Saptarshi Chakraborty, Ph.D.

CONTACT INFORMATION

Department of Biostatistics
718 Kimball Tower
Buffalo, NY 14228, USA

saptarshichakraborty.net
E-mail: chakrab2@buffalo.edu
Google Scholar | ORCID

EMPLOYMENT

State University of New York at Buffalo August 2020 - Present
Assistant Professor (Tenure-Track)
Department of Biostatistics
Buffalo, NY, USA

RESEARCH INTERESTS

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

EDUCATION/ TRAINING

Memorial Sloan-Kettering Cancer Center, New York, NY, USA

Postdoctoral Research, Biostatistics, September 2018 - August 2020

- Project Description: *Statistical modeling of big and high-dimensional cancer genomic data*
- Mentors: Colin B. Begg, Ph.D., and Ronglai Shen, Ph.D.

University of Florida, Gainesville, FL, USA

Ph.D., Statistics, August 2013 - August 2018

- Thesis Topic: *Theory and Applications of Markov Chain Monte Carlo Techniques*
- Advisors: Kshitij Khare, Ph.D. and Samuel W.K. Wong, Ph.D.

Indian Statistical Institute, Kolkata, India

M.S., Statistics, July 2011 - June 2013

- Specialization: Applied Statistics and Data Analysis

Presidency College, Kolkata, India

B.Sc. (Hons.), July 2008 - June 2011

- Major: Statistics; Electives: Mathematics, Economics

PROFESSIONAL EXPERIENCE

Graduate Research Assistant August 2017 – August 2018
Department of Statistics,
University of Florida
Supervisor: Michael J. Daniels, Sc.D.

Graduate Course Instructor January 2017 – May 2017
STA 4321 & STA 5325 (Fall 2017):
Introduction to Probability & Fundamentals of Probability
Department of Statistics
University of Florida

Graduate Teaching Assistant September 2014 – May 2015
STA 2023 : Introduction to Statistics
Instructors: Megan Mocko (Fall 2014), Maria Ripol (Spring 2015)
Department of Statistics
University of Florida

Methodological Research Publication

Refereed Journal Articles

1. Lee, M., **Chakraborty, S.**, and Su, Z. (2021). A Bayesian approach to envelope quantile regression. *To appear in Statistica Sinica.*

2. **Chakraborty, S.**, Begg, CB, and Shen, R (2020). Using the "Hidden" Genome to Improve Classification of Cancer Types. *Biometrics.* 2020; 1–11. Link. arXiv.

3. **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. *Nature Communications 10*, 5506 (2019). Link.

4. **Chakraborty, S.** and Khare, K. (2019). Consistent estimation of the spectrum of trace class data augmentation algorithms. *Bernoulli.* 25(4B), 2019, 3832 -– 3863. Link.

5. **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.

6. 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.

Submitted Articles and Articles under Review

1. **Chakraborty, S.**, Martin, A., Begg, C. B., Shen, R. (2021+) Mining the Hidden Genome to Map Tumor Site of Origin. In revision at the *Nature Communications.*

2. **Chakraborty, S.** and Su, Z. (2021+). A comprehensive Bayesian framework for envelope models. In revision at the *Journal of the American Statistical Association.*

3. **Chakraborty, S.** and Wong, S. W. (2021+). Bayesian analysis of coupled cellular and nuclear trajectories for cell migration. In revision at *Biometrics*.

4. **Chakraborty, S.**, Bhattacharya, B., and Khare, K. (2021+). Estimating accuracy of MCMC variance estimator: a central limit theorem for batch means estimator. Under Review at *Journal of Computational and Graphical Statistics*.

5. **Chakraborty, S.** and Wong, S. W. (2021+). On the circular correlation coefficients for bivariate von Mises distributions on a torus. Under review.

Computational Research Publication

Refereed Journal Articles

1. **Chakraborty, S.** and Wong, S. W. (2019). `BAMBI:` An R package for Fitting Bivariate Angular Mixture Models. To appear in the *Journal of Statistical Software*. Arxiv.

