Jyotishka Datta **April 16, 2024**

CONTACT INFORMATION

Department of Statistics
Virginia Tech
250 Drillfield Drive
Blacksburg, VA 24061

E-mail: jyotishka@vt.edu
Webpage: https://jyotishkadatta.wordpress.com/
Git: https://github.com/DattaHub

RESEARCH INTERESTS

Bayesian methodology and theory, Statistical Computing, Sparse signal recovery, Global-local shrinkage priors, Changepoint detection, Default Bayes, Discrete data, High-dimensional data, Geospatial data, Compositional data. Applications in Cancer Genomics, Epidemiology, Neuroscience, Bioinformatics, Criminology and Ecology.

PROFESSIONAL EXPERIENCE

January 2021 to present: Assistant Professor, Department of Statistics, Virginia Polytechnic Institute and State University, Blacksburg, VA.

- May 2022 to present: Core Faculty Member, Center of Biostatistics and Health Data Science, Virginia Polytechnic Institute and State University, Blacksburg, VA.

July 2016 to December 2020: Assistant Professor, Department of Mathematical Sciences, University of Arkansas, Fayetteville.

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

2008 - 2009: *Associate Manager*, Analytics, Credit Cards, Global Retail & Commercial Banking, India. Barclays Bank, PLC, Mumbai, India.

EDUCATION

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

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

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

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

AWARDS AND HONORS

- ***National and International***:
  - ***Dayanand Naik award*** from the ASA-VA chapter (Virginia chapter of American Statistical Association). 2023.
    'The Dayanand Naik award recognizes an individual (who works, or resides in Virginia, during the time of the award) for outstanding research contributions and service to the Commonwealth of Virginia in statistics and related fields.'
  - ***Lay Nam Chang Dean's Discovery Fund***, Virginia Tech, 2022-23.
  - ***Robert and Sandra Connor Endowed Faculty Fellowship***, University of Arkansas, 2018-19. News article.

- ***William J. Studden Publication Award*** for an outstanding publication in a mathematical statistics journal, 2013, Department of Statistics, Purdue University.
- ***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.
- ***Competitive 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
- ***Miscellaneous***:
  - 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

Google Scholar: (Citations: 2330, h-index: 15, i10 index: 20 as of April 16, 2024).

PAPERS IN REFEREED JOURNALS (SINCE VT)

[1] Sagar K. N., Banerjee, S., **Datta, J.**, and Bhadra A. (2024), "Maximum a Posteriori Estimation in Graphical Models Using Local Linear Approximation", In Press. ***STAT***.

[2] **Datta, J.**, Banerjee S., and Dunson D. (2023), "Nonparametric Bayes multiresolution testing for massive-dimensional rare events". ***Journal of Nonparametric Statistics***. 1–15. https://doi.org/10.1080/10485252.2024.2309978

[3] Bhadra, A., **Datta, J.**$^{+}$, Polson, N. G., Sokolov, V., Xu, J. (2023), (*alphabetical) "Merging Two Cultures: Deep and Statistical Learning". ***WIREs Computational Statistics.*** 16(2), e1647. https://doi.org/10.1002/wics.1647

[4] Sagar K. N., Banerjee, S., **Datta, J.**, and Bhadra A. (2023), "Precision Matrix Estimation under Horseshoe-like Penalty". **18 (1)**, 1-46. ***Electronic Journal of Statistics.*** https://doi.org/10.1214/23-EJS2196.

♠ Winner (Sagar K.N.) International Biometric Society Eastern North American Region's (ENAR) Distinguished Student Paper Awards, ENAR 2022.

[5] Kundu, R., **Datta, J**., Ray, D., Bhattacharyya, R., Mishra, S., Zimmermann, L., and Mukherjee, B. (2023), Comparative impact assessment of COVID-19 policy interventions in five South Asian countries using reported and estimated unreported death counts during 2020-2021, ***PLOS Global Public Health.*** **3(12)**: e0002063. https://doi.org/10.1371/journal.pgph.0002063.

[6] Boss, J., **Datta, J.**, Wang, X., Park, S., Kang, J., Mukherjee, B. (2023), "Group Inverse-Gamma Gamma Shrinkage for Sparse Regression with Block-Correlated Predictors". ***Bayesian Analysis***, **1(1)**, 1-30.

♠ Winner (Jonathan Boss), International Biometric Society Eastern North American Region's (ENAR) Distinguished Student Paper Awards, ENAR 2021.

[7] Jelesko, J.; Thompson, K.; Magerkorth, N.; Verteramo, E.; Becker, H.; Flowers, J.; Sachs, J.; **Datta, J.**; Metzgar, J. (2023), "Poison Ivy (*Toxicodendron radicans*) Leaf Shape Variability: Why Plant Avoidance-By-Identification Recommendations Likely Do Not Substantially Reduce Poison Ivy Rash Incidence." ***Plants, People, Planet.***, 6(1), 210–220. https://doi.org/10.1002/ppp3.10439.

[8] Ek, A., Drawve, G., Robinson, S., **Datta, J.** (2023), "Quantifying the Effect of Socio-Economic Predictors and the Built Environment on Mental Health Events in Little Rock, AR. " ***ISPRS International Journal of Geo-Information.***; **12(5)**: 205. open-access.

