Aaron J. Fisher

Statistics & Machine Learning Researcher

aaronjfisher.com | afishe27@alumni.jh.edu | @ajfishr | Boston, MA

Education

PhD in Biostatistics (2016), *Johns Hopkins Bloomberg School of Public Health*, Baltimore, MD

- Advisors: Vadim Zipunnikov & Brian Caffo
- Awards & Scholarships:
  - Margaret Merrell Award for outstanding research by a biostatistics doctoral student
  - June B. Culley Award for outstanding achievement on an oral examination paper

BA in Economics (2010), *University of Rochester*, Rochester, NY

- Summa cum laude

Professional experience

Sr Statistical Scientist (2020-Present), *Foundation Medicine Inc., Decision Sciences*, Boston, MA

- Personalized treatment effect estimation for clinical decision support tools

Principal Statistician (2019-2020), *Takeda Pharmaceuticals, Statistics and Quantitative* Sciences, Boston, MA

- Analysis of wearable devices in early-stage clinical trials (with Dmitri Volfson)

Postdoctoral Research Fellow (2016-2019), *Department of Biostatistics at the Harvard T.H. Chan School of Public Health*, Boston, MA

- Research on interpretability for machine learning models (with Francesca Dominici & Cynthia Rudin)

Statistical Consultant (2016), *Pfizer*, Boston, MA

- Analysis of wearable devices and temperature probes in human sleep studies

Skills

Statistics & Machine Learning: Causal inference, nonparametric theory, online testing, matrix decompositions, regression in a RKHS, Bayesian regression trees, random forests, neural networks, finite sample bounds, adaptive clinical trials, non-convex quadratic programming.

Computing: R package development, git, Stan, Python, PyTorch, MATLAB, Stata, shell scripting, LaTeX

Summary of selected projects

Personalized Effects: Estimates of conditional average treatment effects allow interventions to be tailored to individuals. A well-known challenge in this field is that estimates can be unstable when the treatment groups are imbalanced. My research highlighted that *inverse-variance weights* mitigate this problem, producing estimators that attain "oracle-efficiency" under the weakest conditions currently available. This work also sheds light on the performance of the popular "R-Learning" method, which implicitly uses a form of inverse variance weighting.

Interpretability for machine learning models: Black-box, proprietary prediction models can provoke distrust and produce predictions that are difficult to combine with supplementary information in order to make decisions. In response, we proposed a method for estimating how much unknown proprietary models rely on different covariates, which we applied to the criminal recidivism model COMPAS. The technical aspects of our work combined approaches from finite sample theory, U-statistics, covering numbers, and non-convex quadratic programming. Our paper was published in *JMLR* and has been cited over 1,570 times.

Visual intuition for influence functions used in causal inference: Influence functions have emerged as a popular, though highly technical, framework for combining machine learning with statistical inference, especially within the field of real world (observational) data analysis. In order to make influence functions more approachable, our educational paper, published in *The American Statistician*, builds intuition based on rigorous, visual illustrations.

Fast bootstrap PCA for millions of covariates: In order to quantify uncertainty for the dominant patterns in brain MRI data, with millions of voxels measured for each patient, we developed a fast bootstrap principal component analysis procedure. Our approach reduces computation time from the order of days to the order of minutes with no loss in accuracy. This work was published in *JASA T&M*, and my associated R package (bootSVD) has been downloaded over 44,000 times.

Selected academic papers (see Google Scholar for full list)

Submitted Manuscripts & Working Papers

A. J. Fisher (2024). The Connection Between R-Learning and Inverse-Variance Weighting for Estimation of Heterogeneous Treatment Effects. arXiv:2307.09700. (link.)

A. J. Fisher & V. Fisher (2023). Three-way Cross-Fitting and Pseudo-Outcome Regression for Estimation of Conditional Effects and other Linear Functionals. arXiv:2306.07230. (link.)

Peer-reviewed publications

A. J. Fisher (2023). Online False Discovery Rate Control for LORD++ and SAFFRON Under Positive, Local Dependence. *Biometrical Journal.* (link.)

A. J. Fisher (2022). Online Control of the False Discovery Rate under "Decision Deadlines." *International Conference on Artificial Intelligence and Statistics* (AISTATS). (link.)

A. J. Fisher & E. H. Kennedy (2020). Visually Communicating and Teaching Intuition for Influence Functions. *The American Statistician*. (link.)

A. J. Fisher, C. Rudin, F. Dominici (2019). All Models are Wrong, but Many are Useful: Learning a Variable's Importance by Studying an Entire Class of Prediction Models Simultaneously. *The Journal of Machine Learning Research*. (link; 1572 citations as of April , 2024, including citations to previous arXiv versions)

A. J. Fisher & M. Rosenblum (2018). Stochastic Optimization of Adaptive Enrichment Designs for Two Subpopulations. *Journal of Biopharmaceutical Statistics*. (link.)

R. Y. Coley, **A. J. Fisher**, M. Mamawala, H. B. Carter, K. J. Pienta, S. L. Zeger (2017). A Bayesian Hierarchical Model for Prediction of Latent Health States from Multiple Data Sources with Application to Active Surveillance of Prostate Cancer. *Biometrics.* (link.)

A. J. Fisher, B. Caffo, B. Schwartz, V. Zipunnikov (2016). Fast, Exact Bootstrap Principal Component Analysis for $p > 1$ Million. *Journal of the American Statistical Association TM.* (link.)

A. J. Fisher, G. B. Anderson, R. Peng, J. Leek (2014). A randomized trial in a massive online open course shows people don't know what a statistically significant relationship looks like, but they can learn. *PeerJ.* (link; 13,876 unique visitors as of April 8, 2024.)

Service

2024-Present: American Statistical Association's (ASA) Committee on Data Science and Artificial Intelligence

2020-Present: Journal of Machine Learning Research (JMLR) Editorial Board

2023: Top 10% of Reviewers at AISTATS

2022: Top 10% of Reviewers at AISTATS

A more complete list of reviews is available on my Web of Science page.