Saptarshi Chakraborty

Curriculum Vitae

CONTACT INFORMATION

Department of Biostatistics
School of Public Health and Health Professions
State University of New York at Buffalo
718 Kimball Tower
Buffalo, NY 14214, USA

EDUCATION AND TRAINING

2020 **Postdoctoral Research in Statistical Genomics**, *Memorial Sloan Kettering Cancer Center*, New York, NY, USA.

2018 **Ph.D. in Statistics**, *University of Florida*, Gainesville, Florida, USA.

2013 **M.S. in Statistics**, *Indian Statistical Institute*, Kolkata, India.

2011 **B.Sc. (Hons.) in Statistics**, *Presidency College*, Kolkata, India.

EMPLOYMENT

Since 2020 **Assistant Professor of Biostatistics (Tenure Track)**, *State University of New York at Buffalo*, Buffalo, NY, USA.

RESEARCH INTERESTS

Computational Statistics, Bayesian modeling, statistical modeling in cancer genomics and computational biology, Markov chain Monte Carlo, statistical machine learning, big and high-dimensional data, dimension reduction, statistical software development, drug-safety, electronic health-record data, statistical consultation in biomedical research.

REFEREED PUBLICATIONS

Statistics in Genomics and Computational Biology

1. **Chakraborty, S**., Martin, A., Guan, Z., Begg, C. B., and Shen, R. (2021). Mining mutation contexts across the cancer genome to map tumor site of origin. *Nat Commun* **12,** 3051. Link.
2. **Chakraborty, S.**, Ecker, B. L., Seier, K., Aveson, V. G., Balachandran, V. P., Drebin, J. A., D'Angelica, M. I., Kingham, T. P., Sigel, C. S., Soares, K. C., Vakiani, E., Wei, A. C., Chandwani, R., Gonen, M., Shen, R., Jarnagin, W. R. (2021). Genome-derived Classification Signature for Ampullary Adenocarcinoma to Improve Clinical Cancer Care. *Clinical Cancer Research.* **(27)** (21) 5891-5899. Link.
3. **Chakraborty, S.** Tian, L, Tseng, Y, and Wong, S. W. (2021). Bayesian analysis of coupled cellular and nuclear trajectories for cell migration. *Biometrics* 1-12. Link.
4. **Chakraborty, S.**, Begg, C. B., and Shen, R. (2020). Using the "Hidden" Genome to Improve Classification of Cancer Types. *Biometrics. 2020;1–11.* Link. arXiv.
5. **Chakraborty, S.**, Arora A., Begg, C. B. and Shen, R. (2019). Using Somatic Variant Richness to Mine Signals from Rare Variants in the Cancer Genome. *Nat Commun* **10**, 5506. Link.

Markov Chain Monte Carlo Theory

6. **Chakraborty, S.**, Bhattacharya, B., and Khare, K. (2022). Estimating accuracy of the MCMC variance estimator: asymptotic normality for batch means estimators. *To appear in Statistics and Probability Letters.* arXiv.
7. **Chakraborty, S.** and Khare, K. (2019). Consistent estimation of the spectrum of trace class data augmentation algorithms. *Bernoulli.* 25(4B), 2019, 3832–3863. arXiv. Link.
8. **Chakraborty, S.** and Khare, K. (2017). Convergence properties of Gibbs samplers for Bayesian probit regression with proper priors, *Electronic Journal of Statistics* 11, 177-210. Link.

Multivariate Analysis and Robust Statistical Methods

9. Lee, M., **Chakraborty, S.**, and Su, Z. (2021). A Bayesian approach to envelope quantile regression. *To appear in Statistica Sinica.* Link.
10. Maji, A., **Chakraborty, S.**, and Basu, A., (2017). Statistical Inference based on the Logarithmic Power Divergence. *Society For Application Of Statistics And Allied Sciences*, 2, 39–51. Link

Computational Research

11. **Chakraborty, S.** and Wong, S. W. (2021). BAMBI: An R package for Fitting Bivariate Angular Mixture Models. *Journal of Statistical Software*, *99*(11), 1–69. Link.

