Lucas Janson

Associate Professor of Statistics

Harvard University
Science Center, Room 710
One Oxford Street
Cambridge, MA 02138-2901

Phone: (617) 998 5658
Fax: (617) 495 1712
Email: ljanson@fas.harvard.edu
Homepage: http://lucasjanson.fas.harvard.edu

Biosketch

Lucas Janson is an Associate Professor in the Department of Statistics and an Affiliate in Computer Science at Harvard University, where he works on high-dimensional inference and statistical machine learning.

Employment

Associate Professor, Harvard University Department of Statistics, Cambridge, MA (2023–present).
Note that at Harvard, Associate Professor is an untenured position.

Assistant Professor, Harvard University Department of Statistics, Cambridge, MA (2017–2023).
Due to substantive disruptions to scholarship, teaching, advising, mentoring, and service/citizenship for all tenure-track faculty, resulting from the COVID-19 pandemic, Harvard University delayed my associate review by two years and gave me one course of teaching relief.

Quantitative Analyst, Goldman Sachs: Quantitative Investment Strategies, New York, NY (2011–2012)

Education

Stanford University (Stanford, CA), Ph.D. in Statistics, 2017. Advisor: Prof. E. Candès.

Stanford University (Stanford, CA), M.S. in Statistics, 2011.

Stanford University (Stanford, CA), B.S. in Mathematics (with Honors) and Physics, 2011.

Affiliations and Memberships

Affiliate, Computer Science area of the Harvard School of Engineering and Applied Sciences, Cambridge, MA (2020–present).

Associate Member, Broad Institute of MIT and Harvard, Cambridge, MA (2019–present).

Faculty Affiliate, Quantitative Biology Initiative at Harvard, Cambridge, MA (2018–present).

Faculty Associate, The Salata Institute for Climate and Sustainability at Harvard University (formerly the Harvard University Center for the Environment), Cambridge, MA (2017–present).

Selected Awards

- Clark Award (Awarded by the Harvard Faculty of Arts and Sciences; $6,000; 2023).

- Best Paper Award in Mathematics (Awarded by The International Congress of Basic Science of China for "Panning for gold: 'model-X' knockoffs for high dimensional controlled variable selection"; $25,000; 2023; *declined because no co-authors were able to attend the required in-person award ceremony*).

- Bernoulli Society New Researcher Award 2023 in Mathematical Statistics (awarded in 2022).

- National Science Foundation DMS-2134157 on "A Theory of Learned Representations in Artificial and Natural Neural Networks" (Co-PI with Boaz Barak, Cengiz Pehlevan, Demba Ba; $1,100,000; 2022–2024).

- National Science Foundation CBET-2112085 on "AI Institute in Dynamic Systems" (SP with PI/Co-PIs Nathan Kutz, Hod Lipson, Steven Brunton, Na Li, Krithika Manohar; $20,000,000; 2021–2026).

- National Science Foundation DMS-2045981 on "CAREER: Beyond Conditional Independence: New Model-Free Targets for High-Dimensional Inference" (PI; $400,000; 2021–2026).

- National Science Foundation IIS-1924984 on "NRI: FND: Robust Grasping by Integrating Machine Learning with Physical Models" (Co-PI with Robert Howe; $749,998; 2019–2022).

- Harvard Milton Fund on "Statistically Removing Population Structure from GWAS" ($50k; 2019).

- Jerome H. Friedman Applied Statistics Dissertation Award (2017).

- Joint Statistical Meetings Student Travel Award (2016).

- Qualcomm Innovation Fellowship (5.5% acceptance rate; 2015).

- Statistics Department Teaching Assistant Award (2015).

Teaching and Training

Courses Taught

- Harvard CS/Stat 184: Introduction to Reinforcement Learning, Fall '22, '23 – Undergraduate Level.

- Harvard Stat 211: Statistical Inference I, Fall '17, '18, '19, '20, '21, '22, '23 – Graduate Level.

- Harvard Stat 195: Statistical Machine Learning, Fall '18, '19, '20 – Undergraduate Level.

Non-Lecture-Based Courses and Reading Groups Run

- Harvard Stat 305: Statistical Consulting, Fall '18, '19, '20, '21, '22, '23, Spring '19, '20, '21, '22, '23 – Graduate Level.

- Harvard Reading Group on Generative AI for Statisticians, Summer '23 – Graduate Level.

- Harvard Reading Group on Convex Optimization, Spring '21 – Graduate Level.

- Harvard Stat 300: Big Ideas in Statistics, Summer '20 – Graduate Level.

