Lihua Lei

Sequoia Hall, 390 Serra Mall, Stanford University
Stanford, CA 94305-4020

Website: https://lihualei71.github.io
Email: lihualei@stanford.edu

Academic Appointments

Assistant Professor of Economics Stanford Graduate School of Business	7/2022 - present
Assistant Professor of Statistics (by courtesy) Department of Statistics	11/2022 - present
Faculty Fellow Stanford Institute for Economic Policy Research	12/2022 - present
Postdoctoral Researcher, Stanford University Advisor: Emmanuel Candès	9/2019 - 7/2022

Education

Ph.D. Statistics, University of California, Berkeley Advisors: Peter, J. Bickel, and Michael, I. Jordan	8/2014 - 8/2019
B.S. Statistics, Peking University Undergraduate Thesis Advisor: Song Xi, Chen	9/2010 - 6/2014
B.A. Economics (double major), Peking University	9/2011 - 6/2014

Research

Journal Publications [1]

1. Bates, S.*, Candès, E. J.*, **Lei, L.**, Romano, Y.*, and Sesia, M.* (2022+) Testing for Outliers with Conformal p-values. *ArXiv e-prints, abs/2104.08279*. To appear in *Annals of Statistics (AoS)*.

2. Candès, E. J.*, **Lei, L.**, Ren, Z.* (2022+) Conformalized Survival Analysis. *ArXiv e-prints, abs/2103.09763*. To appear in *Journal of the Royal Statistical Society: Series B (JRSS-B)*.

3. Horváth, S.*, **Lei, L.**, Richtárik, P., and Jordan, M. I. (2022) Adaptivity of Stochastic Gradient Methods for Nonconvex Optimization. *SIAM Journal on Mathematics of Data Science (SIMODS)*, 4(2), 634-648.

4. Fithian, W. and **Lei, L.** (2022) Conditional calibration for false discovery rate control under dependence. To appear in *Annals of Statistics (AoS)*.

---

[1] * = alphabetical ordering or equal contribution

5. Bates, S.*, Angelopoulos, A.*, **Lei, L.***, Malik, J., and Jordan, M. I. (2021) Distribution-Free, Risk-Controlling Prediction Sets. *Journal of the ACM (JACM)*, 68(6), 1–34.

6. Loper, J. H.*, **Lei, L.***, Fithian, W., and Tansey, W. (2021) Smoothed Nested Testing on Directed Acyclic Graphs. To appear in *Biometrika*.

7. **Lei, L.** and Candès, E. J. (2021) Conformal Inference of Counterfactuals and Individual Treatment Effects. *Journal of the Royal Statistical Society: Series B (JRSS-B)*. 83(5), 911-938.

8. **Lei, L.** and Ding, P. (2020) Regression Adjustment in Randomized Experiments With A Diverging Number of Covariates. *Biometrika*. 108(4), 815-828.

9. Li, T., **Lei, L.**, Bhattacharyya, S., Sarkar, P., Bickel, P. J., and Levina, E. (2020) Hierarchical community detection by recursive partitioning. To appear in *Journal of the American Statistical Association (JASA)*.

10. **Lei, L.** and Bickel, P. J. (2020) An Assumption-Free Exact Test For Fixed-Design Linear Models With Exchangeable Errors. *Biometrika*, 108(2), 397-412.

11. **Lei, L.**, Ramdas, A., and Fithian, W. (2020) A general interactive framework for FDR control under structural constraints. *Biometrika*, 108(2), 253-267.

12. **Lei, L.** and Jordan, M. I. (2020) On the Adaptivity of Stochastic Gradient-Based Optimization. *SIAM Journal on Optimization (SIOPT)*, 30(2), 1473-1500.

13. D'Amour, A., Ding, P., Feller, A., **Lei, L.**, and Sekhon, J. (2019) Overlap in High Dimensional Observational Studies. *Journal of Econometrics (JoE)*, 221(2), 644-654.

14. **Lei, L.**, and Fithian, W. (2018). AdaPT: An Interactive Procedure For Multiple Testing With Side Information. *Journal of the Royal Statistical Society: Series B (JRSS-B)*, 80(4), 649-679.

15. **Lei, L.**, Bickel, P. J., and El Karoui, N. (2018). Asymptotics For High Dimensional Regression M-Estimates: Fixed Design Results. *Probability Theory and Related Fields (PTRF)*, 172(3-4), 983-1079.

16. Chen, S. X., **Lei, L.**, and Tu, Y. (2016). Functional Coefficient Moving Average Model with Applications to forecasting Chinese CPI. *Statistica Sinica*, 26, 1649-1672.

Conference Publications

1. Elibol, M., **Lei, L.**, and Jordan, M. I. (2020). Variance Reduction with Sparse Gradients. To appear in *International Conference on Learning Representations (ICLR)*.

