Alexander N. D'Amour

Office
Google Cambridge
355 Main St, Floor 5
Cambridge, MA 02142

alexdamour@google.com
(650)793-4899
alexdamour.com

Home
11 Hazel St
Cambridge, MA 02138

Appointments

2018–Present **Research Scientist, Google Brain, Cambridge, MA.**

2016–2018 **Neyman Visiting Assistant Professor of Statistics, UC Berkeley, Berkeley, CA.**

Education

2010–2017 **PhD in Statistics, Harvard University, Cambridge, MA.**
Dissertation: Superpopulation Generalization in Social Network Analysis.

2007–2008 **SM in Applied Mathematics, Harvard University, Cambridge, MA.**

2004–2008 **AB in Applied Mathematics, Harvard University, Cambridge, MA.**

Research

Papers Under Review

A. D'Amour, P. Ding, A. Feller, L. Lei, and J. Sekhon. "Overlap in Observational Studies with High-Dimensional Covariates." https://arxiv.org/abs/1711.02582.

A. Franks, **A. D'Amour**, and A. Feller. "Flexible Senstivitiy Analysis for Observational Studies Without Observable Implications." https://arxiv.org/abs/1809.00399.

Published Papers

2017 A. Miller, N. Foti, **A. D'Amour**, and R. Adams. "Reducing Reparameterization Gradient Variance." In "Advances in Neural Information Processing Systems 30," 2017. https://arxiv.org/abs/1705.07880.

2017 A. Franks, **A. D'Amour**, D. Cervone, and L. Bornn. "Meta-Analytics: Tools for Understanding the Statistical Properties of Sports Metrics." *Jounal of Quantitative Analysis in Sports*, 2017.

2016 D. Cervone, **A. D'Amour**, L. Bornn, and K. Goldsberry. "A Multiresolution Stochastic Process Model for Predicting Basketball Possession Outcomes." *Journal of the American Statistical Association*, 111(514):585–599, 2016.

2014 D. Cervone, **A. D'Amour**, L. Bornn, and K. Goldsberry. "POINTWISE: Predicting Points and Valuing Decisions in Real Time with NBA Optical Tracking Data." *MIT Sloan Sports Analytics Conference*, 2014.

2014 G. Li, R. Lai, **A. D'Amour**, D. Doolin, Y. Sun, V. Torvik, A. Yu, and L. Fleming. "Disambiguation and Co-authorship Networks of the U.S. Patent Inventor Database." *Research Policy*, 43(6):941–955, 2014.

2012 M. Lipsitch, O. Abdullani, **A. D'Amour**, W. Xie, D. Weinberger, E. Tchetgen, and J. Scott. "Estimating Rates of Carriage Acquisition and Clearance and Competitive Ability for Pneumococcal Serotypes in Kenya With a Markov Transition Model." *Epidemiology*, 23(4):510–519, 2012.

2008 R. Acharya, A. Ahmed, **A. D'Amour**, H. Lu, C. Morris, B. Oglevee, A. Peterson, and R. Swift. "Improving Major League Baseball Park Factor Estimates." *Journal of Quantitative Analysis in Sports*, 4(2), 2008.

Papers in Preparation

A. D'Amour and E. Airoldi. "The Effective Estimand." Dissertation chapter. Working draft available at https://alexdamour.com.

A. D'Amour and E. Airoldi. "Misspecification, Sparsity, and Superpopulation Inference for Sparse Social Networks." Dissertation chapter. Working draft available at https://alexdamour.com.

A. D'Amour and E. Airoldi. "Causal Inference with Social-Interaction-Valued Outcomes." Dissertation chapter. Working draft available at https://alexdamour.com.

Invited Talks

2018 "Model-Free Statistical Assessment of Population Overlap." Atlantic Causal Inference Conference at CMU., May 2018. Slides available at: https://goo.gl/1cdXB9.