- Harvard Stat 314: Timely Topics in Statistics, Fall '18 – Graduate Level.

- Harvard Stat 315: Modern High-Dimensional Inference, Spring '18 – Graduate Level.

- Harvard Stat 303: The Art and Practice of Teaching Statistics, Fall '17, Spring '18 – Graduate Level.

- Stanford STATS 302: Qualifying Exams Workshop, Summer '16 – Graduate Level.

- Stanford STATS 390: Consulting Workshop, Summer '14, '16 – Graduate Level.

Undergraduate Student Supervision

- Kevin Du (Fall 2023–present).
- Grace van Oiste (Fall 2023–present).
- Maximilian Li (Spring 2023–present).
- Vladimir Petrov (Spring 2022–present; HCRP Awardee).
- Virginia Ma (Spring 2022–Spring 2023; PRISE Awardee, Department of Statistics Concurrent Masters Award).
- Massimo Aufiero (Spring 2021–Spring 2022).
- Yash Nair (Spring 2021–Spring 2022; PRISE Awardee, Hoopes Prize, Undergraduate Department of Statistics Prize, NSF GRFP).
- Junu Lee (Spring 2020–Spring 2022; Herschel Smith Fellowhip, NSF GRFP).
- Asher Spector (Fall 2019–Spring 2021; Hoopes Prize, Undergraduate Department of Statistics Prize, NSF GRFP).
- Stephen Casper (Fall 2020–Spring 2021; Hoopes Prize).
- Jason Huang (Fall 2021).
- David Yang (Fall 2019–Spring 2020).
- Abdul Saleh (Fall 2019–Spring 2020).
- Jessica Huang (Fall 2019–Spring 2020).
- Ryan Plunkett (Fall 2018–Spring 2019).
- Wentong Zhang (Fall 2017–Spring 2018).

Masters Student Supervision

- Alec Meade (Fall 2019–Spring 2020).

Doctoral Student Supervision

- Nathan Cheng (Spring 2023–present).
- Benjamin Schiffer (Fall 2022–present).
- Souharya Sengupta (Fall 2022–present).
- Yicong Jiang (Fall 2021–present).
- Ritwik Bhaduri (Fall 2021–present).
- Biyonka Liang (Fall 2020–present; Cox Scholarship Honorable Mention).
- Alexandre Bayle (Spring 2020–present).
- Dae Woong Ham (Fall 2019–present).
- Kelly Zhang (Spring 2019–present; Siebel Scholar).
- Feicheng Wang (Fall 2018–Fall 2022).
- Lu Zhang (Spring 2018–Summer 2022; QBio Student Fellowship, Dempster Award, SLDS student paper award).
- Wenshuo Wang (Spring 2018–Spring 2021; Dempster Award).
- Dongming Huang (Fall 2017–Spring 2020).

Postdoctoral Student Supervision

- Shahriar Talebi (Fall 2023–present).
- M Ganesh Kumar (Fall 2023–present).
- Nikhil Vyas (Fall 2022–present).
- Siyuan Ma (Fall 2019–Summer 2020).

University Service Activities

- Harvard CMSA Conference on Big Data Organizing Committee (2023).
- Harvard Data Science Committee (2021–present).
- Statistics Department Tenure-track Search Committee (2021–2022).
- Electrical Engineering Tenure-track Search Committee (2021–2022).
- Co-organizer of Machine Learning Foundations Seminar (2021–present).
- Statistics Department Graduate Studies Committee (2020–present).
- Harvard Data Science Initiative Postdoctoral Fellow application reviewer (2020).
- William F. Milton Fund proposal reviewer (2020).
- Research Computing Faculty Advisory Group (2019–present).
- First-Generation Faculty Mentorship Program (2019–2021).
- Organizer of Harvard Free Statistical Consulting Service (2018–present).
- Paper Selection Committee for Dempster Award (2018–2021).
- Statistics Department Ph.D. Admissions Committee (2017–2021).
- Statistics Department Undergraduate Studies Committee (2017–present).
- Statistics Department Colloquium Organizer (Spring 2018–Fall 2018; Fall 2021).

Professional Service Activities

Workshops and Conference Committees

- Program Committee, ACM-IMS Foundations of Data Science Conference (2020).
- Scientific Committee, Conference on Statistical Learning and Data Science/Nonparametric Statistics (2020).
- Organizing Committee, Eighth Annual New England Machine Learning Day (2019).

