Aditya K. Mishra

Contact Information	201 Saint Pauls Ave, Apt 6L Jersey City, NJ 07306	+1 201-234-1456 amishra@flatironinstitute.org https://bit.ly/39zKPHR https://github.com/amishra-stats

Research Interests Variable selection, multivariate analysis, reduced-rank regression, regularization, singular value decomposition, microbiome data analysis, variational inference.

Education

Flatiron Institute, Simons Foundation, New York, NY

Flatiron Research Fellow, Sept 2017 - Present

- Objective: *Developing statistical model for analyzing the microbiome data*
- Mentor: Dr. Christian L. Muller

University of Connecticut, Storrs, CT

M.S. and Ph.D., Statistics, Aug 2013 - Aug 2017

- Thesis Topic: *Sequential estimation of multivariate association*
- Advisors: Professor Dipak Dey and Professor Kun Chen

Indian Institute of Technology, Kharagpur, Kharagpur, India

B.S and M.S., Statistics and Informatics, Jun 2005 - Aug 2010

- Topic: *A study of identification of intron-exon boundary in eukaryotic genes.*
- Advisor: Professor Umesh C Gupta
- Best thesis award

Publications

Mishra, A., Dey, D., Chen, K. (2017) *Sequential co-sparse factor regression. Journal of Computational and Graphical Statistics, 26(4), 814-825.*

Mishra, A., Dey, D., Chen, Y., Chen, K. (2020) *Generalized co-sparse factor regression. Computational Statistics & Data Analysis, 157, p.107127.*

Mishra, A., Müller, C. (2019) *Robust regression with compositional covariates. arXiv preprint arXiv:1909.04990.* **Under review**.

Manuscript in Preparation

Mishra, A., Buja, A., Müller, C. (2021) Negative binomials factor regression model for microbiome data analysis.

Mishra, A., Müller, C., McNichol, J., Blei, D. (2021) Embedding model for learning microbial associations.

Badri, M., Morton, J., **Mishra, A.**, Eckles, D., Bonneau, R. (2021) Instrumental variable regression for the microbiome data analysis. (methodology contribution)

Mutshinda, C., **Mishra, A.**, Finkel, Z., Widdicombe, C., Irwin, A. (2021) Using a two-component bayesian model to decouple the abiotic controls of phytoplankton occurrence and abundance from in situ data.

Paper in Preparation

Mishra, A., Müller, C. (2021) *Robust microbial associations network model.*

Mishra, A., Hyun, S., Bien, J., Müller, C. (2021) *Statistical workflows for marine data analysis*

R Package

Mishra, A., Chen, K. (2017) secure: Sequential co-sparse factor regression. R package https://CRAN.R-project.org/package=secure.

Mishra, A., Chen, K. (2020) gofar: Generalized co-sparse factor regression. R package https://github.com/amishra-stats/gofar.

Mishra, A., Müller, C. (2019) robregcc: Robust regression with compositional covariates. R package https://cran.r-project.org/package=robregcc.

Mishra, A., Hyun, S., Bien, J., Müller, C. (2019) cmap4r: CMAP for R Users. R package https://github.com/simonscmap/cmap4r. *Popular among marine biologist.*

Presentations

Invited/Seminar Talk

1. "Robust regression with compositional covariates". ISI World Statistics Congress, Kuala Lampur, Malaysia Aug 2019
2. "Application of generalized co-sparse factor regression in microbiome data analysis". WNAR/IMS Annual Meeting, Portland, Oregon June 2019
3. "Learning association: with application to marine microbiome data". ITMAT, Perelman School of Medicine, Univ. of Pennsylvania, PH June 2019
4. "Statistical methods for the analysis of ocean microbiome data". CBIOMES Annual Meeting, NYC, NY May 2019
5. "Robust regression with compositional covariates". New England Statistics Symposium, Hartford, CT May 2019
6. "Sparse reduced rank model for multivariate count data". SAMSI Operator Splitting Workshop, NYC, NY March 2019
7. "Robust regression with compositional covariates". Dept. of Economic Sciences Colloquium at IIT Kanpur, Kanpur, India Jan 2019
8. "A divide and conquer strategy for low-rank and sparse factor regression". Dept. of Mathematics & Statistics Colloquium at IIT Kanpur, Kanpur, India Jan 2019
9. "Robust regression with compositional covariates". National Centre for Biological Sciences Colloquium, Bangalore, India Jan 2019
10. "Robust regression with compositional covariates". Fall Colloquium Seminars at Columbia Biostatistics, NYC, NY Sept 2018
11. ASA Conference on Statistical Learning and Data Science, NYC, NY June 2018
12. "Estimation of Inverse Covariance Matrix in Compositional Data". SAMSI Operator Splitting Workshop, Raleigh, NC March 2018
13. "Sequential co-sparse factor regression". Computational Bayesian Statistics Journal Club Aug 2020
14. "Negative binomial factor regression with application to microbiome data analysis". Biomedical Statistics and Data Science, LMU/HMGU, Munich, Germany Jan 2021