2018 "Overlap in High Dimensions." European Causal Inference Meeting in Florence, IT., April 2018. **Best early career talk award**. Slides available at: https://goo.gl/cK34mc.

2017 "Overlap in Observational Studies with High-Dimensional Covariates." Berkeley Division of Biostatistics Seminar, October 2017.

2017 "Advances in Basketball Analytics Using Player-Tracking Data." Invited talk at Berkeley Consortium for Data Analytics in Risk, October 2017.

2017 "Overlap in High Dimensions." Invited talk at Atlantic Causal Inference Conference at UNC, May 2017.

2017 "Prediction is Not Enough: Designing decision-support statistics for causal inference." Invited talk at Clarify Health Solutions in San Francisco, CA, May 2017.

2016 "Advances in Basketball Analytics Using Player-Tracking Data." Invited talk at Boston ML Meetup, July 2016.

2015 "Prediction is Not Enough: Designing decision-support statistics for causal inference and attribution." Invited talk at Lumos Labs in San Francisco, CA, October 2015.

2014 "POINTWISE: Predicting Points and Valuing Decisions in Real Time with NBA Optical Tracking Data." Research Paper Competition Finalist Presentation at MIT Sloan Sports Analytics Conference 2014, February 2014.

2013 "Sparse is Different: Covariate Effect Estimation on Sparse Networks." Invited talk presented in the PED Seminar series at MIT Lincoln Laboratory in Lexington, MA, November 2013.

2013 "Multi-Concept Item Response Theory." Talk given at Knewton, Inc. in New York, NY, April 2013.

2012 "Analysis of Sparsity: An Observation Model for Interaction Data." Invited talk presented to the ISR and Tactical Systems Division of MIT Lincoln Laboratory in Lexington, MA, June 2012.

2009 "Dataverse Network Patent Network Database Project." Invited talk and workshop given at the University of Trento X Summer School in Networks and Innovation in Trento, IT, July 2009.

Conference Talks and Posters

2014 "Sparsity Misspecification and Robust Covariate Effect Estimation for Sparse Social Networks." Talk given at the *Joint Statistical Meetings* in Boston, MA, August 2014.

2014 "Real-Time Prediction of Basketball Outcomes Using High-Resolution Spatio-Temporal Tracking Data." Poster presented at the *International Society for Bayesian Analysis World Meeting 2014* in Cancun, Mexico, July 2014.

2012 "Consistent Estimation of Counting Processes on Sparse Networks." Poster presented at the *Neural Information Processing Systems* conference "Social network and social media analysis: Methods, models and applications" workshop in South Lake Tahoe, CA, December 2012.

2011 "Why We Think Obama ♡ Reagan: Feature Selection in Topic Models Using Most Similar Dimension." Talk given at the *New England Statistics Symposium* in Storrs, CT, April 2011.

Media

2014 K. Goldsberry. "Behind DataBall: A Discussion on the Methodology of Expected Possession Value." Grantland. http://grantland.com/the-triangle/behind-databall-a-discussion-on-the-methodology-of-expected-points-value/, 2014.

2014 K. Goldsberry. "DataBall." Grantland. http://grantland.com/features/expected-value-possession-nba-analytics/, 2014.

2013 C. Duffy. "Bayesian." Interview for radio show "You're the Expert", recorded at Oberon Theater, Cambridge, MA, January 2013.

Teaching

Awards

Fall 2014 **2014 Pickard Memorial Teaching Fellow.** Departmental award for sustained excellence in teaching.

Fall 2011, 2012 Spring 2013, 2014 **Harvard University Certificate of Distinction in Teaching.** University commendation for receiving excellent student evaluations.

Courses

2017-2018 **STAT 278B: Causal Inference Reading Group.**
Co-instructor of reading group studying contemporary issues in causal inference. Topics have included high-dimensional causal inference, interference, and optimization approaches.

Spring 2017, Fall 2017 **STAT 153: Introduction to Timeseries.**
Upper-division undergraduate course introducing time-domain and frequency-domain approaches to timeseries analysis.