Proposal Reviewer

- National Science Foundation

Outreach

- Activity Co-leader, Harvard Data Adventure Day (2023).
- Pre-doctoral Mentor, National Alliance for Doctoral Studies in the Mathematical Sciences (2021–present).
- Quoted in "Like hitting a bullseye with closed eyes" by Juan Siliezar in Harvard Gazette (2021).
- Guest Lecture "How Journalists Can Understand Data Science", ENG CNSR: Narrative Science Journalism (2019).
- Evaluation Chair, SAILORS: Stanford Artificial Intelligence Laboratory Outreach Summer (2016).
- Consultant, Stanford Statistics Free Consulting Service, over 100 consultees helped (2012-2016).
- Member, Statistics for Social Good Working Group at Stanford University (2013-2016).
- Judge, Seton Middle School Science Fair (2014-2016).
- Guest instructor, 30-minute presentation about data science to 6th grade class at Taft Community School (2015).

Reviewer

Journal of the Royal Statistical Society: Series B, Journal of the American Statistical Association, Annals of Statistics, Biometrika, Annals of the Institute of Statistical Mathematics, Bernoulli, Biometrical Journal, Biometrics, Electronic Journal of Statistics, Journal of Computational and Graphical Statistics, Scandinavian Journal of Statistics, SIAM Journal on Mathematics of Data Science, Statistics in Medicine, Statistica Sinica, Statistics & Probability Letters, Statistical Science, Stat, TEST, ACM-IMS Foundations of Data Science Conference (FODS), Information and Inference, INFORMS Journal on Computing, Communications Methods and Measures, Conference on Learning Theory (COLT), Conference on Neural Information Processing Systems (NeurIPS), Journal of Machine Learning Research, Discrete & Computational Geometry, Econometric Theory, Journal of Climate, Proceedings of the National Academy of Sciences (PNAS), PLOS Biology, International Journal of Robotics Research, American Control Conference (ACC), IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE Conference on Decision and Control (CDC), IEEE International Conference on Robotics and Automation (ICRA), IEEE Robotics and Automation Letters, IEEE Transactions on Aerospace and Electronic Systems, IEEE Transactions on Automation Science and Engineering, IEEE Transactions on Mechatronics, IEEE Transactions on Signal Processing.

Publications

* denotes alphabetized author order

In Preparation/Under Review

[P7] W. Wang, **L. Janson***, L. Lei*, and A. Ramdas*. Total Variation Floodgate for Variable Importance Inference in Classification. 2023. [https://arxiv.org/abs/2309.04002]

[P6] Y. Nair and **L. Janson**. Randomization Tests for Adaptively Collected Data. 2023. [https://arxiv.org/abs/2301.05365]

[P5] B. Liang and **L. Janson**. Powerful Partial Conjunction Hypothesis Testing via Conditioning. 2022. [https://arxiv.org/abs/2212.11304]

[P4] M. Aufiero and **L. Janson**. Surrogate-Based Global Sensitivity Analysis with Statistical Guarantees via Floodgate. 2022. [https://arxiv.org/abs/2208.05885]

[P3] A. Spector and **L. Janson**. Controlled Discovery and Localization of Signals via Bayesian Linear Programming. 2022. [`https://arxiv.org/abs/2203.17208`]

[P2] K. Zhang, **L. Janson**, S. Murphy. Statistical Inference After Adaptive Sampling in Non-Markovian Environments. 2022. [`https://arxiv.org/abs/2202.07098`]

[P1] L. Zhang and **L. Janson**. Floodgate: Inference for Model-Free Variable Importance. 2020. [`https://arxiv.org/abs/2007.01283`]

Journal Articles

[J21] D. W. Ham, K. Imai*, and **L. Janson***. Using Machine Learning to Test Causal Hypotheses in Conjoint Analysis. *Political Analysis* (to appear), 2023+.

[J20] R. F. Barber* and **L. Janson***. Testing Goodness-of-fit and Conditional Independence with Approximate Co-sufficient Sampling. *Annals of Statistics*, 50(5), 2514–2544, 2022.

[J19] F. Wang and **L. Janson**. Exact Asymptotics for Linear Quadratic Adaptive Control. *Journal of Machine Learning Research*, 22(265):1–112, 2021.

[J18] W. Wang and **L. Janson**. A Power Analysis of the Conditional Randomization Test and Knockoffs. *Biometrika*, 109(3):631–645, 2022.

[J17] S. Ma, B. Ren, H. Mallick, Y.S. Moon, E. Schwager, S. Maharjan, T. Tickle, Y. Lu, R. Carmody, E. Franzosa, **L. Janson**, and C. Huttenhower. A Statistical Model for Describing and Simulating Microbial Community Profiles. *PLOS Computational Biology*, 17(9):1–27, 2021.