[9] Bhaduri, R., Kundu, R., Purkayastha, S., Kleinsasser, M., Beesley, L.,Mukherjee, B. and **Datta, J**. (2022), "Extending the Susceptible-Exposed-Infected-Removed (SEIR) model to handle the *fa*lse *n*egative rate and *sy*mptom-based administration of COVID-19 diagnostic tests: *SEIR-fansy*" ***Statistics in Medicine*.** https://doi.org/10.1002/sim.9357.

[10] Harris, C., Drawve, G., Thomas, S., **Datta, J.**, Steinman (2022): " Innovative Data in Communities and Crime Research: An Example at the Intersection of Racial Segregation, Neighborhood Permeability, and Crime", 1-18, ***Journal of Crime and Justice***.

[11] Chaudhuri, J., Biswas, S., Gangopadhyay, G., Biswas, T., **Datta, J.**, Biswas, A.; Datta, A., Mukherjee, A., Bhattacharya, P., Hazra, A. (2022). "Correlation of ATP7B gene mutations with clinical phenotype and radiological features in Indian Wilson Disease patients", **122 (1)**, 181-190, ***Acta Neurologica Belgica.***

[12] **Datta, J.,** and Mukherjee, B. (2021). "Discussion on "Regression Models for Understanding COVID-19 Epidemic Dynamics with Incomplete Data"", Invited discussion, ***Journal of American Statistical Association***. 116 (536), 1583-1586.

[13] Rezaeiahari, M.; Brown, C. C.; Ali, M. M.; **Datta, J**.; Tilford, J. M.; (2021) "Understanding Racial Disparities in Severe Maternal Morbidity Using Bayesian Network Analysis". *Accepted*, **PLoS One** ;16(10):e0259258. [full-text].

[14] Li, Y., **Datta, J.,** Craig, B.A., and Bhadra, A. (2021). "Joint mean–covariance estimation via the horseshoe". ***Journal of Multivariate Analysis***. 183 (2021): 104716.[preprint].

[15] Gu, X., Mukherjee, B., Das, S., **Datta, J.** (2021). "COVID-19 prediction in South Africa: Understanding the unascertained cases–the hidden part of the epidemiological iceberg". ***Journal of Statistical Research***.

♠ This was an **invited paper** for the special issue to celebrate 50-year independence of Bangladesh.

[16] Deshwal, A., Phan, P., **Datta, J.**, Kannan, R., Suresh Kumar, T.K., "A Meta-Analysis of the Protein Components in the Rattlesnake Venom". ***Toxins***, **13 (6)**, 372.

[17] Steinman, H., Drawve, G., **Datta, J.**, Harris, C. T., and Thomas, S. A. (2021): "Risky Business: Examining the 80-20 Rule in Relation to a RTM Framework". (Criminal Justice Review), **46 (1)**, 20-39.

PAPERS IN REFEREED JOURNALS (BEFORE VT)

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

[19] Bhadra, A., **Datta, J.**, Polson, N. G., & Willard, B. T (2020), (*alphabetical), "Global-local mixtures - A Unifying Framework". https://doi.org/10.1007/s13171-019-00191-2, ***Sankhya A*** *- J. K. Ghosh Memorial Issue*.

♠ Prof. Christian Robert covered this paper on one of his blog articles.

[20] Drawve, G., Harris, C., Thomas, S. A., **Datta, J.**, Cothren, J. (2020): "Current and New Frontiers: Exploring how Place Matters through Arkansas NIBRS Reporting Practices". (*C*rime & Delinquency), 67 (6-7), 941-969.

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

[22] Bhadra, A., **Datta, J.**, Polson, N. G., & Willard, B. T (2019), (*alphabetical), "Lasso Meets Horseshoe - A Survey" **34(3)**, 405-427. ***Statistical Science***. [full-text]

[23] Bhadra, A., **Datta, J.,** Polson, N. G., & Willard, B. T (2019), (*alphabetical), "Horseshoe Regularization for Feature Subset Selection". https://doi.org/10.1007/s13571-019-00217-7, ***Sankhya B***.

[24] Bhadra, A., **Datta, J.,** Polson, N. G., & Willard, B. T (2017), (*alphabetical) "The Horseshoe+ Estimator of Ultra-Sparse Signals", ***Bayesian Analysis***. **12 (4)**, 1105-1131. [full-text]

[25] Reddy, A., Zhang, J., Davis, N. S., Moffitt, A. B., Love, C. L., Waldrop, A., Leppa, S., Pasanen, A., Meriranta, L., Karjalainen-Lindsberg, M.-L., Nørgaard, P., Pedersen, M., Gang, A. O., Høgdall, E., Heavican, T. B., Lone, W., Iqbal, J., Qin, Q., Li, G., Kim, S. Y., Healy, J., Richards, K. L., Fedoriw, Y., Bernal-Mizrachi, L., Koff, J. L., Staton, A. D., Flowers, C. R., Paltiel, O., Goldschmidt, N., Calaminici, M., Clear, A., Gribben, J., Nguyen, E., Czader, M. B., Ondrejka, S. L., Collie, A., Hsi, E. D., Tse, E., Au-Yeung, R. K. H., Kwong, Y.-L., Srivastava, G., Choi, W. W. L., Evens, A. M., Pilichowska, M., Sengar, M., Reddy, N., Li, S., Chadburn, A., Gordon, L. I., Jaffe, E. S., Levy, S., Rempel, R., Tzeng, T., Happ, L. E., Dave, T., Rajagopalan, D., **Datta, J.**, Dunson, D. B., & Dave, S. S. (2017). Genetic and functional drivers of diffuse large B cell lymphoma. ***Cell***, 171(2), 481-494. https://doi.org/10.1016/j.cell.2017.09.027

♠ This article was featured on EurekAlert!, the newsletter from AAAS, link.

