MENGYU LI

Address: No. 59, Zhongguancun Street, Haidian District, Beijing 100872, P.R. China

Email: limengyu516@ruc.edu.cn   Tel.: (+86) 18811793569

GitHub / Google Scholar

EDUCATION

Institute of Statistics and Big Data, Renmin University of China, Beijing, China

- Ph.D. Candidate in Statistics   Sep. 2020 – present
- Supervisor: Cheng Meng, Assistant Professor

School of Statistics, Beijing Normal University, Beijing, China

- Bachelor in Statistics   Sep. 2016 – Jun. 2020
- Thesis advisor: Junlong Zhao, Professor

Faculty of Science, University of British Columbia, Vancouver, Canada

- Summer Program   Jun. 2018 – Jul. 2018

RESEARCH INTERESTS

Optimal transport; Subsampling in big data; AI in health care

AWARDS & HONORS

- **Spark Award (Encouragement Award)**   2023
  Huawei Technologies Co., Ltd. Collaborated with Jun Zhu and Cheng Meng

- **Student Paper Award** in Joint Statistical Meetings   2023
  Statistics in Imaging Section, American Statistical Association (**five winners per year**)

- **Student Paper Award** in Joint Statistical Meetings   2022
  Sections on Statistical Computing and Graphics, American Statistical Association (**five winners per year**)

Presented by RUC

- Joint Ph.D. Program (50,000 RMB)   2023 – 2024
- Outstanding Innovative Talents Cultivation Funded Program (70,000 RMB)   2021 – 2024
- First-Class Academic Scholarship   2021 – 2023
- Social Work and Volunteering Scholarship   2023
- Study Excellence Scholarship   2022
- Merit Student   2021

Presented by BNU

- First-Class (or Second-Class) Jingshi Scholarship   2017 – 2019
- Academic Competition Scholarship   2019
- Social Contribution Scholarship   2017

PUBLICATIONS & WORKING PAPERS

Published

1. **Mengyu Li**, Jun Yu, Tao Li & Cheng Meng (2023). "Importance sparsification for Sinkhorn algorithm." *Journal of Machine Learning Research*, 24(247), 1–44.
   Student Paper Award in 2023 JSM, Statistics in Imaging Section.

2. **Mengyu Li**, Jun Yu, Hongteng Xu & Cheng Meng (2023). "Efficient approximation of Gromov-Wasserstein distance using importance sparsification." *Journal of Computational and Graphical Statistics*, in press.

3. **Mengyu Li** & Junlong Zhao (2022). "Communication-efficient distributed linear discriminant analysis for binary classification." *Statistica Sinica*, 32, 1343–1361.

Under Review & Submitted

1. **Mengyu Li**, Jun Yu, Tao Li & Cheng Meng. "Core-elements for large-scale least squares estimation." Under review in *IEEE Transactions on Neural Networks and Learning Systems*.
   Student Paper Award in 2022 JSM, Sections on Statistical Computing and Graphics.

2. **Mengyu Li**, Jingyi Zhang & Cheng Meng. "Nonparametric additive models for billion observations." Under review in *Journal of Computational and Graphical Statistics*.

3. Yukuan Hu, **Mengyu Li**, Xin Liu & Cheng Meng. "Sampling-based methods for multi-block optimization problems over transport polytopes." Under review in *Mathematics of Computation*.

4. Xinlai Kang, **Mengyu Li**, Xuqiang Chen, Fangyu Li & Cheng Meng. "A Hausdorff regression paradigm for interval privacy." Under review in *IEEE Signal Processing Letters*.

In Preparation

1. Chengshuo Du, **Mengyu Li**, Tao Li, Jun Yu & Cheng Meng. "Ensemble pruning with conditional Wasserstein distance measure."

2. Jun Zhu, **Mengyu Li** & Cheng Meng. "Optimal transport based signal data augmentation and data fusion."

3. Xintao Yu, **Mengyu Li**, Shanwen Wang & Cheng Meng. "Optimal transport with p-adic numbers."

4. Gaoming Chen, Hongyu Wang, Haoxian Liang, **Mengyu Li**, Shanwen Wang & Cheng Meng. "Hirachical Gromov-Wasserstein distance with application in 3D point cloud data matching."

INVITED TALKS

- **Capital Statistics and Data Science High-Level Forum**, Beijing, China   Jul. 2023
  Efficient algorithms for large-scale optimal transport problems

TEACHING ACTIVITIES

Teaching Assistant

- Data Mining   Spring 2021
  My duties included lecturing exercise lessons, homework preparation and grading, etc.

PROFESSIONAL ACTIVITIES

Review for journals & conferences (more than 15 articles)

- Statistica Sinica

- Journal of Computational and Graphical Statistics

- Journal of Machine Learning Research

- Journal of Applied Statistics

- Neural Information Processing Systems (NeurIPS)

- International Conference on Machine Learning (ICML)

- International Conference on Learning Representations (ICLR)

PROGRAMMING & DEVELOPMENT

Programming

- R, Python, MATLAB

Package development

- **Spar-Sink** (`https://github.com/Mengyu8042/Spar-Sink`)
  `Python3` implementation of the paper "Importance sparsification for Sinkhorn algorithm".

- **Spar-GW** (`https://github.com/Mengyu8042/Spar-GW`)
  `Python3` implementation of the paper "Efficient approximation of Gromov-Wasserstein distance using importance sparsification".

- **Core-elements** (`https://github.com/Mengyu8042/Core-elements`)
  `R` implementation of the paper "Core-elements for large-scale least squares estimation".