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

Academic Positions

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

Publications and Preprints

In the following, * indicates alphabetical author ordering.

Preprints
1. **Gao, L.L.***, Ye, J.J.*, Zeng, S.*, and Zhou, J.* (2022+) Necessary and sufficient conditions for multiple objective optimal regression designs.

Published or Accepted

2. Neufeld, A., **Gao, L.L.**, Popp, J., Battle, A., and Witten, D. (2022+) Inference after latent variable estimation for single-cell RNA sequencing data, to appear in *Biostatistics*.
3. **Gao, L.L.**, Bien, J., and Witten, D. (2022+) Selective inference for hierarchical clustering, to appear in *JASA: Theory and Methods*.

4. 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.

5. **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.]**

6. **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*.

7. 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.

8. **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.]**

9. **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.

10. 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.

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

12. **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.

Software

multiviewtest: Hypothesis Tests for Association Between Subgroups in Two Data Views

- R-Package, hosted on [CRAN](#).

clusterpval: Inference for Estimated Clusters in R

- R-Package, hosted on [Github](#).
- Tutorials available at http://www.lucylgao.com/clusterpval/index.html.

Presentations

Invited

1. UConn Statistics Seminar (November 2022, virtual)
2. The Fifth ICSA-Canada Chapter Symposium, session on "Statistical Learning Methods and Applications" (July 2022, Banff AB)
3. 2022 IMS Annual Meeting, session on "Selective Inference" (June 2022, London UK)

4. 2022 Classification Society Annual Meeting, session on "Biostatistics" (June 2022, Waterloo ON)
5. UMass Amherst Statistics Seminar (March 2022, virtual)
6. University of Bristol Statistics Seminar (March 2022, virtual)
7. USC Biostatistics Seminar (February 2022, virtual)
8. UBC Statistics Seminar (February 2022, virtual)
9. Astro@UofT Statistics and Machine Learning Journal Club (October 2021, virtual)
10. HEC Montreal Decision Sciences Seminar (October 2021, virtual)
11. UofT Statistics Seminar (September 2021, virtual)
12. ICSA Applied Statistics Symposium 2021, session on "Advances in selective and simultaneous inference" (September 2021, virtual)
13. JOBIM 2021, session on "Post-selection inference: valid double-dipping" (July 2021)
14. Harvard-BCH Machine Learning and AI Journal Club (May 2021, virtual)
15. UCSF Biostatistics Seminar (March 2021, virtual)
16. International Seminar on Selective Inference (August 2020, virtual)
17. ETH Zurich Young Data Science Researcher Seminar (June 2020, virtual)
18. University of Waterloo Statistics Seminar (January 2020)
19. McGill University Statistics Seminar (December 2019)
20. University of Victoria Statistics Seminar (September 2018)
21. JSM 2018 in Vancouver BC, session on "Discovering Homology in Multi-View Data: New Statistical Methods for Data Integration" (July 2018)

Contributed

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

- (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* .

Service

Internal

- (2021-2022) Clinical Research Ethics Board Member at the University of Waterloo
- (2020-2022) Department Seminar Committee Member at the University of Waterloo
- (2017-2018) Peer Mentoring Program Member at the University of Washington
- (2017-2018) Co-organizer for working groups at the University of Washington Biostatistics department (Witten Lab and SLAB Lab)

External

- Reviewer for *Journal of Classification, Computational Statistics and Data Analysis, Biometrika, Journal of the American Statistical Association: Theory and Methods, Journal of Machine Learning Research, Statistical Science, Journal of Computational and Graphical Statistics, Journal of the Royal Statistical Society: Series C, Electronic Journal of Statistics, Biostatistics, Statistics in Medicine, and Statistical Papers*

Graduate Student Supervision

&

Graduate Committee Memberships

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)

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

PhD Dissertations, non-chair
- As internal committee member:

   1. Jason Hou-Liu, Department of Statistics and Actuarial Science, University of Waterloo, PhD Candidate (2021-current)

   2. Marzieh Mussavi Rizi, Department of Statistics and Actuarial Science, University of Waterloo, PhD Candidate (2021-current)

   3. Trang Bui, Department of Statistics and Actuarial Science, University of Waterloo, PhD Candidate (2022-current)

4. Yiran Wang, Department of Statistics and Actuarial Science, University of Waterloo, PhD Candidate (2022-current)

Teaching

Instructor at the University of Waterloo

STAT 330: Mathematical Statistics   **Spring 2021**

Two undergraduate sections with ~300 students. Lectures were delivered asynchronously online due to the COVID-19 pandemic. Average student course evaluation score: 4.5/5.

STAT 330: Mathematical Statistics   **Winter 2021**

Two undergraduate sections with ~250 students. Lectures were delivered asynchronously online due to the COVID-19 pandemic. Average student course evaluation score: 4.2/5.

Guest Lectures at the University of Washington

STAT 435: Intro to Statistical Machine Learning   **Winter 2019**

Undergraduate section with ~20 students. My guest lecture topic was unsupervised learning methods.

BIOST 546: ML for Biomedical & Public Health   **Winter 2019**

Graduate section with ~30 students. My guest lecture topic was unsupervised learning methods.

BIOST 311: Regression Methods in Health Sciences   **Spring 2018**

Undergraduate section with ~10 students. My guest lecture topic was regression methods for correlated data.

Teaching Assistant at the University of Washington

BIOST 310: Biostatistics in Health Sciences   **Spring 2018**

Undergraduate section with ~125 students. I taught discussion/tutorial sections once a week for ~30 students, held office hours, and graded assignments. I also developed material for discussion/tutorial sections.

Industry Positions

Data Scientist Intern   **Summer 2018**
LinkedIn Co. Experimentation Science Team