



THE UNIVERSITY OF GEORGIA  
DEPARTMENT OF STATISTICS

## *Colloquium Series*

**Thursday, August 31, 2023**

**4:00 PM, Room 204, Caldwell Building**

**Dr. Aditya Mishra**

Department of Statistics, University of Georgia

### **On integrative statistical learning approach for cancer genomics and microbial science**

The modeling of interrelated responses utilizing correlated predictors within high-dimensional contexts emerges as a compelling conundrum in the realm of integrative statistical learning, resonating across diverse scientific inquiries. This dilemma is manifest in scenarios such as genomics, where insights into the modulation of gene expression in yeast cells across multiple temporal points via transcription factors are sought; ocean microbiome investigations, delving into the influence of environmental and geochemical variables on microbial abundance; gut microbiome analyses, unraveling the interplay between host attributes and microbial prevalence; and cancer genomics, probing the interrelations among omics datasets, including microbiome and metabolomics information. Effectively addressing this challenge involves harnessing a multivariate analysis framework, synergized with a low-rank, sparse coefficient matrix, to infer the latent associations underpinning the phenomena. However, the simultaneous imposition of orthogonality and sparsity constraints renders the decipherment of such a decomposition arduous. To surmount this, we introduce a divide-and-conquer strategy for inferring the coefficient matrix from data, entailing decomposition into unit-rank matrices with sparse left and right singular vectors, estimated via a sequential (greedy) approach. Our method accommodates diverse outcome types, assuming conditional independence and adherence to an appropriate exponential dispersion family. Implementation is realized through R packages "secure" (Gaussian outcomes), "GO-FAR" (exponential outcomes), and "NB-FAR" (negative binomial outcomes), with demonstrated efficacy on multi-omics data encompassing colorectal cancer, yeast cell cycle, and Tara Ocean Expedition genomics datasetse.

#### **About the Speaker**

Dr. Aditya Mishra is an environmental AI cluster hire faculty in the Department of Statistics. Prior to joining the department, Dr. Mishra was a computational/data scientist at the Platform for Innovative Microbiome and Translational Research at the University of Texas MD Anderson Cancer Center. At MD Anderson, Aditya was involved in developing and applying statistical methods to understand the role of the human microbiome in cancer onset, progression, and response to therapy. Aditya's research

mainly focuses on developing statistical methods and computational tools using the framework of High-dimensional statistics, Multivariate analysis, Reduced-rank regression, Regularization, Robust statistics, Multi-omics, Computational biology, Causal inference and Variational inference. He has extensive interdisciplinary research experience in a variety of fields, including microbiome data analysis, genomics, cancer genomics, public health, ocean microbiology, and dietary intervention study. As a part of environmental, Aditya will continue to understand and investigate the role of microbiome various context including cancer genomics, marine science, soil science and agriculture science.



For more information, please contact us at:

Phone: 706.542.5232 E-Mail: [stat@uga.edu](mailto:stat@uga.edu)

Parking is available in the North Campus Parking Deck.

For a UGA Campus map, please see: <http://aviary.camplan.uga.edu/CampusMap/Default.aspx>