Jyotishka Datta

CONTACT INFORMATION	Department of Mathematical Sciences University of Arkansas SCEN 219 Fayetteville, AR 72701, USA

CONTACT INFORMATION

Department of Mathematical Sciences
University of Arkansas
SCEN 219
Fayetteville, AR 72701, USA

Mobile: +1-765-398-2914
E-mail: jd033@uark.edu
Webpage: https://sites.uark.edu/jd033/
Git: https://github.com/DattaHub

RESEARCH INTERESTS

Bayesian methodology and theory, Sparse signal recovery, Global-local shrinkage priors, Default Bayes, Discrete data, High-dimensional data, Geo-spatial prediction, Bioinformatics, Compositional data, Applied probability, and Bayesian nonparametrics.

PROFESSIONAL EXPERIENCE

2016 to present: Assistant Professor,
Department of Mathematical Sciences, University of Arkansas, Fayetteville.

2014 to 2016: Postdoctoral Associate.
Department of Statistical Science, Duke University, Durham, NC.
Statistical and Applied Mathematical Sciences Institute, Durham, NC.

- **Postdoctoral advisors:** Prof. David B. Dunson (Statistical Science), and Prof. Sandeep S. Dave (Medicine), Duke University.
- **SAMSI Program:** Beyond Bioinformatics.

EDUCATION

2009 - 2014: Ph.D. in Statistics, Purdue University, West Lafayette, IN.

- Dissertation Topic: "Some Theoretical and Methodological Aspects of Multiple Testing, Model Selection and Related Areas",
- **Ph.D. advisor:** Prof. Jayanta K. Ghosh.

2003 - 2008: B.Stat and M. Stat, Indian Statistical Institute, Kolkata, India.

- Dissertation Topic: "Efficiency Versus Robustness - An Weighted Likelihood Equation Approach", Advisor: Prof. Ayanendranath Basu.

AWARDS AND HONORS

- Robert and Sandra Connor ***Endowed Faculty Fellowship***, University of Arkansas, 2018-19.
- ***Honorable Mention Award for Best Theoretical Poster*** at the O'Bayes 2013: The Tenth International Workshop on Objective Bayesian Statistics, December 15-19, Durham, USA.
- ***William J. Studden Publication Award*** for an outstanding publication in a mathematical statistics journal, 2013, Department of Statistics, Purdue University.
- ***Travel Awards***:
  - 19th IMS Meeting of New Researchers in Statistics and Probability, 2016
  - International Indian Statistical Association 2016 Conference
  - ASA-Kutner faculty poster session at the SRCOS 2016 Summer Research Conference
  - O-Bayes 2013 : The Tenth International Workshop on Objective Bayesian Statistics
- Award for Academic Excellence, Indian Statistical Institute, Kolkata, 2008.
- Ranked **8th** and **10th** in State Level Joint Entrance Examination in **Engineering** and **Medicine** (out of approximately two hundred thousand students), 2003.

PUBLICATIONS (STATISTICS)

[1] Bhadra, **Datta**, Li and Polson (2020). (*alphabetical[1]), "Horseshoe Regularization for Machine Learning in Complex and Deep Models". https://doi.org/10.1111/insr.12360, ***International Statistical Review***. [preprint].

---

[1](Articles co-authored with Prof. Polson have alphabetically ordered author-list.)

[2] Bhadra, **Datta**, Polson, and Willard (2020), (*alphabetical), "Global-local mixtures - A Unifying Framework". https://doi.org/10.1007/s13171-019-00191-2, *Sankhya A - J. K. Ghosh Memorial Issue*. [blog article on the paper]

[3] Bhadra, **Datta**, Li, Polson, and Willard (2019), (*alphabetical), "Prediction Risk for Global-Local Shrinkage Regression". **20 (78)**, 1-39, *Journal of Machine Learning Research*. [full-text].

