Lucy L. Gao

Department of Statistics
University of British Columbia

lucy.gao@stat.ubc.ca   http://lucylgao.com

Education

PhD in Biostatistics, 2020
University of Washington
Advisor: Dr. Daniela Witten

BSc in Combined Mathematics & Statistics, 2015
University of Victoria

Work Experience

Assistant Professor, 2022-current

Department of Statistics
University of British Columbia

Adjunct Assistant Professor, 2022-current

Department of Statistics and Actuarial Science
University of Waterloo

Assistant Professor, 2020-2022

Department of Statistics and Actuarial Science
University of Waterloo

Data Scientist Intern, Summer 2018
LinkedIn Experimentation Science

Publications and Preprints

Preprints

1. Keller, A., **Gao, L.L.**, Witten, D., Dunham, M. Condition-dependent fitness effects of large synthetic chromosome amplifications, biorXiv:2023.06.08.544269.
2. Dharamshi, A., Neufeld, A., Motwani, K., **Gao, L.L.**, Witten, D. and Bien, J. (2023+) Generalized data thinning using sufficient statistics, arXiv:2303.12931.
3. Neufeld, A., Dharamshi, A., **Gao, L.L.** and Witten, D. (2023+) Data thinning for convolution-closed distributions, arXiv:2301.07276.

4. **Gao, L.L.***, Ye, J.J.*, Yin, H.*, Zeng, S.*, and Zhang, J.* (2023+) Moreau Envelope Based Difference-of-weakly-Convex Reformulation and Algorithm for Bilevel Programs, arxiv:2306.16761.

Published or Accepted

5. **Gao, L.L.***, Ye, J.J.*, Zeng, S.*, and Zhou, J.* (2023+) Necessary and sufficient conditions for multiple objective optimal regression designs, to appear in *Statistica Sinica.*

6. **Gao, L.L.**, Bien, J., and Witten, D. (2024+) Selective inference for hierarchical clustering, to appear in *JASA: Theory and Methods.* **[2023 JASA Reproducibility Award.]**

7. Neufeld, A., **Gao, L.L.**, Popp, J., Battle, A., and Witten, D. (2024) Inference after latent variable estimation for single-cell RNA sequencing data, *Biostatistics*, 25(1), 270-287..

8. Neufeld, A., **Gao, L.L.** and Witten, D. (2022) Tree-values: selective inference for regression trees, *Journal of Machine Learning Research*, 23(305), 1–43.

9. **Gao, L.L.**, Witten, D. and Bien, J. (2022) Testing for association in multi-view network data, *Biometrics*, 78(3), 1018-1030. **[2020 ASA Statistical Learning and Data Science Section Student Paper Award.]**

10. **Gao, L.L.***, Ye, J.J.*, Yin, H.*, Zeng, S.*, and Zhang, J.* (2022) Value function based difference-of-convex algorithm for bilevel hyperparameter selection problems, *ICML 2022*.

11. Liu, P., **Gao, L.L.** and Zhou, J. (2022) R-optimal designs for multi-response regression models with multi-factors. *Communications in Statistics - Theory and Methods*, 51(2), 340-355.

12. **Gao, L.L.**, Bien, J. and Witten, D. (2020) Are clusterings of multiple data views independent? *Biostatistics*, 21(4), 692-708. **[2019 ASA Biometrics Section Student Paper Award.]**

13. **Gao, L.L.*** and Zhou, J.* (2020) Minimax D-optimal designs for multivariate regression models with multi-factors. *Journal of Statistical Planning and Inference*, 209, 160-173.

14. Hsu, E. K., Shaffer, M. L., **Gao, L.**, Sonnenday, C., Volk, M. L., Bucuvalas, J. and Lai, J. C. (2017) Analysis of liver offers to pediatric candidates on the transplant wait list. *Gastroenterology*, 153(4), 988-995.

15. **Gao, L. L.** and Zhou, J. (2017) D-optimal designs based on the second-order least squares estimator. *Statistical Papers*, 58(1), 77-94.

16. **Gao, L. L.** and Zhou, J. (2014) New optimal design criteria for regression models with asymmetric errors. *Journal of Statistical Planning and Inference*, 149, 140-151.

In the above, * indicates alphabetical author ordering. Publications on statistical methodology follow the convention of junior authors listed first and senior authors second, with contribution-based ordering.