[J16] A. Spector and **L. Janson**. Powerful Knockoffs via Minimizing Reconstructability. *Annals of Statistics*, 50(1):252–276, 2022.

[J15] M. Liu, E. Katsevich, **L. Janson***, and A. Ramdas*. Fast and Powerful Conditional Randomization Testing via Distillation. *Biometrika*, 109(2):277–293, 2022.

[J14] S. Bates*, E. Candès*, **L. Janson***, and W. Wang*. Metropolized Knockoff Sampling. *Journal of the American Statistical Association*, 116(535):1413–1427, 2021.

[J13] D. Huang, **L. Janson**. Relaxing the Assumptions of Knockoffs by Conditioning. *Annals of Statistics*, 48(5):3021–3042, 2020.

[J12] E. Candès*, Y. Fan*, **L. Janson***, and J. Lv*. Panning for Gold: Model-X Knockoffs for High-dimensional Controlled Variable Selection. *Journal of the Royal Statistical Society: Series B*, 80(3):551–577, 2018.

[J11] Y. Chow, M. Ghavamzadeh, **L. Janson**, and M. Pavone. Risk-Constrained Reinforcement Learning with Percentile Risk Criteria. *Journal of Machine Learning Research*, 18(167):1–51, 2018.

[J10] **L. Janson**, B. Ichter, and M. Pavone. Deterministic Sampling-Based Motion Planning: Optimality, Complexity, and Performance. *International Journal of Robotics Research*, 37(1):46–61, 2018.

[J9] S. Tamang, A. Milstein, H. Sørensen, L. Pedersen, L. Mackey, J. Betterton, **L. Janson**, and N. Shah. Predicting Patient "Cost Blooms" in Denmark: a Longitudinal Population-Based Study. *BMJ Open*, 7(1), 2017.

[J8] **L. Janson**, R. Foygel Barber, and E. Candès. EigenPrism: Inference for High-Dimensional Signal-to-Noise Ratios. *Journal of the Royal Statistical Society, Series B*, 79(4):1037–1065, 2017.

[J7] **L. Janson***, and W. Su*. Familywise Error Rate Control Via Knockoffs. *Electronic Journal of Statistics*, 10(1):960–975, 2016.

[J6] **L. Janson**, W. Fithian, and T. Hastie. Effective Degrees of Freedom: A Flawed Metaphor. *Biometrika*, 102(2):479–485, 2015.

[J5] G. Poultsides, T. Tran, E. Zambrano, **L. Janson**, D. Mohler, M. Well, R. Avedian, B. Visser, J. Lee, K. Ganjoo, E. Harris, J. Norton. Sarcoma Reconstruction With and Without Vascular Reconstruction: A Matched Case-Control Study. *Annals of Surgery*, 262(4):632–640, 2015.

[J4] S. Gholami, **L. Janson**, D. Worhunsky, T. Tran, M. Squires III, L. Jin, G. Spolverato, K. Votanopoulos, C. Schmidt, S. Weber, M. Bloomston, C. Cho, E. Levine, R. Fields, T. Pawlik, S. Maithel, B. Efron, J. Norton, and G. Poultsides. Number of Lymph Nodes Removed and Survival after Gastric Cancer Resection: An Analysis from the US Gastric Cancer Collaborative. *Journal of the American College of Surgeons*, 221(2):291–299, 2015.

[J3] **L. Janson**, E. Schmerling, A. Clark, and M. Pavone. Fast Marching Tree: a Fast Marching Sampling-Based Method for Optimal Motion Planning in Many Dimensions. *International Journal of Robotics Research*, 34(7):883–921, 2015.

[J2] **L. Janson** and B. Rajaratnam. A Methodology for Robust Multiproxy Paleoclimate Reconstructions and Modeling of Temperature Conditional Quantiles. *Journal of the American Statistical Association*, 109(505):63–77, 2014.

[J1] **L. Janson***, M. Klein*, H. Lewis*, A. Lucas*, A. Marantan*, and K. Luna. Undergraduate Experiment in Superconductor Point-Contact Spectroscopy with a Nb/Au Junction. *American Journal of Physics*, 80(2):133–140, 2012.

Refereed Conference Proceedings

[C15] F. Wang and **L. Janson**. Rate-matching the Regret Lower-bound in the Linear Quadratic Regulator With Unknown Dynamics. In *Conference on Decision and Control*, Singapore, December 2023. [`https://arxiv.org/abs/2202.05799`]

[C14] A. Koenig, Z. Liu, **L. Janson**, R. Howe. The Role of Tactile Sensing in Learning and Deploying Grasp Refinement Algorithms. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Kyoto, Japan, October 2022.