[4] Bhadra, **Datta**, Polson, and Willard (2019), (*alphabetical), "Lasso Meets Horseshoe - A Survey" **34(3)**, 405-427. *Statistical Science*. [full-text]

[5] Bhadra, **Datta**, Polson, and Willard (2019), (*alphabetical), "Horseshoe Regularization for Feature Subset Selection". https://doi.org/10.1007/s13571-019-00217-7, *Sankhya B*. [preprint]

[6] Bhadra, **Datta**, Polson, and Willard (2017), (*alphabetical) "The Horseshoe+ Estimator of Ultra-Sparse Signals", *Bayesian Analysis*. **12 (4)**, 1105-1131. [full-text]

[7] **Datta** and Dunson (2016), "Bayesian inference on quasi-sparse count data", *Biometrika*, **103 (4)**: 971-983. [full-text]

[8] Bhadra, **Datta**, Polson, and Willard (2016), (*alphabetical) "Default Bayesian analysis with global-local shrinkage priors", *Biometrika*, **103 (4)**: 955-969. [full-text]

[9] **Datta**, and Ghosh (2014), "Bootstrap – An Exploration." *Statistical Methodology*: **20**, 63-72.

[10] **Datta**, and Ghosh (2013), "Asymptotic Properties of Bayes Risk for the Horseshoe Prior". *Bayesian Analysis* **8(1)**, 111-132. [full-text].

PUBLICATIONS (CANCER GENOMICS)

[11] Reddy, Anupama, *et al*. (2017) "Genetic and Functional Drivers of Diffuse Large B Cell Lymphoma". *Cell*, **171.2**: 481-494. Featured on EurekAlert!, the newsletter from AAAS, link.

[12] Moffitt *et al.* (2017). "Enteropathy-associated T cell lymphoma subtypes are characterized by loss of function of SETD2", *Journal of Experimental Medicine*, **214(5)**, 1371-86.

[13] McKinney, Moffitt, *et al.* (2017) "The Genetic Basis of Hepatosplenic T Cell Lymphoma". *Cancer Discovery*, **CD-16-0330**.

[14] Healy, *et al.* (2016). "GNA13 loss in germinal center B cells leads to impaired apoptosis and GCB cell persistence and promotes lymphoma in vivo". *Blood*, **127(22)**, 2723-2731.

PUBLICATIONS (INTER-DISCIPLINARY)

[15] **Criminology:** Steinman; Drawve; Datta; Harris; Thomas (2020): "Risky Business: Examining the 80-20 Rule in Relation to a RTM Framework". (*C*riminal Justice Review)

[16] **Pediatrics:** Chaudhuri, Biswas, Datta, . . ., Chakarabrty (2016). "Evaluation of malnutrition as a predictor of adverse outcomes in febrile neutropenia associated with pediatric hematological malignancies." *Journal of Paediatrics and Child Health*, **52 (7)**, 704-709.

[17] **Neuroscience:** Parthasarathy, Datta, Torres, Hopkins, and Bartlett (2014). "Age-Related Changes in the Relationship Between Auditory Brainstem Responses and Envelope-Following Responses." *Journal of the Association for Research in Otolaryngology*. **15 (4)**, 649-661.

[18] **Geology:** Libohova, Winzeler, Lee, Schoeneberger, Datta, and Owens (2016). "Geomorphons: Landform and property predictions in a glacial moraine in Indiana landscapes". *Catena 2016*, v.142.

REFEREED BOOK CHAPTERS	[19] **Datta** and Ghosh (2015), "In Search of Optimal Objective Priors for Model Selection and Estimation". In S. Upadhyay, U. Singh, D. Dey, & A. Loganathan (Eds.), *Current Trends in Bayesian Methodology with Applications*, 225-239. Chapman & Hall/CRC Press.
	[20] Dasgupta, Ghosh, Chakravarty, and **Datta** (2015), "Some Remarks on Pseudo Panel Data". *Growth Curve and Structural Equation Modeling*, 25-34. Springer International.
ARTICLES UNDER REVIEW (STATISTICS)	[21] Li, **Datta**, Craig, and Bhadra, (2020+). "Joint Mean-Covariance Estimation via the Horseshoe with an Application in Genomic Data Analysis". (*Submitted to* Journal of Multivariate Analysis). [preprint].
	[22] Harris; Drawve; Thomas; Datta; Steinman (2020+): "Lines of Black and White: Racial Segregation, Neighborhood Permeability, and Crime" (*Submitted to* Social Science Research).
MANUSCRIPTS IN PREPARATION	[1] **Datta**, Ovaskainen and Dunson (2020+), "Sparse generalized Dirichlet distributions for high-dimensional probabilities"
	[2] Sengupta, **Datta**, Chen (2020+), "Proximity Block-models for Network Data".
	[3] Boss, **Datta**, Kang, Kim, Mukherjee (2020+), "Group Inverse Gamma Gamma Shrinkage".
	[4] Abba, Bhadra, **Datta**, and Polson (2020+), (*alphabetical), "Bayesian Square-root Lasso".
	[5] **Datta**, Shi and Bandopadhyay, D. (2020+), "Shrinkage and Selection for Compositional Data".
	[6] **Datta** and Dunson (2020+), "Nonparametric Bayes multiresolution testing for massive-dimensional rare events".
CONFERENCE PUBLICATIONS	[7] LeBow V., Bernhardt-Barry, M. L., and **Datta**, J. (2018), Improving Spatial Visualization Abilities Using 3D Printed Blocks Paper presented at 2018 ASEE Annual Conference & Exposition , Salt Lake City, Utah. URL.
OTHER PUBLICATIONS	[8] Datta and Drawve, "Does Machine Learning Reduce Racial Disparities in Policing?", IISA Newsletter, December, 2016.
	[9] Datta and Ghosh, "Optimal Objective Priors for Linear Models", Indian Bayesian Society Newsletter, Vol XI, No. 1, May, 2014.
FUNDING	**External**

