

Keren Li

ASSISTANT PROFESSOR OF STATISTICS

Department of Mathematics
University of Alabama at Birmingham
Birmingham, AL 35294, USA
Mobile: 872-588-2017

Email1: kli@uab.edu
Email2: li.keren.cn@gmail.com

ACADEMIC/PERSONAL WEBSITE: sites.uab.edu/klilab
GOOGLE SCHOLAR: scholar.google.com
GITHUB REPOSITORIES: github.com/kerenli

Research Interests

Distributed machine learning; Federated learning; Deep learning; Generalized linear models; Variable selection; Optimal design; Graphic models; Bioinformatics; Financial math.

Working Experience

- 2022– Assistant Professor of Statistics, Department of Mathematics, University of Alabama at Birmingham (UAB), Birmingham, AL
- 2023– Associate Scientist, Informatics Institute, University of Alabama at Birmingham, Birmingham, AL
- 2018-2022 Postdoctoral Fellow, NSF-Simons Center for Quantitative Biology & Department of Statistics and Data Science, Northwestern University, Evanston, IL
- 2015-2018 Teaching Assistant and Researching Assistant, University of Illinois at Chicago (UIC), Chicago, IL
- 2014-2015 Visiting Scholar, University of Illinois at Chicago, Chicago, IL
- 2010-2014 Instructor, Chongqing University of Science and Technology (CQUST), Chongqing, China
- 2007-2010 Instructor, Beihai College of Beihang University (BHBH), Guangxi, China

Education

- 2018 PHD in Statistics, University of Illinois at Chicago, Chicago, IL
- 2004 MS in Mathematics, Louisiana State University, Baton Rouge, LA
- 2001 BA in Mathematics, Nankai University, Tianjin, China

Mentors

Postdoc: Ji-Ping Wang, Professor of Statistics and Chair, Adjunct Professor of Molecular BioSciences, Faculty member of NSF-Simons Center for Quantitative, Northwestern University

PhD : Jie Yang, Professor of Statistics and Director of Graduate Studies, University of Illinois at Chicago

Publications

- 2023 **Li, K.**, “Representative Learning: Anchored Score-Matching Representative”, (preprint).
- 2023 Zheng, D., **Li, K.**, Yang, J. , Response-Aided Score-Matching Approaches for Big Data Analysis and Model Selection under Generalized Linear Models, submitted to *Statistica Sinica*.
- 2023 Huang, Y., **Li, K.**, Mandal, A., Yang, J., A New Algorithm for D-optimal Designs under General Parametric Statistical Models with Mixed Factors, *arxiv:2309.09367*.
- 2022 **Li, K.**, Carroll, M., Vafabakhsh, R., Wang, X., Wang, J., “DNAcycP: a Novel Tool for DNA Cyclizability Prediction”, *Nucleic Acids Research*, March 2022; gkac162. [DOI:10.1093/nar/gkac162](https://doi.org/10.1093/nar/gkac162). Online app [DNAcycP](#). Python script available on GitHub: [kerenli/dnacyp](https://github.com/kerenli/dnacyp).
- 2022 **Li, K.**, Yang, J. , “Score Matching Representative Approach for Big Data Analysis with Generalized Linear Model”, *Electronic Journal of Statistics*, 2022; 16(1):592-635. [DOI:10.1214/21-EJS1965](https://doi.org/10.1214/21-EJS1965).
- 2020 **Li, K.**, Hope, M., Wang, X., Wang, J., “Ribo-DiPA: A Novel tool for differential pattern analysis in Ribo-seq data”, *Nucleic Acids Research*, December 2020; 48(21):12016–12029. [DOI:10.1093/nar/gkaa1049](https://doi.org/10.1093/nar/gkaa1049). R package [RiboDiPA](#) on *Bioconductor*.
- 2021 Jiang, L., Zhou, S., **Li, K.**, Wang, F., and Yang, J., “A New Nonparametric Estimation of Risk-Neutral Density and its Application in Variance Swaps”, *Frontiers in Applied Mathematics and Statistics*, 2021 January; 6:68. [DOI:10.3389/fams.2020.611878](https://doi.org/10.3389/fams.2020.611878)
- 2018 Zabawa, L., **Li, K.**, Chmell S., “Patient Dissatisfaction Following Total Knee Arthroplasty: External Validation of a New Prediction Model”, *European Journal of Orthopaedic Surgery & Traumatology*, 2019 May; 29(4):861-867. [DOI:10.1007/s00590-019-02375-w](https://doi.org/10.1007/s00590-019-02375-w)
- 2017 **Li, K.**, Yang, J., “D-optimal Sampling method for Big Data with Multinomial Logistic Models”, (preprint).