Collaborative Applied Research

12. Ambruster, C.E., Brauer, A.L., Humby, M.S., Shao, J., **Chakraborty, S.** (2021). Prospective assessment of catheter-associated bacteriuria in nursing home residents: clinical presentation, epidemiology, and colonization dynamics. *JCI Insight.* Oct 8;6(19):e144775. Link.
13. Cassidy, D. J., **Chakraborty, S.**, Panda, N., McKinley, S. K., Mansur, A., Hamdi I., Mullen, J., Petrusa, E., Phitayakorn, R., and Gee, D. (2020). The Surgical Knowledge "Growth Curve": Predicting ABSITE Scores and Identifying "At-Risk" Residents. *Journal of Surgical Education.* Link.
14. Barnard, A. M., Wilcox, R., Forbes, S.C., Daniels, M. J., **Chakraborty, S.**, Lott, D., J., Senesac, C. R., Arora, H., Sweeny, L., Walter, G. H., and Vandenborne, K. H. E. (2020). MR biomarkers predict clinical function in Duchenne muscular dystrophy. *Neurology*, *94*(9), e897-e909. Link.
15. Rooney, W. D., Berlow, Y. A., Triplett, W. T., Forbes, S. C., Willcocks, R. J., Wang, D, Arora, H, Senesac, C, Lott, D. J., Finkel, R., Russman, B. S., Finanger, E. L., **Chakraborty, S.**, O'Brien, E, Moloney, B, Barnard, A, Sweeney, H. L., Daniels, M. J., Walter, G. A., and Vandenborne, K. (2020). Modeling disease trajectory in Duchenne muscular dystrophy. *Neurology*, *94*(15), e1622-e1633. Link.
16. Vaziri, S., Awan, O., Porche, K., Scott, K., Sacks, P., Dru, A. B., **Chakraborty, S.**, Khare, K., Hoh, B., and Rahman, M. (2019). Reimbursement Patterns for Neurosurgery: Analysis of the NERVES Survey Results from 2011-2016. *Clinical Neurology and Neurosurgery*, p.105406. Link.
17. Chatterjee, N., Nair, P.K.R., **Chakraborty, S.**, and Nair, V.D. (2018). Changes in soil carbon stocks across the Forest-Agroforest-Agriculture/Pasture continuum in various agroecological regions: A meta-analysis. *Agriculture, Ecosystems and Environment*, 266, 55-67. Link.
18. Vaziri, S., Wilson, J., Abbatematteo, J., Kubilis, P., **Chakraborty, S.**, Kshitij, K., and Hoh, D. J. (2017). Predictive performance of the American College of Surgeons universal risk calculator in neurosurgical patients. *Journal of Neurosurgery*, 1-6. Link.

STATISTICAL SOFTWARE DEVELOPMENT

1. `BAMBI`: An R package for Bivariate Angular Mixture Models. *Downloaded over 35,000 times.*
2. `variantprobs`: An R package for estimating probabilities and expected numbers of mutations in the tumor genome.
3. `hidgenclassifier`: An R package implementing Key functions for Bayesian hidden genome classifiers. Includes functions for preprocessing genommic data, fitting and predicting from hidden genome classifiers.
4. `pvLRT`: An R package for likelihood ratio test based methods for pharmacovigilance.