Internal

- **Robert and Sandra Connor Endowed Faculty Fellowship** from the University of Arkansas, 2018-19. ($5,000)
- **Datta, J.** and Bernhardt-Barry, M. L., "Predicting Soil Type from Non-destructive Geophysical Data", December 2018, Provost's Collaborative Research Grant ($2,200).
- **Datta, J.**, Drawve, G., Harris, C., and Thomas, S. (*alphabetical). November 2017. "Participant Field Training with Little Rock Police Department." Provost's Collaborative Research Grant ($ 2,000).
- **Datta, J.**, M. A. Abba* (*graduate student). November 2016. "Multiresolution Nonparametric Bayesian Hotspot Detection." Provost's Collaborative Research Grant ($2,000).
- *Summer Research Grant*, Department of Statistics, Purdue University, **2011**-**2013**.

INVITED TALKS
- August 1-6, 2020 (forthcoming): Invited Session (Bayesian methods in structured data and high dimensional problem: some recent advances), TBD, Joint Statistical Meeting at Philadelphia, PA.
- June 29-July 3, 2020 (postponed to 2021): Invited Session, TBD, ISBA 2020 World Meeting at Kunming, China.
- May 16-20, 2020 (forthcoming): Invited Session ,"Bayesian Shrinkage for Continuous & Discrete Data– a Tale of Two Cities", ICSA Applied Statistics Symposium 2020. Houston, TX.
- March 2020 (forthcoming, virtual): Invited session "Innovative Statistical Approaches for High-Dimensional Omic and Microbiomic Data", Title: "Sparse Generalized Dirichlet Distributions for Microbiome Compositional Data",in ENAR 2020, Nashville, Tennessee.
- December 2019: Invited Session "Bayesian Modeling and Computation", Title: 'Bayesian Shrinkage for Continuous & Discrete Data – a Tale of Two Cities" in IISA 2019 Conference, Mumbai, India.
- August 2019: Special Invited Session in Memory of Prof. J.K. Ghosh, Title: "Bayesian Sparse Signal Recovery: Gaussian Models and Beyond", in Joint Statistical Meeting, Denver, Colorado.
- August 2019: Invited Talk (Innovative Approaches for High-dimensional Omics and Neuroimaging Data) in Joint Statistical Meeting, Denver, Colorado.
- May 2019: Invited Talk, "Bayesian Shrinkage for Continuous & Discrete Data– a Tale of Two Cities", Department of Biostatistics, University of Michigan, Ann Arbor.
- January 2019: Invited Session (Multiple Testing) in Young Statisticians' Meet: Data Science in Action: January 4-5, 2019, Indian Statistical Institute, Kolkata, India.
- December 2018: Plenary Session in 10th International Calcutta Triennial Symposium, December 27-30, 2018, Kolkata, India.
- April 2018, "Bayesian Sparse Signal Recovery: Horseshoe Regularization", Departmental Statistics Colloquium, Florida State University.
- December 2017: "Horseshoe Regularization for Feature Subset Selection", 2017, IISA International Conference on Statistics at Hyderabad, India.
- December 2017: "Horseshoe Regularization for Feature Subset Selection", ERCIM WG Meeting, CMStatistics 2017 Conference at London, UK.
- August 2017: "Detecting rare mutational hotspots by multiscale BNP method", Joint Statistical Meeting, Baltimore, Maryland.
- January 2017: "Sparse signal recovery and default Bayesian analysis using global-local shrinkage priors", Applied Statistics Unit, Indian Statistical Institute, Kolkata.
- August, 2016: "Default Bayesian analysis for global-local shrinkage priors", IISA Conference, Corvallis, Oregon.
- August 2016: "Shrinkage Priors for High-Dimensional Sparse Poisson Means", Joint Statistical Meeting, Chicago, Illinois.
- February, 2016: "Shrinkage Priors for High-Dimensional Sparse Poisson Means" (STAT 701 Talk): Duke University.