Talks

- 2023 “Unveiling Collective Intelligence: Navigating Representative Learning for Federated Insights”, Informatics Institute Powertalk Seminar Series, University of Alabama at Birmingham
- 2023 “Big Data, Distributed Learning, and Representative”, Mississippi State University
- 2023 “DNAcycP: A Deep Learning Attempt at Mechanical Properties of DNA”, The 2023 Western North American Region of The International Biometric Society / Institute of Mathematical Statistics (WNAR/IMS) Annual Meeting

- 2022 “Representative Approaches in Distributed Learning and Federated Learning”, 34th Annual University of Alabama System Applied Mathematics Meeting
- 2022 “Representative Approaches for Generalized Linear Models in Distributed Learning”, Department of Computer Science, University of Alabama at Birmingham
- 2022 “Response-Aided Score-Matching Approaches for Big Data Analysis”, 2022 International Conference on Statistical Distributions and Applications
- 2022 “Response-Aided Score-Matching Approaches for Big Data Analysis under Generalized Linear Models”, SIAM Conference on Mathematics of Data Science (MDS22)
- 2021 “Score-Matching Representative Approach for Big Data Analysis and Its Extension”, Sixth International Conference on Establishment Statistics
- 2020 “RiboDiPA: Differential pattern analysis in Ribo-seq data”, 2020 Conference on Quantitative Approaches in Biology
- 2020 “Differential pattern analysis in Ribo-seq data”, Northwestern University
- 2020 “Score-Matching Representative Approach for Big Data Analysis and its Extension”, SIAM Conference on Mathematics of Data Science (MDS20)
- 2019 “A new statistical method to investigate translational regulation using Ribo-profiling data”, 2019 Conference on Quantitative Approaches in Biology
- 2019 “Score Matching Representative Approach for Big Data Analysis with Generalized Linear Model”, Northwestern University
- 2019 “A new statistical method to investigate translational regulation using Ribo-profiling data”, 2019 Joint Statistical Meetings
- 2019 “Score Matching Representative”, 2019 International Conference on Statistical Distributions and Applications
- 2019 “A new statistical method to investigate translational regulation using Ribo-profiling data”, 2019 ICSA Midwest & NIC-ASA Joint Fall Meeting
- 2018 “Pre-Knowledge Based Lasso for Gaussian Graphical Models”, 2018 Conference on Quantitative Approaches in Biology
- 2014 “Simple Parallel Statistical Computing in R”, University of Illinois at Chicago

Grant Proposals

SUBMITTED PROPOSALS

Title: Representative Learning: A New Distributed Learning Architecture

Role: Principal Investigator

Funding Agency: Simons Foundation

Date of Submission: January 2024

Requested Amount: \$42,000

Title: A New Distributed Learning Structure: Representative Learning

Role: Principal Investigator

Funding Agency: University of Alabama at Birmingham

Date of Submission: December 2023

Requested Amount: \$9,614

Title: Machine Learning Solutions for Modeling Multi-Omic Molecular Mechanisms Underlying Phenotypes