[26] Moffitt, A. B., Ondrejka, S. L., McKinney, M., Rempel, R. E., Goodlad, J. R., Teh, C. H., Leppa, S., Mannisto, S., Kovanen, P. E., Tse, E., Au-Yeung, R. K. H., Kwong, Y-L., Srivastava, G., Iqbal, J., Yu, J., Naresh, K., Villa, D., Gascoyne, R. D., Said, J., Czader, M. B., Chadburn, A., Richards, K. L., Rajagopalan, D., Davis, N. S., Smith, E. C., Palus, B. C., Tzeng, T. J., Healy, J. A., Lugar, P. L., **Datta, J.**, Love, C., Levy, S., Dunson, D. B., Zhuang, Y., Hsi, E. D., & Dave, S. S. (2017). "Enteropathy-associated T cell lymphoma subtypes are characterized by loss of function of SETD2", ***Journal of Experimental Medicine***, **214(5)**, 1371-86.

[27] McKinney, M., Moffitt, A. B., Gaulard, P., Travert, M., De Leval, L., Nicolae, A., Raffeld, M., Jaffe, E. S., Pittaluga, S., Xi, L., Heavican, T., Iqbal, J., Belhadj, K., Delfau-Larue, M. H., Fataccioli, V., Czader, M. B., Lossos, I. S., Chapman-Fredricks, J. R., Richards, K. L., Fedoriw, Y., Ondrejka, S. L., Hsi, E. D., Low, L., Weisenburger, D., Chan, W. C., Mehta-Shah, N., Horwitz, S., Bernal-Mizrachi, L., Flowers, C. R., Beaven, A. W., Parihar, M., Baseggio, L., Parrens, M., Moreau, A., Sujobert, P., Pilichowska, M., Evens, A. M., Chadburn, A., Au-Yeung, R. K. H., Srivastava, G., Choi, W. W. L., Goodlad, J. R., Aurer, I., Basic-Kinda, S., Gascoyne, R. D., Davis, N. S., Li, G., Zhang, J., Rajagopalan, D., Reddy, A., Love, C., Levy, S., Zhuang, Y., **Datta, J.,** Dunson, D. B., & Dave, S. S. (2017) "The Genetic Basis of Hepatosplenic T Cell Lymphoma". ***Cancer Discovery***, **CD-16-0330**.

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

[29] Healy, J. A., Nugent, A., Rempel, R. E., Moffitt, A. B., Davis, N. S., Jiang, X., Shingleton, J. R., Zhang, J., Love, C., **Datta, J.**, McKinney, M. E., Zhang, T. J., Wettschureck, N., Offermanns, S., Walzer, K. A., Chi, J-T., Rasheed, S. A. K., Casey, P. J., Lossos, I. S., & Dave, S. S.(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.

[30] Bhadra, A., **Datta, J.**, Polson, N. G., & Willard, B. T (2016), (*alphabetical) "Default Bayesian analysis with global-local shrinkage priors", ***Biometrika***, **103 (4)**: 955-969. [full-text]

[31] Chaudhuri, J., Biswas, T., **Datta, J.,** Sabui, T.K., Chatterjee, S., Ray, S., Raychaudhuri, D., Mandal, K., Chatterjee, K. and Chakraborty, S., (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.

[32] Libohova, Z., Winzeler, H. E., Lee, B., Schoeneberger, P. J., **Datta, J.,** and Owens, P. R. (2016). "Geomorphons: Landform and property predictions in a glacial moraine in Indiana landscapes". ***Catena***, **142**, 66-76.

[33] Parthasarathy, **Datta, J.**, 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.

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

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

BOOK CHAPTERS SINCE VT

[1] Young, S., **Datta, J.**, Kar, B., Huang, X., Williamson, M., Tullis, J., and Cothren, J. (2021), "Challenges and limitations of geospatial data and analyses in the context of COVID-19". *"Human Dynamics in Smart Cities"*, Springer.

BOOK CHAPTERS BEFORE VT

[2] **Datta, J.** and Ghosh, J. K. (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.

[3] Dasgupta, R., Ghosh, J. K., Chakravarty, S., and **Datta, J.** (2015), "Some Remarks on Pseudo Panel Data". *Growth Curve and Structural Equation Modeling*, 25-34. Springer International.

REFEREED CONFERENCE PROCEEDINGS

[1] Chakraborty, Verma, Sahoo, and **Datta, J.** (2020), "FairMixRep: Self-supervised Robust Representation Learning for Heterogeneous Data with Fairness constraints", IEEE International Conference on Data Mining Workshop (ICDMW). 2020. preprint.

[2] LeBow V., Bernhardt-Barry, M. L., and **Datta, J.** (2018), "Improving Spatial Visualization Abilities Using 3D Printed Blocks". 2018 ASEE Annual Conference & Exposition , Salt Lake City, Utah. full-text.

OTHER PUBLICATIONS

[1] **Datta, J.**, "*Vox Populi* : Exploring the Strengths and Limits of Collective Decision-Making", Part of a quarterly column 'Datta's Data Dives', IISA Newsletter, April, 2024.

[2] **Datta, J.**, "Paradoxes", Part of a quarterly column 'Datta's Data Dives', IISA Newsletter, December, 2023.

[3] **Datta, J.** and Drawve, G., "Does Machine Learning Reduce Racial Disparities in Policing?", IISA Newsletter, December, 2016.