—

Presentations

Invited

1. JSM 2023, session on "Recent Developments of Novel Statistical Methods of Analyzing Single-Cell Genomic Data for Precision Medicine and Public Health" (August 2023)
2. North America Machine Learning, Optimization and Statistics Symposium 2023 (June 2023)
3. SAGE/Biostat Seminar at the University of Alberta (March 2023, virtual)
4. University of Victoria Statistics Seminar (February 2023)
5. CMStatistics 2022, session on "Advances in multivariate analysis: Clustering, factor models, and more" (December 2022, hybrid virtual and London UK)
6. UConn Statistics Seminar (November 2022, virtual)
7. The Fifth ICSA-Canada Chapter Symposium, session on "Statistical Learning Methods and Applications" (July 2022, Banff AB)
8. 2022 IMS Annual Meeting, session on "Selective Inference" (June 2022, London UK)
9. 2022 Classification Society Annual Meeting, session on "Biostatistics" (June 2022, Waterloo ON)
10. UMass Amherst Statistics Seminar (March 2022, virtual)
11. University of Bristol Statistics Seminar (March 2022, virtual)
12. USC Biostatistics Seminar (February 2022, virtual)
13. UBC Statistics Seminar (February 2022, virtual)
14. Astro@UofT Statistics and Machine Learning Journal Club (October 2021, virtual)
15. HEC Montreal Decision Sciences Seminar (October 2021, virtual)
16. UofT Statistics Seminar (September 2021, virtual)
17. ICSA Applied Statistics Symposium 2021, session on "Advances in selective and simultaneous inference" (September 2021, virtual)
18. JOBIM 2021, session on "Post-selection inference: valid double-dipping" (July 2021)
19. Harvard-BCH Machine Learning and AI Journal Club (May 2021, virtual)
20. UCSF Biostatistics Seminar (March 2021, virtual)
21. International Seminar on Selective Inference (August 2020, virtual)
22. ETH Zurich Young Data Science Researcher Seminar (June 2020, virtual)
23. University of Waterloo Statistics Seminar (January 2020)

24. McGill University Statistics Seminar (December 2019)
25. University of Victoria Statistics Seminar (September 2018)
26. JSM 2018 in Vancouver BC, session on "Discovering Homology in Multi-View Data: New Statistical Methods for Data Integration" (July 2018)

Contributed

1. SSC 2023 in Ottawa ON, session on "Methods for Genetic Association Studies" (June 2022)
2. JSM 2020, session on "SLDS Student Paper Awards" (July 2020, virtual)
3. WNAR 2020, student paper competition session (June 2020, virtual)
4. JSM 2019 in Denver CO, session on "ASA Biometrics Section JSM Travel Awards (II)" (August 2019)
5. WNAR 2019 in Portland OR, student paper competition session (June 2019)
6. AT&T Graduate Student Symposium in New York City NY (December 2017)
7. BIRS Robustness Theory and Methodology: Recent Advances and Future Directions workshop, in Banff AB (September 2016)
8. JSM 2015 in Seattle WA, session on "New Developments and Applications in Design of Experiments" (August 2015)
9. CSCC (Canadian Statistics Student Conference) 2014 in Toronto ON (May 2014)

Research Funding as Principal Investigator

Internal

New Faculty Start-up Funding: $60,000 CAD 2022
University of British Columbia

New Faculty Start-up Funding: $60,000 CAD 2020-2022
University of Waterloo

External

NSERC Discovery Grants: $90,000 CAD 2021-2026
Natural Sciences and Research Council of Canada

NSERC Discovery Launch Supplement: $12,500 CAD 2021-2022
Natural Sciences and Research Council of Canada

Awards and Scholarships

- (2023) **JASA Reproducibility Award**, *$500 USD*.

- (2020) **University of Washington SPH Outstanding PhD Student**, awarded to a graduating Biostatistics PhD student in recognition of excellence in Coursework & Exam Performance, Thesis & Publications, Written & Oral Skills, and Leadership.
- (2020) **ASA Statistical Learning and Data Science Section Student Paper Award**, for "*Testing for association in multi-view network data*, *$1,000 USD.*
- (2019) **ASA Biometrics Section Student Paper Award**, for "*Are clusterings of multiple data views independent?*", *$1,000 USD.*
- (2016-2019) **NSERC PGSD-3**, a doctoral scholarship from the Natural Sciences and Engineering Council of Canada, *$63,000 CAD.*
- (2015) **CIHR Summer Studentship Award**, an award providing funding for undergraduate research from the Canadian Institutes of Health Research, *$6,250 CAD.*
- (2014) **NSERC Undergraduate Student Research Award**, an award providing funding for undergraduate research from the Natural Sciences and Engineering Council of Canada, *$5,625 CAD.*
- (2013) **NSERC Undergraduate Student Research Award**, *$5,625 CAD* .

Graduate Student Supervision

Master's Essays, supervisor
1. Qiaoyu Liang (2020-2021, University of Waterloo)
   - First position after graduation: PhD student in Statistics at the University of Toronto

2. Steven Mak (2021-2022, University of Waterloo)
3. Thu Nguyen (2022-2023, University of British Columbia)

PhD Dissertations, chair
1. Andrew Kenig (2021-current, University of Waterloo, co-chair w/ Shoja'eddin Chenouri)

Teaching

Instructor at the University of British Columbia

STAT 545A: Exploratory Data Analysis	**Fall 2023**
STAT 545B: Exploratory Data Analysis	**Fall 2023**
STAT 545A: Exploratory Data Analysis	**Fall 2022**
STAT 545B: Exploratory Data Analysis	**Fall 2022**

Instructor at the University of Waterloo

STAT 330: Mathematical Statistics	**Winter 2021**
STAT 330: Mathematical Statistics	**Summer 2021**