- December, 2015 - January, 2016: "Sparse Signal Recovery for Discrete & Continuous Data" (Job Talk): Binghamton University, University of Arkansas at Fayetteville, and Clemson University.
- May,2015: "Multiscale Bayesian cluster detection and testing for whole genome sequencing studies", Transition workshop for "Beyond Bioinformatics", SAMSI, North Carolina.
- August, 2014: "Sparse and Ultra-Sparse Signal Recovery: The Horseshoe and The Horseshoe+Prior", Department of Statistical Science, Duke University.
- January, 2014: "Shrinkage priors for multiple testing and model selection", University of Texas - M. D. Anderson Cancer Center, Houston, TX.
- November, 2013: "In Search of Optimal Objective Priors for Model Selection and Estimation", Mathematical Statistics Seminar, Purdue University.
- May, 2013: "Two-groups and One-Group Models for Multiple Testing", National Institute of Biomedical Genomics, Kalyani, India.

MENTORING

Undergraduate Students

- Honors Thesis Committee: Vanessa Lebow, Winson Chee, Dhruba Dasgupta, Christopher Peterson.
- Academic Advising: Jodi Mitchell, Bruce Dunning, Alex Coleman, Rosario Dispensa, Kaylee Henry, David O'Hearn, Lauren Pearce.

Graduate Students

- Primary Advisor (PhD): James Roddy.
- Primary Advisor (MS): Ek Alfieri, Apu Chandra Das, Mohamed Abba, Josh Price, Kai Cui.
- Committee Member (MS): Anne Lin, Ji Li, Michael Ellis, James Willbanks, Ruizhe Yin, Shanshan Zhang, Mahboubeh Madadi, Gina Riggio (Cell and Molecular Biology Program).
- Committee Member (PhD):
  1. Ghadeer Mahdi, Department of Mathematical Sciences. (Chair: Dr. Avishek Chakraborty)
  2. Sarah Jones, Food Science. (Chair: Dr. Kristen Gibson)
  3. Thomas Yeargin, Food Science (Chair: Dr. Kristen Gibson)

TEACHING EXPERIENCE

- Fall 2016 - now, Instructor, Department of Mathematical Sciences, University of Arkansas. Teaching duties: 2 + 1 courses for first two years, then 2 + 2 courses third year onwards.
  - STAT 5443 (Computational Statistics): Spring 2017, 2018 and 2019. Advanced Graduate course. Syllabus: http://dattahub.github.io/stat5443/syllabus.html.
  - STAT 4033, (Nonparametric Statistics). Fall 2016, 2017, 2018, 2019. Audience: Undergraduate and Graduate students from quantitative disciplines. Syllabus: http://dattahub.github.io/stat4033/list.html.
  - STAT 3013 (Introduction to Probability): Fall 2017, 2018, 2019, Spring 2017, 2018, 2019. Undergraduate Course. Apps: Central Limit Theorem and Glivenko-Cantelli Lemma

- Summer 2014, **Instructor**, Department of Statistics, Purdue University.
  - Stat 301, Introduction to Statistics, Course Coordinator: Meghan Tooman.
  - Responsibilities: Designing and holding recitations and lab sessions for using SPSS for undergraduate students, grading homework, lab exercises, and midterm and final exams.

- Spring 2012-Spring 2013, **Teaching Assistant (Lab Instructor)**, Department of Statistics, Purdue University.
  - Stat 598Z, Introduction to Computing for Statisticians , Instructor: Prof. S. V. N. Vishwanathan.
  - Stat 598G, Introduction to Computational Statistics, Instructor: Prof. Sergey Kirshner.
  - Responsibilities: Holding lab sessions for teaching Statistics using SPSS to small groups of undergraduate students, grading homework, lab exercises, and midterm tests.
  - Lab Website: https://learning.cs.purdue.edu/courses/sp2013/598z/lab.