2. Ye, Y., **Lei, L.**, and Ju, C. (2018). HONES: A Fast and Tuning-free Homotopy Method For Online Newton Step. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*.

3. **Lei, L.**, Ju, C., Chen, J., and Jordan, M. I. (2017). Nonconvex Finite-Sum Optimization Via SCSG Methods. In *Proceedings of the 30th Advances in Neural Information Processing Systems (Neurips)*.

4. **Lei, L.** and Jordan, M. I. (2017). Less than a Single Pass: Stochastically Controlled Stochastic Gradient. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*.

5. **Lei, L.** and Fithian, W. (2016). Power of Ordered Hypothesis Testing. In *Proceedings of the 33th International Conference on Machine Learning (ICML)*.

Under Revision

1. Arkhangelsky, D.*, Imbens, G. W.*, **Lei, L.**, and Luo, X.* (2022) Double-Robust Two-Way-Fixed-Effects Regression For Panel Data. *Arxiv e-prints, abs/2107.13737.* Under R&R at *Quantitative Economics (QE)*.

Preprints and Submissions

1. Sahoo, R., **Lei, L.**, and Wager, S. (2022) Learning from a Biased Sample. *Arxiv e-prints, abs/2209.01754.*

2. Marandon, A., **Lei, L.**, Mary, D., and Roquain, R. (2022) Machine learning meets false discovery rate. *Arxiv e-prints, abs/2208.06685.*

3. Luo, Y., Fithian, W, and Lei, L. (2022) Improving knockoffs with conditional calibration. *Arxiv e-prints, abs/2208.09542.*

4. Angelopoulos, A.*, Bates, S.*, Fisch, A.*, **Lei, L.**, and Schuster T.* (2022) Conformal Risk Control. *Arxiv e-prints, abs/2208.02814.*

5. Angelopoulos, A.*, Bates, S.*, Candès, E. J.*, Jordan, M. I.*, and **Lei, L.*** (2021) Learn then Test: Calibrating Predictive Algorithms to Achieve Risk Control. *Arxiv e-prints, abs/2110.01052.*

6. Yang, C., **Lei, L.**, Ho, N., and Fithian, W. (2021) BONuS: Multiple multivariate testing with a data-adaptive test statistic. *ArXiv e-prints, abs/2106.15743.*

7. **Lei, L.***, Li, X.*, and Lou, X.* (2020) Consistency of Spectral Clustering on Hierarchical Stochastic Block Models. *ArXiv e-prints, abs/2004.14531.*

8. **Lei, L.** (2019) Unified $\ell_{2\to\infty}$ Eigenspace Perturbation Theory for Symmetric Random Matrices. *ArXiv e-prints, abs/1909.04798.*

Software

1. `adaptMT`: R package on Adaptive P-value Thresholding (on CRAN);

2. `cfcausal`: R package on conformal inference of counterfactuals and individual treatment effects (on github);

3. `dbh`: R package on dependence-adjusted Benjamini-Hochberg and general step-up procedures (on github);

4. `mkn`: R package on multiple knockoffs based inference (on github).

5. `ovalue`: R package on distribution-free assessment of population overlap (O-values) for observational studies (on github).

Awards

ICSA Junior Research Award, 2022.

IMS (Institute of Mathematical Statistics) New Researcher Travel Award, 2022.

Rising Star in Data Science (University of Chicago), 2021.

Eric Lehmann Citation (Annual Dissertation Award in Theoretical Statistics), 2019.

Citadel Fellowship, 2017-2018.

Outstanding Graduate Student Instructor Award, 2016.

ICML (International Conference on Machine Learning) Travel Award, 2016.

ACIC (Atlantic Causal Inference Conference) Travel Award, 2018, 2022.

Scholarship from The Sally and Terry Speed Graduate Support Fund, 2015.

Invited Talks

2022: Chamberlain Econometrics Seminar, Online Causal Inference Seminar, UC Berkeley (Econometrics), Joint Statistical Meetings, ICSA China Conference, Workshop on Interactive Causal Learning, IMS Annual in Probability and Statistics, Simons Institute for the Theory of Computing (UC Berkeley), New England Statistics Symposium, Tsinghua IIIS Seminar on Foundations Of Data Science, UNC CIRG Seminar, AEA/ASSA Annual Meeting

2021-22 job talks: University of Michigan (Statistics), USC (Math), Rutgers (Statistics), NYU Courant (Data Science), USC(Statistics), University of Minnesota (Statistics), Northwestern (Statistics), Stanford GSB (Economics), Columbia (Statistics), UPenn Wharton (Statistics), CMU (Statistics), Harvard (Statistics), University of Chicago (Data Science), Chicago Booth (Econometrics), Purdue (Statistics), UC Berkeley (Statistics), Yale (Statistics)

