Louis Cammarata

Cambridge, MA • +1 857-707-8309 • louis_cammarata@g.harvard.edu • Web site • LinkedIn

EDUCATION

Harvard University — *Cambridge, MA*
PhD, Statistics — Expected May 2023
Dissertation topic: Network modeling and machine learning with applications to genomics
Advised by Professor Tracy Ke (Harvard Statistics) and Professor Caroline Uhler (MIT EECS)

Massachusetts Institute of Technology, Institute for Data, Systems, and Society — *Cambridge, MA*
MSc, Technology and Policy, 2018 Best Thesis Award — May 2018

École Polytechnique — *Palaiseau, FRANCE*
MSc, Biological Engineering (ranked 34th out of 506) — May 2016
BSc, Mathematics and Sciences — May 2015

FELLOWSHIPS & AWARDS

Bok Center Statistics Pedagogy Fellowship, Harvard University — 2021

Bok Center Certificate of Distinction in Teaching Award, Harvard University — 2019, 2020, 2021

Keystone Symposia Travel Fellowship, Keystone Symposia — 2019

Technology & Policy Program Best Thesis Award, Massachusetts Institute of Technology — 2018

Outstanding Investment Award, École Polytechnique — 2015

PUBLICATIONS & PREPRINTS

Statistical Theory & Methods

Cammarata, L., Jiang, K., Jin, J. and Ke, Z.T. (2022). Estimating Dynamic Mixed Memberships by Trajectory Embedding. Manuscript in preparation.

- Develop dynamic mixed-membership estimation algorithm using spectral methods in the degree-corrected mixed membership stochastic block model and apply algorithm to international trade network dynamics
- Manuscript 80% complete (available upon request)

Cammarata, L. and Ke, Z.T. (2022). Power Enhancement and Phase Transitions for Global Testing of the Mixed Membership Stochastic Block Model. *Bernoulli*.

- Derived fundamental statistical limit of global detection of network communities (i.e., testing whether network has more than one community) and designed a novel statistically optimal hypothesis test
- 79-page article with proofs

Applied Statistics & Machine Learning

Cammarata, L., Shivashankar, G.V. and Uhler, C. (2022). Adhesome Receptor Clustering is Accompanied by the Co-localization of the Associated Genes in the Cell Nucleus. Manuscript in preparation.

- Characterize map between protein assemblies on cell membrane and 3D organization of genome in nucleus, specifically for adhesome in IMR90 fibroblasts using multiple genomic/epigenomic data modalities (RNA-seq, ChIP-seq, in situ Hi-C)
- Project 90% complete (available upon request)

Belyaeva*, A., **Cammarata***, **L.**, Radhakrishnan*, A., Squires, C., Yang, K.D., Shivashankar, G.V. and Uhler, C. (2021). Causal network models of SARS-CoV-2 expression and aging to identify candidates for drug repurposing. *Nature Communications*.

- Developed drug repurposing pipeline using overparametrized representation learning, network analysis and causal structure learning and applied it to identify *in silico* candidate drug targets for treatment of COVID-19
- Media coverage in MIT News Office and other specialized online media
- Co-first authors marked with *

Zhang, J., **Cammarata***, **L.**, Squires*, C., Sapsis, T.P. and Uhler, C. (2022). Active Learning for Optimal Intervention Design in Causal Models. *arXiv*.

- Designed active learning procedure to push system from source state to target state when this system can be described with a structural causal model (relevant applications in cell reprogramming)
- Equal contribution marked with *

Delaney*, C., Schnell*, A., **Cammarata***, **L.**, Yao-Smith, A., Regev, A., Kuchroo, V.K. and Singer, M. (2019). COMET: a tool for marker-panel selection from single-cell transcriptomic data. *Molecular Systems Biology*.