[4] **Datta, J.** and Ghosh, J.K., "Optimal Objective Priors for Linear Models", Indian Bayesian Society Newsletter, Vol XI, No. 1, May, 2014.

PAPERS UNDER REVIEW

[1] **Datta, J.**, and Bandopadhyay, D. (202x), "Shrinkage and Selection for Compositional Data". Submitted to *Journal of the Indian Society for Probability and Statistics*.

[2] **Datta, J**. and Polson N. (202x). "Quantile Importance Sampling". Revision invited at *Brazilian Journal of Probability and Statistics.* pre-print.

[3] **Datta, J**. and Polson N. (202x). "Inverse Probability Weighting: the Missing Link between Survey Sampling and Evidence Estimation". pre-print.

[4] Guha, N. and **Datta, J.** (202x), "Consistent Model Selection and Change Point Recovery for High-dimensional Changing Linear Regression". pre-print.

[5] Bhadra A., Sagar K. N., Banerjee, S., and **Datta, J.** (202x), "Graphical *Evidence*". pre-print.

MANUSCRIPTS IN PREPARATION

[1] **Datta, J.**, Heiner, M., Ovaskainen, O. and Dunson, D.B. (202x), "Sparse generalized Dirichlet distributions for high-dimensional probabilities".

[2] Sengupta, S., **Datta, J.**, Chen, Y. (202x), "Proximity Block-models for Network Data".

[3] Guha, N. and **Datta, J.** (202x), "A Random Projection Based Technique for Change Point Estimation in Ultra-high Dimension".

[4] Nichols, Q., **Datta, J.**, Davy, B., Davy, K., Hanlon, A., Volpe, S. (202x), "Predicting Hemoglobin A1c Percentage of Asian-American Adults, from Body Composition, Sedentary Behavior, and Dietary Magnesium Intake, using 2011 to 2020 NHANES Data".

[5] Nichols, Q., **Datta, J.**, Davy, B., Davy, K., Hanlon, A., Volpe, S. (202x), "Predicting Hemoglobin A1c 6.5% in Asian-American Adults: Evaluating the Role of Body Composition, Sedentary Behavior, and Dietary Magnesium Intake Using Machine Learning, Decision Tree, and Logistic Regression Approaches with NHANES 2011 to 2020 Data".

FUNDING

External [**key**: ^{vt} - while at VT; ^{ua} - while at the UoA.]

G1 Predict Align Prevent: Developing a spatial child maltreatment risk machine learning model for the State of Texas ("Geographic Location")^{vt}. **Upbring Inc.** (via University of Arkansas), PI: Drawve, G., Role: PI of subcontract. (1/2023 - 3/2024). Amount: $44,694 (subcontract), $85,991 (Total), Percent credit: 100%.

G2 Collaboration with VA's FAACT Project to inform Real-world Research and Translation^{vt}. **Qlarion Inc.**, Role: Co-PI. (8/2022 - 10/2022). Amount: $23,332. Percent credit 0%, 1 summer month for Datta.

G3 Shrinkage on Simplex: Bayesian Inference for Sparse Compositional Data^{vt}. Lay Nam Chang Dean's Discovery Fund. **Virginia Tech Foundation Inc.**, Role: PI, (5/2022 - 6/2025). Amount: $15,592, Percent credit: 100%.

G4 New Directions in Bayesian Change-point Analysis^{vt}. **NSF-DMS**-2015460, Statistics (via University of Massachusetts), PI: Guha, N., Role: PI of subcontract. (8/2020 - 7/2024). Amount: $43,891 (subcontract), $139,984 (Total), Percent credit: 100%.

G5 Spring Lecture Series 2019-2020^{ua}, **NSF-DMS**-1853458, Statistics, Role: co-PI, (4/2019 - 5/2022). Amount: $9,956. (Percent credit: 0%, Conference grant).

G6 Child Maltreatment in Little Rock: Aligning Services with Risk^{ua}, **Arkansas Children's Trust Fund**, Role: co-PI, (10/2019 - 3/2019). Amount: $ 20,000. Percent credit: 33.33%.

G7 Child Maltreatment Pilot Project in Little Rock, Arkansas^{ua}. **Arkansas Children's Trust Fund**, Role: co-PI, (1/2019 - 8/2019). Amount: $ 27,000. Percent credit: 33.33%.

Internal

G8 Faculty Mentoring Grant^{vt}. **Virginia Tech, Office of the Vice Provost for Faculty Affairs**, Role: PI, (3/2021 - 8/2023.) Amount: $ 1,500, Percent credit: 100%.

G9 Robert and Sandra Connor Endowed Faculty Fellowship^{ua}, **University of Arkansas**, Role: PI, (8/2018- 8/2019). Amount: $ 5,000. Percent credit: 100%.

G10 Predicting Soil Type from Non-destructive Geophysical Data^{ua}, **University of Arkansas, Provost's Collaborative Research Grant**, Role: co-PI, (12/2018 - 12/2019). Amount: $ 2,200. Percent credit: 100%.

G11 Participant Field Training with Little Rock Police Department^{ua}. **University of Arkansas, Provost's Collaborative Research Grant**, Role: co-PI, (11/2017- 12/2018). Amount: $ 2,000. Percent credit: 33.33%.

G12 Multiresolution Nonparametric Bayesian Hotspot Detection^{ua}, **University of Arkansas, Provost's Collaborative Research Grant**, Role: co-PI, (11/2016 - 12/2017). Amount: $ 2,000. Percent credit: 100%.