2021: One World YoungStatS Webinar, Journal of Royal Statistical Society Webinar, Stanford Causal Inference Group, University of Washington (Statistics), UC Davis (Econometrics), Harvard (Econometrics Workshop), Temple University (Statistics), NYU Data Science Lunch Seminar, ICSA Applied Statistics Symposium, International Conference on Multiple Comparison Procedures, ICML Workshop on "Neglected Assumptions of Causal Infernece", IFDS Summer Workshop, INFORMS Virtual Healthcare Conference, The Hong Kong University of Science & Technology (Statistics), WNAR Annual Meeting, ETH Zürich (Data Science), UC Berkeley (Statistics), Stanford (Statistics), University of Chicago (Statistics), Stanford RAIN Seminar, UChicago Crime Lab, Chicago Booth (Econometrics), University of Maryland, Baltimore County (Statistics), Dataiku Research, Northwestern (Biostatistics and Statistics), University of Manchester (Biostatistics), The London School of Hygiene & Tropical Medicine, Computer Vision Talks Webinar, Facebook Core Data Science, UCLA (Big Data and ML), University of Missouri (Biostatistics), Leiden Univerity, Twitter, Stitch Fix, Rutgers (Statistics), UPenn Wharton (Statistics)

2020: Lyft, John Hopkins University Causal Inference Group, Florida State University (Guest Lecture), UC Berkeley (Guest Lecture), University of North Carolina at Chapel Hill (Biostatistics), Peking University (Biostatistics), Pacific Causal Inference Conference, Online Causal Inference Seminar, International Seminar on Selective Inference, Southwestern University of Finance and Economics, BAAI Conference, Stanford University (Journal Club), Stanford Data Studio, UC Berkeley (Guest Lecture), WNAR Annual Meeting, UC Berkeley Causal Inference Group

2019: SAMSI Workshop, Facebook Core Data Science, Berkeley-Stanford Econometrics Jamboree, Joint Statistical Meetings, ICSA Applied Statistics Symposium, Yale (Polisci), UC Davis (Statistics), Stanford University (Guest Lecture)

2018: USC (Statistics), University of Michigan (Guest Lecture), University of Michigan (Statistics), International Conference on Econometrics and Statistics, Atlantic Causal Inference Conference, UC Davis (Guest Lecture), Stanford University (Guest Lecture)

2017: Berkeley-Stanford Econometrics Jamboree

Reviewing[2]

Statistics Journals (44): (#papers in parentheses) Annals of Statistics (AoS, 6), Journal of the American Statistical Association (JASA, 6), Biometrika (8), Journal of the Royal Statistical Society-Series B (JRSS-B, 4), Journal of the Royal Statistical Society-Series A (JRSS-A, 1), Bernoulli (1), Electronic Journal of Statistics (EJS, 2), Biometrics (3), Journal of Causal Inference (JCL, 2), Stat (1), Springer Book (1), SpringerBriefs (1), International Journal of Biostatistics (IJB, 1), Journal of Computational and Graphical Statistics (JCGS, 2), Statistica Sinica (1), International Journal of Approximate Reasoning (IJA, 1), Statistics in Biosciences (SIBS, 1), Journal of Statistical Planning and Inference (JSPI, 2).

Economics Journals (6): (#papers in parentheses) Econometrica (2), Journal of Political Economy (JPE, 1), The Review of Economics and Statistics (ReStat, 2), Journal of Applied Econometrics (JAE, 1).

Other Journals (22): SIAM Journal on Mathematics of Data Science (SIMODS, 1), Operations Research (1), Management Science (1), Journal of Machine Learning Research (JMLR, 7), Transactions on Pattern Analysis and Machine Intelligence (TPAMI, 3), Computational Optimization and Applications (COAP, 1), Optimization Methods and Software (GOMS, 1), Artificial Intelligence (1), Information and Inference: A Journal of the IMA (IMAIAI, 2), Vietnam Journal of Mathematics (VJOM, 1), IEEE Transactions on Information Theory (1), American Journal of Epidemiology (AJE, 1), BMC Medical Research Methodology (1)

Conferences: (year in parentheses) ICML (2019, 2020, 2021), NeurlPS (2019, 2020, 2021), COLT (2019, 2020, 2021), AISTATS (2019), UAI (2020, 2021)

Other Academic Services

- Co-organizer of the International Seminar on Selective Inference (with Will Fithian, Rina Barber

---

[2]reviews of revisions not included

and Jelle Goeman, 2020-present).

- Institute of Mathematical Statistics (IMS) Committee on Nomations.

- Co-organizer of the GRoup-Of-Why (GROW) seminar, a causal reading group at Stanford University (with Guillaume Basse and Dominik Rothenhäusler, 2019).