Role: Principal Investigator

Funding Agency: University of Alabama at Birmingham
Date of Submission: October 2023
Requested Amount: \$30,000
Collaborators: Dr. Greer Dolby, Dr. Baocheng Geng, Dr. Roman G. Shterenberg

PROPOSALS UNDER PREPARATION

Title: Representative based Distributed Learning and Decision Making with Human in the Loop

Role: Principal Investigator

Funding Agency: Targeted for submission to NSF

Status: Under Preparation

Estimated Submission Date: January 2024

Collaborators: Dr. Baocheng Geng

Title: Distributed Learning Redefined: Integrating Representative Learning into Data Science

Role: Principal Investigator

Funding Agency: Targeted for submission to NSF

Status: Under Preparation

Estimated Submission Date: February 2024

Teaching (Instructor of Record)

2022-2023 UAB: Statistical Techniques for Machine Learning and Big Data, Introduction to Statistics
2017 UIC: Introduction to Probability
2010-2015 CQUST: Financial Mathematics I, Financial Mathematics II, Mathematical Modeling, Introduction to Computational Statistics, Mathematical Modeling, Applied Differential Equations, Linear Algebra, Probability, Advanced Mathematics I, Advanced Mathematics II
2008-2010 BHBH: Calculus I, Calculus II, Linear Algebra, Probability, Operations Research, Business Statistics

Developed/Developing New Courses

2023 MA 789 Statistical Machine Learning
2022-2023 MA 489/589 Statistical Techniques for Machine Learning and Big Data
2022-2023 MA 189 Data Dive into Birmingham

Courses Taught (Teaching Assistant)

2022 Northwestern: Statistical Methods for Bioinformatics and Computational Biology
2015-2017 UIC: Precalculus Mathematics, Calculus I, Calculus II, Introduction to Probability, Applied Statistical Methods II

Professional Service

- 2023 Invited Session Organizer, 2024 WNAR/IMS/Graybill Annual Meeting
- 2023 Program committee member, 22nd International Workshop on Data Mining in Bioinformatics (BIOKDD 2023)
- 2022 Session Chair, SIAM Conference on Mathematics of Data Science (MDS22)
- 2022 Program committee member, 21st International Workshop on Data Mining in Bioinformatics (BIOKDD 2022)

Journal Reviews

- 2023 Guest Editor: special issue “Mathematical Frontiers in Distributed Learning and High- Dimensional Data Analysis” for the journal Mathematics
- 2015– Reviewer: BMJ Open , Frontiers in Genetics, Computational Statistics and Data Analysis, Journal of Statistical Theory and Practice, Statistical Science, Contemporary Biostatistics with Biopharmaceutical Applications (2019 edition)

Departmental/University Service

- 2023 Search Committee Member, tenure track position of Probability, Statistics, or Actuarial Science
- 2023 Working Group Member, BS in Data Science
- 2023 Serve for Faculty Affairs Committee
- 2023 Faculty mentor, the 2023 UAB NSF Summer REU program
- 2022 Search Committee Member, tenure track position of Computational Math
- 2022 Working Group Member, BS in Data Science
- 2017 Co-chair of statistics graduate student committee at UIC, organized student seminars
- 2017 Student assistant coordinator for a new master program in UIC
- 2010-2014 Advisor in CQUST, to organize and train undergraduate students to participate the Mathematical Contest in Modeling (MCM) and the Interdisciplinary Contest in Modeling (ICM).

Mentees

- 2023– Xinlu Li, PhD student, UAB
- 2022– Joseph Casey, PhD candidate, UAB
- 2022– Qianjiao Chen, PhD candidate, UAB

Awards and Honours

- 2017-2018 Graduate Student Research Award, UIC
- 2017-2018 Graduate Student Service Award, UIC
- 2015-2016 Graduate Student Service Award, UIC

Computer skills

EXPERT: R, PYTHON, BASH script, Parallel Computing, L^AT_EX
BASIC: C++, SQL, MATLAB