DEPARTMENTAL SEMINARS

Convention: ♣ Virtual presentation.

Since VT

1. *Global-Local Shrinkage Priors: An Overview and New Directions*, Department of Statistics, Brigham Young University. March 21, 2024.
2. *Recent Developments in Bayesian Shrinkage for Sparse and Structured Data*. Invited Lecture under the IIM Indore Seminar Series. Indian Institute of Management, Indore. Madhya Pradesh, India. December 28, 2023.
3. *Recent Developments in Bayesian Shrinkage for Sparse and Structured Data*. Department of Statistics, North Carolina State University. November 3, 2023. (Followed by a discussion by Prof. Ryan Martin.)
4. *Recent Developments in Bayesian Shrinkage for Sparse and Structured Data*. INFORMS student Colloquium, Department of Industrial Engineering, Virginia Tech. December 2, 2022.
5. *Recent Developments in Bayesian Shrinkage for Sparse and Structured Data*. Department of Mathematics and Statistics, University of Maryland Baltimore County, October 7, 2022. ♣.
6. *Recent Developments in Bayesian Shrinkage for Sparse and Structured Data.* Department of Statistics, University of Georgia. September 29, 2022.
7. *Recent Developments in Bayesian Shrinkage for Sparse and Structured Data*. Department of Statistics, University of Florida. September 22, 2022.
8. *New Directions in Bayesian Shrinkage for Sparse and Structured Data*. B3D (Biostatistics - Biomedical Informatics - Big Data) Seminar Series, Department of Biostatistics at the Harvard T.H. Chan School of Public Health. April 4, 2022. ♣.
9. *New Directions in Bayesian Shrinkage for Sparse and Structured Data*, Department of Statistics, University of Connecticut, February 2, 2022. ♣.
10. *New Directions in Bayesian Shrinkage for Sparse, Structured Data*. Applied Statistics Workshop. University of Tokyo, October 27, 2021. ♣.

11. *New Directions in Bayesian Shrinkage for Structured Data*, Richard F. Barry Seminar, Department of Mathematics and Statistics, Old Dominion University. April 29, 2021. ♣.

Before VT

12. *New Directions in Bayesian Shrinkage for Structured Data*, Department of Statistics and Actuarial Science, The University of Iowa, September 10, 2020. ♣
13. *Bayesian Shrinkage for Continuous and Discrete Data, a Tale of Two Cities*. Department of Biostatistics, University of Michigan, Ann Arbor., May 28, 2019.
14. *Bayesian Sparse Signal Recovery, Horseshoe Regularization*. Department of Statistics, Florida State University. April 2018.
15. *Sparse Signal Recovery & Default Bayesian Analysis*. Applied Statistics Unit, Indian Statistical Institute, Kolkata. January 2017.
16. *Sparse Signal Recovery for Discrete & Continuous Data*, University of Arkansas, Fayetteville.
17. *Sparse Signal Recovery for Discrete & Continuous Data*, Clemson University.
18. *Sparse Signal Recovery for Discrete & Continuous Data*, Binghamton University.
19. *Sparse and Ultra-Sparse Signal Recovery : Horseshoe and Horseshoe+ prior*. Department of Statistical Science, Duke University. August, 2014.
20. *Large Scale Hypothesis Testing for discrete and continuous data*, University of Texas - M. D. Anderson Cancer Center, Houston, TX. January, 2014.
21. *Two Groups and One Group Models for Multiple Testing*, Mathematical Statistics Seminar, Purdue University. November, 2013.
22. *Two Groups and One Group Models for Multiple Testing*, National Institute of Biomedical Genomics, Kalyani, India. May, 2013.

INVITED TALKS AT CONFERENCES

Convention: ♣: Virtual presentation.
Acronyms: IISA: International Indian Statistical Association, ICSA: International Chinese . tatistical Association, JSM: Joint Statistical Meeting, ISBA: International Society of Bayesian Analysis, ENAR: Eastern North American Region. International Biometric Society.

Since VT

1. *Quantile Importance Sampling*, December 16, 2023: Econometrics and Statistics (EcoSta) Conference, Berlin, Germany ♣.
2. *New directions in Bayesian shrinkage for structure learning*, August 8, 2023: Joint Statistical Meeting (JSM), Toronto, Canada.
3. *Structure Learning with Global-Local Prior-Penalty Dual*. August 2, 2023: Econometerics and Statistics (EcoSta) 2023, Tokyo, Japan ♣.
4. *Quantile Importance Sampling*, June 6, 2023: New England Statistical Society (NESS) Symposium, Boston, MA ♣.
5. *Structure Learning with Global-Local Prior-Penalty Dual*. June 2, 2023. Society of Industrial and Applied Mathematicians (SIAM) Conference on Optimization (OP23), Seattle, WA.
6. *New Directions in Bayesian Shrinkage for Sparse, Structured Data,* December 26, 2022: IISA International Conference on Statistics at IISc Bangalore, India ♣.
7. *New Directions in Bayesian Shrinkage for Structure Learning.* August 11, 2022. Joint Statistical Meeting, 2022, Session title: "Bayesian penalized likelihood methods for Gaussian graphical models". Washington, DC.

