# Madhav Sankaranarayanan

## Curriculum Vitae

321 Franklin St., Cambridge - 02139, MA, USA
+1 (857) 867 2727 · Email · LinkedIn · GitHub

## Education

| Program | Institution | Year |
|---|---|---|
| **Doctor of Philosophy**<br>*Biostatistics*<br>Advisor: Rajarshi Mukherjee | **Harvard T.H. Chan School of Public Health**<br>*Boston, MA, USA* | September 2021-present |
| **Masters of Statistics**<br>*Specialization: Theoretical Statistics* | **Indian Statistical Institute**<br>*Kolkata, WB, India* | July 2019-May 2021 |
| **Bachelors of Statistics (Honours)**<br>*Major: Statistics* | **Indian Statistical Institute**<br>*Kolkata, WB, India* | July 2016-May 2019 |

## Research Interests

High Dimensional Inference • Random Matrix Theory • Causal Inference • Statistical Genetics

## Current Projects

- **Analysis of High Dimensional Instrumental Variable models** (in preparation)
  *Madhav Sankaranarayanan*, Rajarshi Mukherjee, Julien Chhor
  - Study the estimation of signal-to-noise ratio in practical high dimensional IV models, such as Mendelian randomization models
  - Prove optimality of exact instrument recovery under a range of dimensional specifications
  - Improve the estimation of causal quantities under particular practically relevant sparsity conditions
- **Asymptotic Inference in Genetic Association Studies using Genetic Correlations of Glycemic Traits** (Submitting to *Biometrika*)[IISA Student Poster Award]
  *Madhav Sankaranarayanan*, Rajarshi Mukherjee, Tamar Sofer, Yana Hrtsenko
  - Study the association of genetic determinants of proteins with glycemic traits
  - Improve the estimation of these associations using the estimated genetic correlation between proteins and traits
  - Construct provably optimal estimators and algorithms, and study polygenic risk scores for these traits
- **Minimax Detection of the Number of Spikes in Large Wigner Matrices** (Submitted to COLT 2025)
  *Madhav Sankaranarayanan*, Rajarshi Mukherjee, Soumendu Sundar Mukherjee
  - Test presence of spikes in a spiked Wigner matrices
  - Deal with the bounded and unbounded dimensionality of spike sparsity and strength
  - Investigate asymptotic properties of various tests
- **Mortality Analysis of Benzodiazepine Initiation in linked Registry-Claims data** (Submitted to *Pharmacoepidemiology and Drug Safety*)
  *Madhav Sankaranarayanan*, Lidia Moura, Mila Sun, Julianne Brooks, Victor Lomachinsky Torres, Sebastien Haneuse, Mamoon Habib
  - Investigate the effect of benzodiazepines on post-stroke mortality in the elderly on a nationwide scale
  - Extend methodology from studies on local hospital cohorts to Medicare dataset
  - Implement an emulated trial design to account for semi-competing risks and overlapping observation times
- **A Distribution-free Mixed-Integer Optimization approach to Hierarchical Modelling of Clustered and Longitudinal data** [arXiv][NESS Student Research Award]
  *Madhav Sankaranarayanan*, Intekhab Hossain, Tom Chen
  - Implement a mixed-integer optimization (MIO) approach for doing cluster-aware regression
  - Compare to linear mixed effects regression (LMEM) in terms of causal recovery and prediction
  - Establish framework for generalization to new data points using classification trees

## Other Projects

- **Middle Meningeal Artery Embolization in adjunction to Surgical Evacuation for treatment of Subdural Hematomas: a Nationwide comparison of outcomes with Isolated Surgical Evacuation** [Neurosurgery]
  Mirhojjat Khorasanizadeh, Seyed Farzad Maroufi, Rajarshi Mukherjee, *Madhav Sankaranarayanan*, Justin M. Moore, Christopher S. Ogilvy
  - Investigate risk of surgical evacuation for chronic subdural hematomas
  - Investigate middle meningeal artery embolization (MMAE) as a novel treatment approach
  - Perform meta-analysis on multiple small sample studies from hospitals across the country

- **Perfect Transformation Based Markov Chain Monte Carlo**
  M.Stat Dissertation Project, advised by Sourabh Bhattacharya
  - Construct computationally efficienct Transformational MCMC procedures to construct valid and effective algorithms
  - Constrast efficacy of these TMCMC procedures with other MCMC procedures
  - Leverage the dimension-reduction property of TMCMC to create a "perfect" sampling methodology for high-dimensional target distributions

