Soumyakanti Pan

PhD Candidate · Department of Biostatistics
UCLA Fielding School of Public Health
span18@ucla.edu | span-18.github.io | SPan-18

Education

University of California Los Angeles *Los Angeles, USA*
Ph.D., *Biostatistics* 2021 - 2025 (expected)

- *Thesis Title:* "Bayesian Modeling and Inference for Complex Dependent Non-Gaussian Data"
- *Thesis Advisor:* Prof. Sudipto Banerjee

Indian Statistical Institute *Kolkata, India*
M.Stat., *Statistics* 2019 - 2021

- First Division with Distinction

Indian Statistical Institute *Kolkata, India*
B.Stat. (Honours), *Statistics* 2016 - 2019

- First Division

Professional Experience

2021-Present **Graduate Student Researcher,** UCLA Biostatistics
2023-2024 **Teaching Assistant,** UCLA Biostatistics
2021-2023 **Special Reader,** UCLA Biostatistics
2019-2021 **Instructor for Competitive Mathematics,** RKMV Narendrapur
2019 **Short-term Trainee,** National Institute of Biomedical Genomics, India

Awards & Achievements

2024 **Dissertation Year Fellowship,** UCLA Division of Graduate Education USD 20,000
2022, 2024 **Summer Mentored Research Fellowship,** UCLA Graduate Division USD 12,000
2021 **University Fellowship,** UCLA Division of Graduate Education USD 14,863
2016-2021 **Stipend and contingency grant,** Indian Statistical Institute
2015 **2nd position,** Advanced Mathematical Ability Test, Calcutta Mathematical Society

Scholarly Manuscripts

Published

1. **Pan, S.**, Das, D., Ramachandran, G., & Banerjee, S. (2024). Bayesian hierarchical modeling and inference for mechanistic systems in industrial hygiene. *Annals of Work Exposures and Health*, *68*(8), 834–845. DOI: 10.1093/annweh/wxae061

Preprints

3. **Pan, S.**, Zhang, L., Bradley, J. R., & Banerjee, S. (2024). Bayesian inference for spatial-temporal non-Gaussian data using predictive stacking. DOI: 10.48550/arXiv.2406.04655 *(submitted to Bayesian Analysis)*
2. **Pan, S.** & Banerjee, S. (2024). spStack: Practical Bayesian geostatistics using predictive stacking in R .
1. **Pan, S.** & Banerjee, S. (2025). Bayesian inference for spatially-temporally misaligned data using predictive stacking: Application to the impact of ozone on asthma-related health emergencies .

Software

1. **spStack**: Bayesian Geostatistics Using Predictive Stacking.
   Fast Bayesian inference for Gaussian and non-Gaussian geospatial data without using MCMC algorithms. This `R` package is written in `C++` with calls to `FORTRAN` routines for optimized linear algebra. Available from **CRAN**.
2. **spStackCOS**: Bayesian inference for change of support models using predictive stacking.
   Offers functionalities for implementing Bayesian inference for spatially-temporally misaligned data without using MCMC. This is an offshoot of the *spStack* package, and currently *under active development*.

Conference Presentations

1. **UCLA-UCI-KAUST Meeting.** Los Angeles, USA, March 2023. Talk.
   Uncertainty quantification of dynamical systems in industrial hygiene.
2. **Joint Statistical Meetings.** Toronto, Canada, August 2023. Talk.
   Topic Contributed Paper session: Recent advances in uncertainty quantification for complex systems.
3. **The Bayesian Young Statisticians Meeting, *j-ISBA*.** Virtual Meeting, November 2023. Short Talk.
   Predictive stacking in Bayesian hierarchical models for geospatial data from the natural exponential family.
4. **Theory and Foundations of Statistics in the Era of Big Data.** Tallahassee, USA, April 2024. Poster.
   Bayesian inference for spatial-temporal count data using predictive stacking.

Other Research Experiences

Indian Statistical Institute - Human Genetics Unit *Kolkata, India*
Advisor: Prof. Saurabh Ghosh 2020-2021

- *Masters Thesis:* "Distribution-free correlation based tests to differentiate between related populations"

Indian Statistical Institute *Kolkata, India*
Advisor: Prof. Diganta Mukherjee 2019-2021

- *Title:* "Modeling and analysis of multi-layered network data on inter-state military alliances, conflicts and trade relations"

Indian Statistical Institute *Kolkata, India*
Joint work with A. Bhattacharyya, S. Ghosh Dastidar 2019

- *Bachelors Thesis:* "Estimation of origin-destination matrix from traffic and ticket counts"

Collaborations and Consultations

- 2, 3, 7, 8-tetrachlorodibenzo-p-dioxin (TCDD) exposure from Agent Orange herbicide sprays and other sources in relation to birth outcomes in Vietnam. (in preparation)
  Collaboration with Sophie Michel, Department of Epidemiology, Fielding School of Public Health, UCLA.
  Co-Author, *Role:* Bayesian modeling and inference of complex spatial survey data.
- Berberian, A. G., Morello-Frosch, R., Karasaki, S., & Cushing, L. J. (2024). Climate justice implications of natech disasters: Excess contaminant releases during hurricanes on the texas gulf coast. *Environmental Science & Technology*, *58*(32), 14180–14192. doi: 10.1021/acs.est.3c10797
  Contribution: Assistance with Bayesian modeling and analysis of spatial data.
- Gupta, N., Bhattacharya, S., Dutta, A., Tauchen, J., Landa, P., Urbanová, K., Houdková, M., Fernández-Cusimamani, E., & Leuner, O. (2024). Synthetic polyploidization induces enhanced phytochemical profile and biological activities in thymus vulgaris l. essential oil. *Scientific Reports*, *14*(5608). doi: 10.1038/s41598-024-56378-7
  Contribution: Consultation for statistical analysis.

Teaching Experience

Spring 2024	**BIOS 250C: Multivariate Biostatistics,** Teaching Assistant	UCLA
Winter 2024	**BIOS 250B: Linear Statistical Models,** Teaching Assistant	UCLA
Fall 2023	**BIOS 241: Spatial Modeling and Data Analysis,** Teaching Assistant	UCLA
Fall 2021, 22	**BIOS 255A: Measure Theoretic Probability,** Special Reader	UCLA
Winter 22,23	**BIOS 255B: Advanced Probability and Statistics,** Special Reader	UCLA
2019-2021	**Topics in Competitive Mathematics,** Instructor, RKMV Narendrapur	India

Professional Activities

Professional Memberships

- International Society for Bayesian Analysis (ISBA).
- American Statistical Association (ASA).
- International Indian Statistical Association (IISA).

Peer Review

- Annals of Applied Statistics.

Technical Skills

Programming	R, C++, and integration via R's Native Interface, Julia, Lean
Software	R, SAS, RStudio, Jupyter
Computing Platforms	Macintosh, Linux, HPC Cluster
Applications	Git, LaTeX, MS Office
Development Tools	Continuous Integration, UNIX command-line interface

References

Dr. Sudipto Banerjee, *Ph.D. advisor*
Senior Associate Dean for Academic Programs
UCLA Fielding School of Public Health
Professor *and* Past-Chair
UCLA Dept. of Biostatistics
Professor of Statistics & Data Science
University of California Los Angeles
sudipto@ucla.edu
http://sudipto.bol.ucla.edu/