8. *New Directions in Bayesian Shrinkage for Structure Learning.* July 8-9, 2022, 6th EAC-ISBA (Eastern Asia Chapter) Conference ♣.
9. *Precision Matrix Estimation under the Horseshoe-like Prior-Penalty Dual*. May 4, 2022: UP-STAT 2022 Conference in Biostatistics, at University at Buffalo ♣.
10. *Shrinkage on Simplex: Bayesian Inference for Sparse and Structured Compositional Data*. December 20, 2021. CMStatistics 2021. Invited Session: EO402: Bayesian methods in structured data and high-dimensional problems. ♣.
11. *New Directions in Bayesian Shrinkage for Structured Data*. September 14, 2021: Invited Session, ICSA 2021: Invited Session 82: Flexible and efficient Bayesian methods for complex data modeling. ♣.
12. *New Directions in Bayesian Shrinkage for Sparse, Structured Data*, September 8, 2021: International Indian Statistical Association Statistics and Data Science Innovations Series Webinar. ♣.
13. *Shrinkage on Simplex: Quantifying Sparsity and Dependence in Compositional Data with a Bayesian framework*. June 29-July 3, 2021: Invited Session, ISBA 2020 World Meeting, Kunming, China♣.

Before VT

14. *Bayesian Shrinkage for Continuous & Discrete Data– a Tale of Two Cities*, Invited Talk, ICSA Applied Statistics Symposium 2020. Houston, TX. December 2020. ♣
15. *Bayesian methods in structured data and high dimensional problem: some recent advances*. Invited Session, JSM 2020, Philadelphia, PA. August 1-6, 2020. ♣
16. *Sparse Generalized Dirichlet Distributions for Microbiome Compositional Data*. Invited Talk in Session "Innovative Statistical Approaches for High-Dimensional Omic and Microbiomic Data", ENAR 2020, Nashville, Tennessee. March 2020. ♣
17. *Bayesian Shrinkage for Continuous & Discrete Data – a Tale of Two Cities*. Invited Talk, Session "Bayesian Modeling and Computation", IISA 2019 Conference, Mumbai, India. December 2019.
18. *Bayesian Sparse Signal Recovery: Gaussian Models and Beyond*, August 2019: Special Invited Session in Memory of Prof. J.K. Ghosh, JSM 2019, Denver, Colorado.
19. *Nonparametric Bayes Multiresolution Testing for Detecting Rare Variants*. Invited Talk in Session: "Innovative Approaches for High-dimensional Omics and Neuroimaging Data" in Joint Statistical Meeting, Denver, Colorado. August 2019.
20. *Sparse Signal Recovery using Global-Local Shrinkage Priors.* Invited Session (Multiple Testing) in Young Statisticians' Meet: Data Science in Action, Indian Statistical Institute, Kolkata, India. January 4-5, 2019.
21. ♠ *New Directions in Bayesian Sparse Signal Recovery*. **Plenary Session** in Tenth International Calcutta Triennial Symposium, December 27-30, 2018, Kolkata, India. December 28, 2018.
22. *Horseshoe Regularization for Feature Subset Selection*. IISA International Conference on Statistics at Hyderabad, India. December 2017.
23. *Horseshoe Regularization for Feature Subset Selection*. ERCIM WG Meeting, CMStatistics 2017 Conference at London, UK. December 2017.
24. *Detecting rare mutational hotspots by multiscale BNP method.* JSM 2017, Joint Statistical Meeting, Baltimore, Maryland. August 2017.
25. *Default Bayesian analysis for global-local shrinkage priors*, August, 2016: IISA 2016 Conference, Corvallis, Oregon.

26. *Shrinkage Priors for High-Dimensional Sparse Poisson Means*, August 2016: JSM 2016, Joint Statistical Meeting, Chicago, Illinois.
27. *Multiscale Bayesian cluster detection and testing for whole genome sequencing studies*. Transition workshop for “Beyond Bioinformatics", SAMSI, North Carolina. May, 2015.

CONTRIBUTED TALKS AND POSTERS

Before VT

1. *Sparse Signal Recovery for Discrete & Continuous Data* and *Detecting rare mutational hotspots by multiscale BNP method. Talk*, Departmental seminar, University of Arkansas, Fayetteville. September, 2016.
2. *Shrinkage Priors for High-Dimensional Sparse Poisson Means*, *Poster presentation*, John W. Tukey 100^{th} Birthday Celebration at Princeton University. September, 2015.
3. *Shrinkage Priors for Sparse High-Dimensional Discrete or Continuous Data*. *Talk*, SAMSI postdoc seminar. September, 2015.
4. *Bayesian Cluster Detection for Rare Variants*. *Poster Presentation*, SAHD (Sensing and Analysis of High Dimensional Data Workshop), Duke University, Durham, NC. July, 2015.
5. *Multiscale Bayesian cluster detection and testing for whole genome sequencing studies*. *Poster presentation*, SRCOS (Southern Research Conference), Carolina Beach, NC. June, 2015.
6. *Multiresolution nonparametric Bayesian cluster detection and association testing for whole genome sequencing studies*, *Poster presentation*, CCPS (Cancer Control and Population Sciences Fair), Duke University, NC. May, 2015.
7. *Multiresolution nonparametric Bayesian cluster detection and association testing for whole genome sequencing studies with applications in CVID. Poster presentation*, The Biology of Genomes Meeting, Cold Spring Harbor Lab, NY. May, 2015.
8. *Ultra-Sparse Signal Recovery through the Horseshoe+ Prior. Talk*, SAMSI. September, 2014.
9. *In Search of Optimal Objective Priors for Model Selection and Estimation. Poster presentation*, O-Bayes 2013, Duke University. December 2013.
10. *Two-groups and One-Group Models for Multiple Testing. Talk*, Machine Learning Seminar, Department of Computer Science, Purdue University. March, 2013.
11. *Asymptotic properties of Bayes risk for the Horseshoe prior. Talk*, Graduate Student Organization Seminar, Department of Statistics, Purdue University. October, 2012.