- **Quantitative Analysis of Polygenic Risk Scores in the Genes for Good Cohort**
  Summer Project as part of Genomics group in BDSI 2019, advised by Matthew Zawistowski and Brooke Wolford
  - Calculate polygenic risk scores for individuals in the GfG dataset, crowdsourced by the School of Public Health
  - Test for traits such as hypertension, rheumatoid arthritis, schizophrenia and left-handedness
  - Elucidate shortcomings of polygenic risk scores, and investigate potential improvements

- **An MCMC-free approach to Post-selective Inference**
  Project advised by Snigdha Panigrahi
  - Provide an approximation algorithm for selective inference, without using an MCMC sampling method
  - Construct confidence intervals which match the inferential power of previous methodologies
  - Extend to general models in randomized settings

## Presentations

### Conferences

- **Asymptotic Inference in Genetic Association Studies using Genetic Correlations of Glycemic Traits**
  Poster Presentation at International Indian Statistical Association Conference 2024
- **Optimal Detection of the Number of Spikes with Application to Genetic Association Testing**
  Presentation at Joint Statistical Meetings 2024
- **A Distribution-free Mixed-Integer Optimization approach to Hierarchical Modelling of Clustered and Longitudinal data**
  Presentation at New England Statistics Symposium 2024
- **Asymptotic Inference in Genetic Association Studies using Genetic Correlations of Glycemic Traits**
  Presentation at Joint Statistical Meetings 2023
- **Quantitative Analysis of Polygenic Risk Scores in the Genes for Good Cohort**
  Poster at Symposium on Big Data, Human Health and Statistics 2019

### Seminars

- Keynote Speaker at StatStart 2024, July 2024
- **Mortality Analysis of Benzodiazepine Initiation**
  Harvard Biostatistics Student Seminar, Feb 2024
- **A Distribution-free Mixed-Integer Optimization approach to Hierarchical Modelling of Clustered and Longitudinal Data**
  Harvard Biostatistics Student Seminar, Apr 2023

## Work Experience

**Center for Value-Based Health Care and Sciences, Mass General Hospital** — **Boston, MA, USA**
*Research Assistant* — *July 2023 - Present*
Co-advised by Sebastien Haneuse and Lidia Moura

**University of Michigan School of Public Health** — **Ann Arbor, MI, USA**
*Summer Research Student* — *June 2019 - July 2019*
Part of the Big Data Science Initiative

**Institute of Mathematical Sciences** — **Chennai, TN, India**
*Visiting Researcher* — *May 2018 - July 2018*
Advised by Gautam Menon

## Teaching and Mentoring Experience

- **Biostatistics Summer Preparatory Course** (August 2024)
  - Instructed classes in operational math (real analysis, linear algebra)
- **StatStart** (July 2023)
  - A one month summer intensive program intended for high school students from underrepresented backgrounds interested in data science and computing
  - Organized by the Department of Biostatistics, Harvard T.H. Chan School of Public Health
  - Instructed the Intro to Statistics and Probability classes
- **Summer Program in Biostatistics and Computational Biology** (June 2023 -July 2023)
  - A 6 week summer program, offering diverse undergraduate students a unique opportunity to learn about the use of quantitative methods for biological, environmental, and medical research alongside Harvard faculty, researchers, and graduate students.
  - Organized by the Department of Biostatistics, Harvard T.H. Chan School of Public Health
  - Mentored students as the Student Research Mentor for the research group advised by Rafael Irizarry
- **Qualifying Examination Preparation** (July 2023)
  - Taught classes on probability (BIOSTAT 230) to Ph.D. students taking their qualifying exam
- **Teaching Assistant** (January 2023 - May 2023)
  - Statistical Inference 2 (BIOSTAT 241)
  - Instructed by Rajarshi Mukherjee
  - Received Certificate of Distinction in Teaching

## Technical Skills

- Programming Language: R, Python, Julia
- Operating Systems: MacOS, Windows, Linux
- Tools: LaTeX, Microsoft Excel, Microsoft Office, Photoshop

## Achievements

- Received a Student Poster Award at International Indian Statistical Association Conference (2024)
- Received a Student Research Award at the New England Statistical Symposium (2024)
- Received Certificate of Distinction in Teaching from the Department of Biostatistics at Harvard T.H. Chan School of Public Health (2023)
- Recipient of Robert Balentine Reed Prize for Excellence in Biostatistical Science (2022)
- Secured Distinction in the B.Stat (Hons.) program at Indian Statistical Institute, Kolkata (2019)
- Awarded the Kishore Vaigyanik Protsahan Yojana scholarship (2017)

## Others

- Hobbies: Puzzles, Quizzing, Badminton, Table Tennis, Chess, Origami
- Languages: English, Tamil, Hindi, Bengali (working knowledge), Konkani (working knowledge)