Collaborative Scientific Research Publication

Refereed Journal Articles

1. 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*, 78(1), 50-59.

2. Vaziri, S., Awan, O., Porche, K., Scott, K., Sacks, P., Dru, A.B., **Chakra- borty, 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.

3. Rooney, W.D., Berlow, YA, 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.

4. Barnard, AM, Wilcox, R., Forbes, SC, Daniels, M. J., **Chakraborty, S.**, Lott, D., J., Senesac, C. R., Arora, H., Sweeny, L., Walter, G. H., and Vandenborne, K. H. E. (2019). MR biomarkers and predictive relationships to clinical function over 48 months in Duchenne muscular dystrophy. *Neurology*, *94*(9), e897-e909. Link.

5. 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.

6. 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.

Submitted Article

1. Batra, A., Barnard, A., Lott, D. J., Willcocks, R., Forbes, S. C., **Chakraborty, S.**, Daniels, M., Arbogast, J., Triplett, W., Henricson, E., Dayan, Y., Schmalfuss, C., Sweeney, L., Byrne, B. J., McDonald, C., Vandenborne, K., Walter, G. A. (2021+). The natural history of cardiac disease progression in the Duchenne muscular dystrophy population as measured by magnetic resonance imaging. *In revision.*

Statistical Software

1. BAMBI: An R package for Bivariate Angular Mixture Models.
2. variantprobs: An R package for estimating probabilities and expected numbers of mutations in the tumor genome.
3. hidgenclassifier: An R package implementing methodologies described in "Mining the Hidden Genome to Map Tumor Site of Origin"

Teaching Experience

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

Awards

- College of Liberal Arts and Sciences Graduate Travel Award, University of Florida, December 2017
- College of Liberal Arts and Sciences Graduate Travel Award, University of Florida, February 2017
- Graduate School Fellowship, University of Florida, August 2013 - August 2017.
- INSPIRE scholarship, Ministry of Science & Technology, Govt. of India, Jun 2008 - Jun 2011.

Oral and Poster Presentations (Presenter's name in bold)

- **Chakraborty, S.**, Su, Z. A Comprehensive Bayesian Framework for Envelope Models. Invited oral presentation given (virtually) at CFE-CMStatistics conference, London, UK, December 2020.
- **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 2020.
- **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.
- **Lee M**, Chakraborty S, Su Z. A Bayesian quantile envelope regression model. Poster Presentation given at the Joint Statistical Meetings, Denver, CO, 2019.
- Chakraborty S, **Shen R**, Begg C. B. Estimating Somatic Variant Richness in the Cancer Genome. Oral Presentation given at the Joint Statistical Meetings, Denver, CO, 2019.
- Chakraborty, S., **Su, Z.** A Comprehensive Bayesian Framework for Envelope Models. Invited oral presentation given at EcoSta conference, Taichung, Taiwan, 2019.
- 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 Re- search Symposium, 2019.
- **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.
- **Chakraborty S**, Shen R., Begg C. B. Estimating Somatic Variant Richness in the Cancer Genome. Oral presentation given at the Epidemiology & Biostatistics Departmental Seminar Series, Memorial Sloan-Kettering Cancer Center, New York, NY, 2019.
- **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.
- **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.
- **Chakraborty, S.**, Khare, K. Consistent estimation of the spectrum of trace class data augmentation algorithms. Contributed oral presentation given at ENAR, Atlanta, GA, 2018.
- **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.

- **Chakraborty, S.**, Wong, S. W. BAMBI: An R package for bivariate angular mixture models. Contributed oral presentation given at ENAR, Washington, D.C., 2017.
- **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.

Service

- Eastern North American Region (ENAR) Correspondent for Biometric Bulletin by International Biometric Society. January 2021 - January 2023.
- Organizer of Student Seminar Series, Department of Statistics, University of Florida, Fall 2017 - Spring 2018