DEBRAJ BOSE

Email: debrajbose96@gmail.com | Phone: 734-881-5074

Website: bosedebraj96.github.io

EDUCATION

University of Michigan, Ann Arbor, Michigan   2019 – 2026 (expected)
Ph.D. in Biostatistics, GPA: 4.0/4.0
Thesis: Statistical Methods for Gene-Level Inference and Causal Gene Prioritization in Human Genetic Studies
Co-advised by: Prof. Michael Boehnke and Prof. William Wen

Indian Statistical Institute, Delhi and Kolkata   2017 – 2019
M.Stat. (Biostatistics Specialization), Overall Score: 79.5%

St. Xavier's College, Kolkata
B.Sc. (Hons) in Statistics, GPA: 8.9/10   2014 – 2017

PROFESSIONAL EXPERIENCE

University of Michigan, Ann Arbor
Graduate Research Assistant, Dept. of Biostatistics   Sep 2019 – Present

- Performed analytic calculations and conducted large-scale simulations for rare-variant association analyses on **UK Biobank exome data (n $\sim$ 408K)** to compare power of burden, SKAT, and SKAT-O methods with single-variant tests under diverse genetic architectures and sample sizes.

- Integrated GWAS, molecular QTLs, and gene-level features using an **EM-based framework** to estimate gene-level posterior probabilities of causality.

- Currently building scalable analysis pipelines in **R** to apply our gene prioritization method in real data.

Indian Statistical Institute, Kolkata
Master's Project   Aug 2018 – May 2019
Advisor: Prof. Subir K. Bhandari

- Proposed optimal histogram construction in one and multiple dimensions by (1) using **k-means clustering and Voronoi tessellations**, and (2) minimization of mean integrated squared error.

- Developed a novel supervised classification method based on **optimal histogram construction** and demonstrated through simulations across multiple settings that these classifiers achieve performance comparable to optimal Bayes classifier.

- Strengthened expertise in statistical learning theory, density estimation, and algorithmic implementation.

Reliance Industries Limited, Navi Mumbai
Summer Intern, Analytics and Strategic Initiatives   May 2018 – July 2018

- Built NLP-based pipelines in **Python** to extract, summarize, and classify information from quarterly financial reports of major oil and gas competitors, and also automated financial table generation and report summarization, reducing manual analysis time and improving data consistency.

- Collaborated with business analysts to translate unstructured text into structured insights for strategic decision-making.

LANGUAGE, PROGRAMMING AND STATISTICAL SKILLS

Languages: English, Bengali, Hindi.
Programming languages and frameworks: R, Python, C++, Linux, Git.

Summary of statistical and computational skills:

- Strong background in analyzing and performing simulations with **large-scale genetic and metabolomic datasets** in high-performance **Linux** computing environments.

- Experience in **statistical modeling and inference in R**, including longitudinal, survival, and clinical trial data analysis.

- Developed visualizations and interactive dashboards using **R Shiny, ggplot2, plotly**, and **flexdashboard**.

POSITIONS OF RESPONSIBILITY

Statistics in the Community (STATCOM), University of Michigan
Communications Chair (2024–2025), Member (2023–2025)

- Led and contributed to pro bono data science projects for nonprofit and community organizations.

- Developed interactive **flexdashboard** visualizations for the Oakland Livingston Human Service Agency to understand county-level demographics, socioeconomic data, and needs of people experiencing hardship and poverty in Southeast Michigan.

- Built an **R Shiny dashboard** to visualize visitor demographics and geographic trends for the Ann Arbor Hands-On Museum.

- Led a team project to analyze and summarize door-to-door survey data in two impoverished neighborhoods for The People's Action, producing a research report to support community advocacy efforts.

- Led a team project to design summary analyses and visualizations, including text summaries and word clouds, for program evaluation surveys for the Center for Success literacy initiative.

Vice-President (2020-2022) and A-Team Member (2019-2025) of University of Michigan Table Tennis Team

AWARDS AND RECOGNITION

· Rackham Graduate School Conference Travel Grants	2022 – 25
· Sabyasachi Roy Gold Medal award for best M.Stat. project	2019
· Book Grant for exemplary performance in M.Stat. examinations	2018 – 19
· INSPIRE scholarship from Government of India	2014 – 2017

PUBLICATIONS

- **Bose D,** Fuchsberger C, Boehnke M. Rare-variant association studies: When are aggregation tests more powerful than single-variant tests? *American Journal of Human Genetics* 112(8), 1948-1961 (2025).

- Yin X, Li J, **Bose D,** ... , Morrison J. Assessing the potential causal effects of 1099 plasma metabolites on 2099 binary disease endpoints. *Nature Communications* 16, 3039 (2025).

- Yin X, **Bose D,** Kwon A, ... , Wen X. Integrating transcriptomics, metabolomics, and GWAS helps reveal molecular mechanisms for metabolite levels and disease risk. *American Journal of Human Genetics* 109(10), 1727-1741 (2022).

- Yin X, Chan LS, **Bose D,** ... , Boehnke M. Genome-wide association studies of metabolites in Finnish men identify disease-relevant loci. *Nature Communications* 13, 1644 (2022).

- Ray D, Salvatore M, Bhattacharyya R, ... , **Bose D,** ... , Mukherjee B. Predictions, role of interventions and effects of a historic national lockdown in India's response to the COVID-19 pandemic: data science call to arms. *Harvard Data Science Review*, Suppl. 1 (2020).