OUTREACH ACTIVITIES

Outreach/extension talks:

1. October 10, 2023: R programming workshop for students from Radford City High School. Link to materials: https://dattahub.github.io/beginneR/.
2. December 15, 2022: Invited Guest Lecture at Radford City High School - Data Visualization.
3. October 21, 2022: Mu-Sigma-Rho Student Seminar at Department of Statistics, Virginia Tech - ‘Overview of Bayesian Shrinkage’.
4. October 20, 2022: Invited Guest Lecture at Radford City High School - ‘Famous Wins and Failures in Statistics’.

5. February 5, 2022: Invited Special Lecture as part of India at 75 - 'My Statistics - My Story'- Lecture Series at NMIMS University, Sunandan Divatia School of Science.
6. July 9, 2021: "Selection and Shrinkage for Continuous And Discrete Data - A Brief Overview". Lecture for undergraduate students in India for Cheenta School of Statistics and Analytics. The video is available on YouTube: https://www.youtube.com/watch?v=fz9hnQE-wls.
7. May 14, 2021: "Understanding Vaccine Efficacy and Effectiveness: A Statistician's Perspective" for the University of Arkansas Honors College course "Vaccine" in May 2021. The video and the slides are available at: https://scholarworks.uark.edu/hnrcvac/7/ and on YouTube: https://www.youtube.com/watch?v=Jrqv59iyLUE.

Committee Service:

8. January 2021 - present: International Indian Statistical Association's Membership And Outreach Committee , Description: I am a member of the International Indian Statistical Association's Membership And Outreach Committee since January 2021. In 2022, Dr. Shariq Mohammed from Boston University and I co-organized the Statistics And Data Science Innovations Webinar Series with different experts' from Industry and Academia. Committee Role: Co-chair.

MENTORING

Graduate Students

Virginia Tech $^{+}$: Chair. *: Committee Member.

- Currently Advising:
  1. Jie Min * (PhD, Statistics).
  2. David Edwards * (MS, PhD, Statistics).
  3. Phil Geun Jin * (PhD, Statistics).
  4. Deepak Kumar * (PhD, Agricultural and Applied Economics).
  5. Michael White * (PhD, Human Nutrition, Foods, and Exercise).
  6. Mohammed Al rezq * (PhD, Industrial and System Engineering)
  7. Bright Samson * (MA, DAAS).
  8. Josiah Gilbert * (MS, Statistics).
  9. Won Hee Kim * (MS, Statistics).
- Graduated:
  1. Sakshar Desai $^{+}$ (MS, Statistics)
  2. Christopher Grubbs * (PhD, Statistics, Graduated August 2023.).
  3. Mohammed Ba-Aoum * (PhD, Industrial and System Engineering, Graduated April 2024)
  4. Katia Tarkhan $^{+}$ (MA, DAAS).
  5. Anna Flowers * (MS, Statistics).
  6. Eric Larsson * (MS, Statistics).

University of Arkansas

- Primary Advisor [MS, STAN (Statistics and Analytics) program]:
  1. Ek Alfieri 'Quantifying the Simultaneous Effect of Socio-Economic Predictors and Build Environment on Spatial Crime Trends'. University of Arkansas (Graduated Summer '23),
  2. Apu Chandra Das. 'Effect of Predictor Dependence on Variable Selection for Linear and Log-Linear Regression'. University of Nebraska Medical Center. (Graduated Fall '20)

3. Mohamed Abdelkader Abba. 'Adapting to Sparsity and Heavy Tailed Data'. North Carolina State University. (Graduated Summer 2018).
4. Josh Price. 'Effect of Cross-Validation on the Output of Multiple Testing Procedures'. Watcher Inc. (Graduated Summer 2019).
5. Kai Cui. ServiceNow. (Graduated Fall 2018).

- Committee Member [MS, STAN program, unless indicated in parentheses.]
  1. Nana Amma Asamoah
  2. April Walker
  3. Md Abul Hayat
  4. Hanna Steinman (Criminology)
  5. Sho-Hsien Su
  6. Waltram Ravelombola
  7. Anne Lin
  8. Ji Li
  9. Michael Ellis
  10. James Willbanks
  11. Ruizhe Yin
  12. Shanshan Zhang
  13. Mahboubeh Madadi
  14. 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)

Undergraduate Students

University of Arkansas

- Honors Thesis Advisor: Kelvin Feng.
  Thesis title: 'Bayesian Grouped Variable Selection: A Case Study'.
- Honors Thesis Committee:
  1. Vanessa Lebow
  2. Winson Chee
  3. Dhruba Dasgupta
  4. Christopher Peterson
- Academic Advising (Math):
  1. Jodi Mitchell
  2. Bruce Dunning
  3. Alex Coleman
  4. Rosario Dispensa
  5. Kaylee Henry
  6. David O'Hearn
  7. Lauren Pearce

TEACHING EXPERIENCE SINCE VT

- Spring 2021 - now, Department of Statistics, **Virginia Tech**.
  - CMDA 2006. Integrated Quantitative Science (Statistics part).
  - CMDA 2014. Data Matter. Undergraduate course on exploring different kinds of data (numerical, qualitative, text and image).