- Developed COMET, a computational tool for combinatorial marker detection from single-cell transcriptomic data available as an online platform and Python package and leveraged tool to identify B cell subpopulations
- Co-first authors marked with *

RELATED PROFESSIONAL EXPERIENCE

PathAI *Boston, MA*
Biomedical Data Science PhD Intern May-Aug 2022

- Executed project to characterize novel biomarkers of tumor micro-environment using deep learning model in Python in collaboration with team of computational biologists and expert pathologists and used markers for prognosis prediction

Google *Virtual/Mountain View, CA*
Data Science PhD Intern (gTrade) May–Aug 2020

- Developed statistical model to predict competition price on Google Ads Exchange using R and SQL on massive data
- Implemented model and sent it for experiments on small percentage of display adds traffic to assess conversion performance

Goldman Sachs *New York City, NY*
Summer Quantitative Analyst (Franchise Analytics team) Jun–Aug 2017

- Designed statistical similarity model to score US corporate bonds and support relative value trading operations

TEACHING EXPERIENCE

Harvard University *Cambridge, MA*
Statistics Teaching Fellow Sep 2019 – May 2022

- Taught STAT303 – The Art and Practice of Teaching and Communicating Statistics (with Prof. Joe Blitzstein and Prof. Morgane Austern) to first-year Statistics PhD students in Fall 2021 and Spring 2022
- Taught STAT210 – Probability I (Fall 2019, Fall 2020 with Prof. Joe Blitzstein), STAT212 – Probability II (Spring 2020 with Prof. Subhabrata Sen) and STAT131 – Time Series (Spring 2020 with Prof. Tracy Ke) to classes of 30-80 students

The Jubilee Institute | MIT Science and Technology Initiative Global Teaching Lab *Amman, JORDAN*
High School Mathematics Teacher Jan 2017

- Designed and taught 5-session introductory "Randomness & Probability" course to 6 cohorts of 25 students (10th grade)

ON-CAMPUS LEADERSHIP

Harvard GSAS Consulting Club | Harvard University *Cambridge, MA*
Vice President of 2022 Harvard/MIT Case Competition Sep 2021–Present

- Organize 12-day flagship event involving 100+ participants in cooperation with consulting firms and industry experts

Harvard Statistics Graduate Committee | Harvard University *Cambridge, MA*
Vice President Sep 2019–Present

- Coordinate yearly Statistics PhD students retreat for ~70 students, faculty and affiliates in Harvard's Statistics Department
- Organize mentoring of ~10 first-year Statistics PhD students by senior graduate students

PolitiX (student non-partisan policy club) | École Polytechnique *Palaiseau, FRANCE*
President Jan 2015–Mar 2016

- Organized monthly conferences with politicians, CEOs and academics for 500+ students – "Outstanding Investment Award"

SELECTED CONFERENCE PRESENTATIONS

Cammarata, L. Adhesome receptor clustering is accompanied by the co-localization of adhesome genes in the cell nucleus. Oral presentation delivered at the Broad Institute Cell Circuits and Epigenomics Seminar Series, Cambridge, MA, June 2022.

Cammarata, L. Network Global Testing: Phase Transitions and Power Enhancement. Oral presentation delivered at the Harvard Mini-Workshop on Algebraic Statistics, Cambridge, MA, April 2022.

Cammarata, L. Drug Repurposing Pipeline Using Overparametrized Representation Learning and Causality. Oral presentation delivered at Purdue University Data Science Week, West Lafayette, IN (virtual), November 2021.

Cammarata, L. Drug Repurposing Pipeline Using Overparametrized Representation Learning and Causality. Oral presentation delivered at Barcelona Supercomputing Center (BSC), Barcelona, Spain (virtual), September 2021.

Cammarata, L. Phase Transitions in Network Global Testing. Oral presentation delivered at Joint Statistical Meetings (JSM), virtual, August 2021.

Cammarata, L. and Belyaeva, A. Causal Network Models of SARS-CoV-2 Expression and Aging to Identify Candidates for Drug Repurposing. Oral presentation delivered at Machine Learning in Computational Biology (MLCB), virtual, November 2020.

Cammarata, L. COMET: Combinatorial Marker Detection from Single-Cell Transcriptomic Data. Spotlight oral presentation delivered at Machine Learning in Computational Biology (MLCB), Vancouver, Canada, December 2019.

Cammarata, L. COMET: Combinatorial Marker Detection from Single-Cell Transcriptomic Data. Poster presented at NeurIPS, Vancouver, Canada, December 2019.

Cammarata, L. COMET: Combinatorial Marker Detection from Single-Cell Transcriptomic Data. Poster presented at Keystone Symposia on Cancer Immunotherapy: Mechanistic Insights to Improve Clinical Benefits (C2), Whistler, Canada, March 2019.

Cammarata, L. Inferring Total Queueing Time Using Only Elapsed Time to Date. Oral presentation delivered at INFORMS Annual Meeting, Houston, TX, October 2017.