021 - current).
 - ☐ Faculty Affairs Committee: 2024-2025.
 - ☐ University Research Committee: 2023-2024.
 - ☐ USC Marshall School of Business Fellows Program: 2023-2024.
 - ☐ Faculty advisor to USC student groups SABA, BALAKA and AIESEC.

TEACHING

- DATA SCIENCE FOR BUSINESS (GSBA524)
 - *introductory data science class for MBA students*
 - *taught in fall 2023* (183 students);
 - *taught in fall 2024* (206 students);

- APPLIED BUSINESS STATISTICS (BUAD310g)
 - *introductory class for undergraduate business students; the format changed over years;*

Year	#. of weekly classes	#. of Students
2022	2 lectures, 4 applied sections	143
2021	2 lectures, 4 applied sections	132
2020	2 lectures, 4 applied sections	139
2018	2 lectures, 4 applied sections	133
2016	2 lectures, 4 applied sections	117
2015	2 lectures, 6 applied sections	196
2014	8 applied sections	258
2013	4 lectures	141

- REGRESSION AND GENERALIZED LINEAR MODELS FOR BUSINESS APPLICATIONS (GSBA604)
 - *taught in fall 2018, fall 2019, fall 2020 and fall 2021;*
 - *co-taught a modified class in fall 2022, 2023, 2024 with Clive Lennox.*
- INTRODUCTION TO BUSINESS ANALYTICS (BUAD 425)
 - *taught in spring 2023 (45 students);*