[C13] T. Lew, **L. Janson**, R. Bonalli, M. Pavone. A Simple and Efficient Sampling-based Algorithm for General Reachability Analysis. In *Learning for Dynamics & Control Conference*, Stanford, California, June 2022.

[C12] K. Zhang, **L. Janson**, and S. Murphy. Statistical Inference with M-Estimators on Adaptively Collected Data. In *Conference on Neural Information Processing Systems*, December, 2021.

[C11] P. Bayle, A. Bayle, **L. Janson***, and L. Mackey*. Cross-validation Confidence Intervals for Test Error. In *Conference on Neural Information Processing Systems*, December, 2020.

[C10] K. Zhang, **L. Janson***, and S. Murphy*. Inference for Batched Bandits. In *Conference on Neural Information Processing Systems*, December, 2020.

[C9] K. Solovey, **L. Janson**, E. Schmerling, E. Frazzoli, and M. Pavone. Revisiting the Asymptotic Optimality of RRT*. In *IEEE Conference on Robotics and Automation*, Paris, France, May 2020.

[C8] A. Elhafsi, B. Ivanovic, **L. Janson**, and M. Pavone. Map-Predictive Motion Planning in Unknown Environments. In *IEEE Conference on Robotics and Automation*, Paris, France, May 2020.

[C7] **L. Janson**, T. Hu, and M. Pavone. Safe Motion Planning in Unknown Environments: Optimality Benchmarks and Tractable Policies. In *Robotics: Science and Systems*, Pittsburgh, Pennsylvania, June 2018.

[C6] E. Schmerling, **L. Janson**, and M. Pavone. Optimal Sampling-Based Motion Planning under Differential Constraints: the Drift Case with Linear Affine Dynamics. In *Conference on Decision and Control*, Osaka, Japan, December 2015.

[C5] **L. Janson**, B. Ichter, and M. Pavone. Deterministic Sampling-Based Motion Planning: Optimality, Complexity, and Performance. In *International Symposium on Robotics Research*, Sestri Levante, Italy, September 2015.

[C4] **L. Janson***, E. Schmerling*, and M. Pavone. Monte Carlo Motion Planning for Robot Trajectory Optimization Under Uncertainty. In *International Symposium on Robotics Research*, Sestri Levante, Italy, September 2015.

[C3] J. Starek, J. Gomez, E. Schmerling, **L. Janson**, L. Moreno, and M. Pavone. An Asymptotically-Optimal Sampling-Based Algorithm for Bi-directional Motion Planning. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Hamburg, Germany, September 2015.

[C2] E. Schmerling, **L. Janson**, and M. Pavone. Optimal sampling-based motion planning under differential constraints: the driftless case. In *IEEE Conference on Robotics and Automation*, Seattle, Washington, May 2015.

[C1] **L. Janson** and M. Pavone. Fast Marching Trees: a fast marching sampling-based method for optimal motion planning in many dimensions. In *International Symposium on Robotics Research*, Singapore, December 2013.

Refereed Workshop Proceedings

[W2] A. Koenig, Z. Liu, **L. Janson**, R. Howe. Tactile Sensing and its Role in Learning and Deploying Robotic Grasping Controllers. In *IEEE International Conference on Robotics and Automation Workshop: RL for Manipulation*, Philadelphia, Pennsylvania, May 2022.

[W1] **L. Janson** and M. Pavone. Fast Marching Trees: a fast marching sampling-based method for optimal motion planning in many dimensions. In *Robotics: Science and Systems Workshop: Robotic Exploration, Monitoring, and Information Gathering*, Berlin, Germany, June 2013.

Discussion Paper Comments

[D2] **L. Janson**. Discussion of 'A Scale-free Approach for False Discovery Rate Control in Generalized Linear Models' by Chenguang Dai, Buyu Lin, Xin Xing, and Jun S. Liu. *Journal of the American Statistical Association*, 118(543):1584–1585, 2023.

[D1] **L. Janson**. Discussion on 'Random Projection Ensemble Classification'. *Journal of the Royal Statistical Society: Series B*, 79(4):1013–1014, 2017.

Ph.D. Thesis

[T1] **L. Janson**. A Model-Free Approach to High-Dimensional Inference. PhD thesis, Stanford University, Department of Statistics, 2017.

Last updated: November 24, 2023