Contributed Talk

1. "Marine data mining with CMAP". Virtual Joint Statistical Meeting, ASA Aug 2020
2. "Marine data mining with CMAP using R". CMStatistics Annual Meeting, London, UK Dec 2019
3. "CMAP for R users". CBIOMES Virtual Meeting Nov 2019
4. ASA Joint Statistical Meetings, Denver, CO July 2019

5. Biometric Society (ENAR) Spring Meeting, Philadelphia, PH March 2019
6. CMStatistics Annual Meeting, Pisa, Italy Dec 2018
7. "Sequential Co-Sparse Factor Regression". ASA Joint Statistical Meetings, Baltimore, MD July 2017
8. Biometric Society (ENAR) Spring Meeting, Washington, D.C. March 2017
9. International Indian Statistical Association Conference, Pune, India Dec 2015

Poster

1. "Marine data mining with CMAP". CBIOMES annual meeting, New York, NY April 2020
2. "Statistical methods for ocean microbiome data". CBIOMES annual meeting, New York, NY April 2019
3. "Statistical methods for marine ecosystem modeling". CBIOMES annual meeting, New York, NY April 2018
4. " Sequential estimation of sparse factor regression". New England Statistics Symposium, New Haven, CT April 2016

Workshop

1. Strategies and Techniques for Analyzing Microbial Population Structures, Marine Biological Laboratory, Woods Hole, MA Aug 2018

Professional Experience

University of Connecticut, Department of Statistics, Storrs, CT

Consultant, Statistical Consulting Service **Sept 2015 to May 2017**

- Responsible for providing end-to-end analysis to both internal and external clients. Have completed work on **4** projects.
- Service requires designing statistical analysis, providing computational support using statisttical tools like R, SAS and SPSS, and presenting result of analysis as per client's objective.

Fractal Analytics, Mumbai, India

Data Scientist, Fractal Science **Nov 2012 to July 2013**

- Developed machine learning algorithm to learn behavioral attribute of product and consumer from transaction data using the concept of finite mixture model. The algorithm was used to develop product named **Customer Genomics**.
- Developed statistical model to predict customer lifetime value in non-contractual setting.
- Developed attribution model in digital marketing to account for incremental impact of different marketing channels leading to sales using the concept of survival analysis.
- Designed **dCrypt** , a self-learning, natural language processing (NLP) algorithm to predict product attribute from its description in CPG domain. The estimated savings of dCrypt is over 15000 analyst hours per annum.

Kavaii, Bangalore, India

Data Scientist **May 2012 to Nov 2012**

- Responsible for developing statistical model to provide predictive analytic solution to client in manufacturing and healthcare sector.

KACST, Riyadh, Saudi Arabia

Research Associate **Dec 2010 to March 2012**

- Developed mathematical model for vocal tract in cases of unvoiced fricatives sound signal and voiced consonant using statistical signal processing.
- Statistical modelling of biological pathway as Bayesian network from high throughput biological data (genetic expression data) using structural learning algorithm.

Awards

Student Awards — University of Connecticut, Department of Statistics

- Pre-Doctoral Fellowship — Summer 2015
- Elizabeth M. McFarlane Fellowship — Summer 2016

Student Awards — University of Connecticut, Graduate School

- Doctoral Dissertation Fellowship — Fall 2016
- Doctoral Student Travel Award for ENAR Spring Meeting 2017 — Spring 2017

Professional Associations

Member, New England Statistical Society (NESS) 2017–
Member, American Statistical Association (ASA) 2017–
Member, Institute of Mathematical Statistics (IMS) 2019–
Member, International Statistical Institute (ISI) 2019–

Teaching Experience

Department of Statistics, University of Connecticut

Instructor

STAT-3375Q-Introduction to Mathematical Statistics I — Fall 2016
STAT-3445-Introduction to Mathematical Statistics II — Spring 2017
STAT-1100Q- Elementary Concepts of Statistics — Summer 2016

Teaching Assistant — Fall 2013–Spring 2016

STAT 1000 - Introduction to Statistics 1
STAT 1100 - Elementary Concepts of Statistics
STAT 3375Q - Introduction to Mathematical Statistics I

Workshop

Statsistical Consulting Service workshop on repeated measure data analysis — May 2016

Service

Referee

- Biometrika
- Journal of Computational and Graphical Statistics
- Sankhya Series B
- Metrika

Flatiron Institute

- Member, Lodestar Curriculum Committee

Research References

Dipak K. Dey
Distinguished Professor
Department of Statistics
University of Connecticut, Storrs, CT
Phone: 860-486-3414
E-mail: dipak.dey@uconn.edu

Kun Chen
Associate Professor
Department of Statistics
University of Connecticut, Storrs, CT
Phone: 860-486-4847
E-mail: kun.chen@uconn.edu

Christian L. Müller
Project Leader in Computational Statistics
Center for Computational Mathematics
Flatiron Institute, Simons Foundation, NY
Phone: 646-603-3716
E-mail: cmueller@flatironinstitute.org

Teaching References

Robert Apruzese
Lecturer
Department of Statistics
University of Connecticut, Storrs, CT
Phone: 860-486-0068
E-mail: robert.apruzese@uconn.edu