PRE-PRINT PUBLICATIONS

1. Mukherjee S, Khare K, and **Chakraborty S**. (2021). Convergence properties of data augmentation algorithms for high-dimensional robit regression. *Submitted.*
2. **Chakraborty, S.**, and Markatou, M. (2021). Likelihood Ratio Test Based Drug Safety Assessment using R package pvLRT.
3. **Chakraborty, S.**, Liu, A., Ball, R., and Markatou, M. (2021). On the Use of the Likelihood Ratio Test Methodology in Pharmacovigilance. *Submitted.*
4. **Chakraborty, S.** and Su, Z. (2021). A comprehensive Bayesian framework for envelope models. *In revision.*
5. **Chakraborty, S.** and Wong, S. W. (2021) On the circular correlation coefficients for bivariate von Mises distributions on a torus. *Under review.* arXiv.
6. Atanasova, K, **Chakraborty, S** (co-first author), Ratnayake, R, Khare, K, Luesch, H. (2021). An epigenetic small molecule screen to target abnormal nuclear morphology in human cells. *In revision.*
7. Batra, A, Barnard, A, Lott, D J, Willcocks, R, Forbes, S C, **Chakraborty, S**, Daniels, M, Arbogast, J, Triplett, W, Henricson, E, Dayan, J G, Schmalfuss, C, Sweeney, L, Byrne, B J, McDonald, C, Vandenborne, K, Walter, G A. (2021). Longitudinal changes in cardiac function in Duchenne muscular dystrophy population as measured by magnetic resonance imaging. *In revision.*

GRANTS

1. R01CA251339: Harnessing rare variants for tumor classification (PI: Shen, R.). **Source:** National Cancer Institute (HHS - NIH). **Role:** Consultant. **Period:** April 2021 - March 2024. Total Award Amount: $404,888.00.
2. 75F40120C00159: Evaluating LRT for Post-Market Surveillance of Adverse Events (PI: Markatou, M.). **Source:** US Food and Drug Administration (FDA). **Role:** Co-investigator **Period:** October 2020 – September 2022. Total Award Amount: $527,735.00.
3. 5UL1T1001412-06: University at Buffalo Clinical and Translational Science Institute (PD/PI: Murphy, T.F.). **Source:** National Institute of Health: National Center for Advancing Translational Sciences (NIH/NCATS). **Role:** Co-investigator. **Period:** December 2019 - December 2024. Total Award Amount: $19,231,451.00.

TEACHING

1. STA 521: Introduction to Theoretical Statistics I. Department of Biostatistics, University at Buffalo. Fall 2020.
2. STA 522: Introduction to Theoretical Statistics II. Department of Biostatistics, University at Buffalo. Spring 2021.
3. STA 4321 & STA 5325: Introduction to Probability & Fundamentals of Probability. Department of Statistics, University of Florida. Fall 2017.

PRESENTATIONS

Invited Presentations (Presenter's name in bold)

1. **Chakraborty, S.**, Guan, Z, Arora, A., Ecker, B., Martin, A., Jarnagin, W, Gonen M, Seier K, Begg C B, and Shen R. Mining mutation contexts across the cancer genome to map tumor site of origin with applications in Ampullary Adenocarcinoma. Invited Oral Presentation given at University at Buffalo Cancer Research Consortium Seminar Series, Buffalo, NY, July 2021.
2. **Chakraborty, S.**, Su, Z. A Comprehensive Bayesian Framework for Envelope Models. Invited oral presentation given (virtually) at CFE-CMStatistics conference, London, UK, December 2020.
3. **Chakraborty S**, Arora, A, Shen R, Begg C. B. Using Somatic Variant Richness to Mine Signals from Rare Variants & Using the "Hidden" genome. Oral presentation given at Epidemiology and Biostatistics departmental seminar series, Memorial Sloan Kettering Cancer, New York, NY, January 2020.
4. Chakraborty S, Arora, A, Begg C. B, **Shen, R.**. Using Somatic Variant Richness to Mine Signals from Rare Variants in the Cancer Genome. Invited oral presentation given at the Joint Statistical Meetings, Denver, CO, USA, August 2019.
5. Chakraborty, S., **Su, Z.** A Comprehensive Bayesian Framework for Envelope Models. Invited oral presentation given at EcoSta conference, Taichung, Taiwan, 2019.
6. **Chakraborty S**, Shen R., Begg C. B. Estimating Somatic Variant Richness in the Cancer Genome. Invited oral presentation given at the Epi- demiology & Biostatistics Departmental Seminar Series, Memorial Sloan-Kettering Cancer Center, New York, NY, 2019.
7. **Chakraborty, S.**, Khare, K. Convergence properties of Gibbs samplers for Bayesian probit regression with proper priors. Invited oral Presentation given at the Conference of Indian Statistical Association, Hydrabad, India, 2017.