- STAT 5525. Data Analytics. Graduate course on different tools and techniques for drawing meaningful inference from data, with a comprehensive review of popular Statistics/ML methods.
- CMDA 4654: Intermediate Data Analytics and ML: Undergraduate course on popular tools for analyzing data and modern Statistical and ML methods.

Teaching Experience Before VT

- Fall 2016 - 2020, Department of Mathematical Sciences, **University of Arkansas**. Teaching duties: 2 + 1 courses for first two years, then 2 + 2 courses third year onward.
  - 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, 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 2011-Spring 2013, 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.
  - 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, 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.

Professional Service Since VT

- **Department Level**: Biostatistics Faculty Search Committee (Fall 2023), Policies and Procedures Committee (Fall 2023), Colloquium Committee (Fall 2021, Spring 2022).
- Co-organized (as the local organizing committee) the *Pushing the Boundary of Data Science through Statistical Modeling and Inference*, a conference in honor of the 70th birthday of Prof. Dipak K. Dey at Blacksburg, VA Link to the Conference Homepage.
- Served as the lead organizer and chair of the program committee for *International Indian Statistical Association Virtual Mini-Conference* in January 2022.
- **Associate Editor** for *Sankhya Series A*, quarterly journal focused on publishing research articles in Mathematical Statistics and Probability, (January 2022 to present).
- **Guest Associate Editor** for the special issue on “Pushing the Boundary of Data Science through Statistical Modelling and Inference" for *New England Journal of Statistics in Data Science* (August 2023 to present)

- Served as a member of the Scientific Program Committee for the following:
  - 2024 annual conference of the *International Indian Statistical Association* (IISA). 27th-31st December 2024, Cochin University of Science and Technology (CUSAT) in Kochi, Kerala, India.
  - *36th New England Statistics Symposium* 2023, Boston, MA. Organized an invited session entitled Bayesian Methods for High-Dimensional Data with Low-Dimensional Structures. (Speakers: Matt Heiner, Maoran Xu, Marco Ferreira (Jake Williams), Jyotishka Datta.)
- Organized the following invited or topic-contributed sessions:
  - *Bayesian Inference for Sparse and Structured Data: Recent Advances* at the *Joint Statistical Meeting* 2024, Portland, Oregon. (Speakers: Malay Ghosh, Christine B. Peterson, Ismael Castillo, Anindya Bhadra Chair: Sayantan Banerjee.)
  - Recent Advances in Bayesian Methods for Complex Structured Data sponsored by SBSS and IISA at *Joint Statistical Meeting* 2023, Toronto, CA. (Speakers: Daniel Kowal, Ray Bai, Antik Chakraborty, Chris Franck, Jyotishka Datta. Chair: Thomas Metzger.)
  - Recent theoretical and methodological advances in high-dimensional inference for *International Indian Statistical Association Conference* (IISA) 2022, Bangalore, India.
- Co-organized and hosted the IISA *Data Science and Statistics Innovation Webinar* Series. Link to the Playlist.
- Served as a reviewer for the following journals: *Journal of Royal Statistical Society (B), Annals of Applied Statistics, Biometrika, Journal of American Statistical Association (Theory and Methods + Application and Case Studies), Journal of Multivariate Analysis; Statistica Sinica, Bayesian Analysis, Bernoulli, Electronic Journal of Statistics, Operation Research, Computational Statistics, Sankhya Series A and B, Entropy, Statistics in Medicine, Journal of Statistical Computation and Simulation, PLoS One.*
- Served as a reviewer for **Machine Learning Conferences:** NIPS, ICML, AIStats.

Professional Service Before VT

- Co-organized the **Spring Lecture Series** 2019 and 2020, Department of Mathematical Sciences at the University of Arkansas. link to SLS webpage.
  - Spring Lecture Series 2019: Principal Speaker: Mike West, Conference theme: "Bayesian Analysis for Multivariate Dynamic Systems: Decouple/Recouple Concept and Strategies". April 18-20, 2019.
  - Spring Lecture Series 2020: Principal Speaker: Igor Prünster, Conference theme: "Discrete Random Structure in Bayesian Nonparametrics", November 10-13, 2020.
- Served as an ad-hoc proposal reviewer for National Science Foundation (2017).
- Organized the following invited or 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.

- Co-founded the University of Arkansas R group with Dr. Grant Drawve for faculty/staff/students.

- **External Committee Service:**
  - Member of Executive Committee and Newsletter Editor-in-Chief, International Indian Statistical Association (IISA) (2017-2020);
  - Student poster competition committee, IISA Meeting 2017, Hyderabad, India. ICSA Applied Statistics Conference 2020.

Corporate Internship Experience

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.

Popular Science Books

[6] Datta, J. (2023). "Jodubabur Tuitiony - Probability o Paradoxer golpo" (Tr. Jodubabu's tuition: tales of probability and paradox.) (1 ed.). Kolkata, India: Language: Bengali. Published by: Guruchandali. http://guruchandali.com/ ISBN: 978-93-92706-30-1.

Description: This is a popular science book in Bengali with an aim to explain statistical concepts and methods in an accessible way, using real-world examples. A bilingual (English-Bengali) review by Prof. Partha Pratim Majumder, National Science Chair, Science and Engineering Board, Government of India ; and Distinguished Professor, John C. Martin Centre for Liver Research and Innovations is here: https://www.guruchandali.com/comment.php?topic=27767

More Information

More information can be found at https://jyotishkadatta.wordpress.com/.