- Spring 2011-Fall 2011, **Teaching Assistant (Lab Instructor)**, Department of Statistics, Purdue University.
  - Stat 301, Introduction to Statistics, Course Coordinator: Ellen Gundlach.
  - Stat 113, Statistics for Society, Course Cordinator: Prof. John Deely.
  - Responsibilities: Teaching recitation sessions for undergraduate students, holding office hours, grading homework, lab exercises, and the midterm.

- Fall 2010, **Teaching Assistant (Grader)**, Department of Computer Science, Purdue University.
  - CS 471, Artificial Intelligence, Instructor: Prof. Alan Qi.
  - Responsibilities: Teaching recitation sessions for undergraduate students (groups of 20), holding office hours, grading homeworks, lab exercises, and the midterm.

INDUSTRIAL PROFESSIONAL EXPERIENCE

Barclays Bank, PLC, Mumbai, India.

Graduate Emerging Manager   **June 2008 to September 2009**
- Developing and testing scorecard for evaluating potential customers, Developing performance based strategies for approval of credit cards and segmentation analysis to identify delinquency behavior.

Systat Softwares Asia Pacific Ltd., Bangalore, India.

Summer Intern   **May 2005 to July 2005**
- Supervisor: Dr. T. Krishnan.
- Worked on Markov Chain Monte Carlo Methods Using SYSTAT 11 and implementation of Transformed Density Rejection Algorithm.

SOFTWARE SKILLS
- Languages: R, MATLAB, PYTHON, STAN, C.
- Statistical softwares: SPSS, SAS, JMP, STATA, MINITAB.

PROFESSIONAL ACTIVITY
- Served as a reviewer for the following journals: Journal of Royal Statistical Society (B), Annals of Applied Statistics; Journal of Multivariate Analysis; Biometrika, Journal of American Statistical Association (Theory and Methods), Statistica Sinica, Biometrika, Electronic Journal of Statistics, Sankhya Series A, Entropy, Machine Learning Conferences: NIPS, ICML, AIStats, PLoS One.
- Served as a proposal reviewer for National Science Foundation (2017).
- Organized the following topic-contributed sessions:
  - *Recent Advances in Bayesian Structure Learning* sponsored by the Section on Bayesian Statistical Science (SBSS) at Joint Statistical Meeting, Denver, CO, 2019.
  - *Scalable Bayesian Inference for structured high-dimensional data*, International Indian Statistical Association Conference (IISA), 2018, Gainsville, Florida.

- *Recent Advances in Bayesian Methodology and Computation for Ultra-High Dimensional Data* sponsored by the Section on Bayesian Statistical Science (SBSS) at Joint Statistical Meeting, Chicgao, IL, 2016.
- Served as a chair for the invited paper session on "High-dimensional Bayesian statistics: spike-and-slab and global-local shrinkage" at Joint Statistical Meeting, 2016.
- Served as a chair for the invited paper session on "Bayesian Model Selection" at Joint Statistical Meeting 2017. Baltimore, MD.
- Served as a chair for the invited paper session on "Modeling Dependence in Large Systems" at IISA 2017 Annual Conference. Hyderabad, India.
- **Committee Service:** Executive Committee and Newsletter Editor, International Indian Statistical Association (IISA); Student poster competition committee, IISA Meeting 2017, Hyderabad, India.

REFERENCES AVAILABLE TO CONTACT

Prof. David B. Dunson (e/mail: dunson@stat.duke.edu)
- Arts and Sciences Distinguished Professor, Department of Statistical Science, Duke University
- ⋆ *Prof. Dunson is my postdoctoral advisor.*

Prof. Nicholas G. Polson (e/mail: ngp@chicagobooth.edu)
- Professor of Econometrics and Statistics, Booth School of Business,University of Chicago
- ⋆ *Prof. Polson is my collaborator on a few ongoing projects.*

Dr. Anindya Bhadra (e/mail: bhadra@purdue.edu)
- Associate Professor, Department of Statistics, Purdue University
- ⋆ *Dr. Bhadra is my collaborator on a few ongoing projects.*

MORE INFORMATION

More information can be found at https://sites.uark.edu/jd033.
My Google Scholar Profile: http://bit.ly/1OTdd9h