Invited Topic-Contributed Presentations

8. **Chakraborty S**, Su, Z. A Comprehensive Bayesian Framework for Envelope Models. Invited topic contributed oral presentation given at Joint Statistical Meeting Virtual Conference, USA, August 2021.

Contributed Presentations (Presenter's name in bold)

9. **Chakraborty S**, Arora, A, Shen R, Begg C. B. Using Somatic Variant Richness to Mine Signals from Rare Variants in the Cancer Genome. Poster presentation given at the annual postdoc symposium at the Memorial Sloan-Kettering Cancer Center, New York, NY, 2019.
10. **Lee M**, Chakraborty S, Su Z. A Bayesian quantile envelope regression model. Poster Presentation given at the Joint Statistical Meetings, Denver, CO, 2019.
11. Awan O, Scott K, **Vaziri S**, Chakraborty S, Kshitij K, Rahman M. Re- imbursement Patterns for Neurosurgery: Analysis of the NERVES Survey Results from 2011-2016. Poster presented at the University of Florida Research Symposium, 2019.

12. **Vaziri S**, Scott K, Awan O, Chakraborty S, Kshitij K, Kubilis P, Hoh D. Risk Calculators in Neurosurgery: Identifying the High Cost Patient. Oral Presentation given at the University of Florida Neurosurgical Research Symposium in Gainesville, FL, 2019.
13. **Vaziri S**, Henson C, Scott K, Awan O, Chakraborty S, Kshitij K, Kubilis P, Hoh D. Predictors of Cost in Patients Undergoing Lumbar Spine Surgery. Oral Presentation given at the CNS Spine Section National Meeting in Miami, FL, 2019.
14. **Vaziri S**, Awan O, Scott K, Chakraborty S, Khare, K., Rahman M. Re- imbursement Patterns for Neurosurgery: Analysis of the NERVES Survey Results from 2011-2016. Oral Presentation given at the AANS in San Diego, CA 2019.
15. **Chakraborty, S.**, Khare, K. Consistent estimation of the spectrum of trace class data augmentation algorithms. Contributed oral presentation given at ENAR, Atlanta, GA, 2018.
16. **Chakraborty, S.**, Wong, S. W. BAMBI: An R package for bivariate angular mixture models. Contributed oral presentation given at ENAR, Washington, D.C., 2017.
17. **Chakraborty, S.**, Khare, K. Consistent estimation of the spectrum of trace class data augmentation algorithms. Contributed oral presentation given at the Joint Statistical Meetings, Baltimore, MD, 2017.

PROFESSIONAL EXPERIENCE

2017-18 **Graduate Research Assistant**, *Department of Statistics, University of Florida*, Gainesville, FL, USA.

2017 **Graduate Course Instructor**, *Department of Statistics, University of Florida*, Gainesville, FL, USA.

2014-15 **Graduate Teaching Assistant**, *Department of Statistics, University of Florida*, Gainesville, FL, USA.

2013-14, 2015-16 **Graduate School Fellow**, *Department of Statistics, University of Florida*, Gainesville, FL, USA.

PROFESSIONAL SERVICES

2021-23 **Biometric Bulletin Correspondent**, Eastern North American Region (ENAR) Section of International Biometric Society (IBS).

2020-Present **Peer Reviewer of Scholarly Journal Articles**, at Biometrics, Journal of the Royal Statistical Society, Journal of the American Statistical Association, Electronic Journal of Statistics, The R Journal, Frontiers in Oncology Electronic Journal of Statistics, and Statistics in Medicine.

2020 **Peer Reviewer of Academic Grant Proposals**, Reviewed Grant at National Science Foundation.