Summer 2017 **STAT 199: Supervised Independent Study and Research.**
Summer independent study course for several undergraduates interested in experimental and observational causal inference.

Fall 2016 **STAT 88: Probability and Mathematical Statistics in Data Science.**
Introductory course to provide technical depth for students taking Foundations of Data Science.

Spring 2014 **Teaching Fellow for Statistics 225: Spatial Statistics.**
Graduate course introducing core topics in spatio-temporal statistical methods, covering both theoretical approaches and computational methods.

Fall 2013, 2014 **Teaching Fellow for Computer Science 109/Statistics 121: Data Science.**
Introductory course concerning the diverse set of skills necessary for modern data science. Focus on prediction, visualization, Bayesian methods, and empirical model validation.

Spring 2013 **Teaching Fellow for Statistics 221: Statistical Computation and Visualization.**
Graduate course on computational methods and visualization for moderns statistical problems. Topics include building probabilistic models, EM algorithm, MCMC techniques, and visualization using d3.

Fall 2011, 2012 **Teaching Fellow for Statistics 220: Bayesian Data Analysis.**
Core graduate course on Bayesian approaches to model building, model selection, inference, and computation.

Spring 2012 **Teaching Fellow for Statistics 107: Financial Statistics.**
Intro-level course applying basic statistical ideas to trading strategies in financial markets.

Summer 2011 **Teaching Fellow for Statistics S100: Introduction to Statistics.**
Intro-level summer course for college and advanced high school students.

Professional Experience

2017–2018 **Advisor at Clarify Health Solutions, Inc.**
Advise the Data Science team on goals and implementation of key analytical components in Clarify's care management and patient engagement platform.

2015–present **Founding Partner at XY, LLC.**
Sports analytics consultancy, with a focus on player-tracking data. Clients include Philadelphia 76ers, LA Dodgers, Arsenal, San Antonio Spurs, Dallas Mavericks.

2015–present **Founding Partner at Damyata, LLC.**
Data science consultancy. Founded with two partners with deep experience as data scientists in the tech space, Damyata focuses on enabling teams of any size to effectively exploit data. We offer advice and solutions over the full stack of technical, organizational, and analytical challenges that come with ramping up data science operations. Clients include Blue Apron and Demand Signals.

2013–2017 **Data Science Consultant at FirstAccess, Inc.**
Team lead on statistical modeling and experimental design for construction of real-time credit scores for microlenders and development of lending strategies that reduce default rates while expanding access to credit.

2014–2016 **Quantitative Analytics Consultant at Legendary Pictures.**

Advisor for quant analytics unit for Legendary Pictures. Developed a market research tool to predict audience response to potential film properties in the pre-greenlighting stage of production. Currently supervising a project to estimate causal effects of marketing engagement with the intention of producing publications.

2012–2014 **Statistical Consultant at Knewton, Inc.**

Developed statistical tools for an EdTech company's core adaptive teaching platform. Simultaneously assesses student proficiency and measures question difficulty across multiple concept areas.

2009–2011 **Statistical Consultant at Harvard School of Public Health.**

Implemented statistical models of infection dynamics for a variety of contexts. Work featured in an *Epidemiology* publication and a software package prepared for the CDC.

2009–2010 **Statistical Programmer at Harvard Institute for Quantitative Social Science.**

Lead architect on NSF-funded project to construct and distribute the full collaboration network of inventors who have held patents in the United States since 1975.

2008 **Senior Analyst at Lehman Brothers, Fixed Income Analytics.**

Implemented financial models in the Fixed Income Capital Markets Division.

Skills

Programming Advanced proficiency in R, Python, C/C++, LaTeX. Experience with relational databases (MySQL, SQLite), graphical databases (Neo4J), and key-value databases (BerkeleyDB). Extensive development experience on Linux systems and the AWS stack.

Languages Japanese (Intermediate), French (Intermediate).

Interests Cycling, cooking, coffee, soccer, sports analytics.