Arjun Sondhi

Education

2019 **PhD, Biostatistics**, *University of Washington*, Seattle, WA.
Research focus: causality and machine learning for network-structured data

2014 **BMath, Statistics & Pure Mathematics**, *University of Waterloo*, Waterloo, ON.

Work Experience

2019-present **Quantitative Scientist**, *Flatiron Health*, New York, NY.
Statistical analysis and methods research for clinico-genomic electronic medical records.

2014-2019 **Research & Teaching Assistant**, *University of Washington*, Seattle, WA.
Developed statistical methods for large network-structured data, with applications to genomics and metabolomics. Taught introductory biostatistics and machine learning courses to public health graduate students.

2018 **Research Scientist Intern**, *Facebook*, Menlo Park, CA.
Developed novel method for improved counterfactual policy evaluation in contextual bandit and reinforcement learning settings.

2017 **Data Scientist Intern**, *Google*, New York, NY.
Implemented machine learning methods to improve causal ad attribution product.

Publications

2020+ David Arbour, Drew Dimmery, and Arjun Sondhi. "Permutation Weighting". Under review.

2020+ Arjun Sondhi and Ali Shojaie. "Two-way network penalized regression with applications to metabolomics profiling data". Under review.

2020+ Jean Feng, Arjun Sondhi, Jessica Perry, and Noah Simon. "Selective prediction-set models with coverage guarantees". Under review.

2020 Arjun Sondhi, David Arbour, and Drew Dimmery. "Balanced off-policy evaluation in general action spaces". Published in *International Conference on Artificial Intelligence and Statistics (AISTATS 2020)*.

2019 Arjun Sondhi and Ali Shojaie. "The Reduced PC-Algorithm: Improved Causal Structure Learning in Large Random Networks". Published in *Journal of Machine Learning Research*.

2017 Arjun Sondhi and Kenneth M. Rice. "Fast permutation tests and related methods, for association between rare variants and binary outcomes". Published in *Annals of Human Genetics*.

2016 Arjun Sondhi and Ali Shojaie. "Causal structure learning with reduced partial correlation thresholding". Published in *IEEE Conference on Data Science and Advanced Analytics (DSAA 2016)*.

Software

`glmfunk` R package with C++ backend implementing generalized linear models for two-way network-structured data.
Available on GitHub: `https://github.com/asondhi/glmfunk`

`AUtests` R package implementing a variety of association tests for rare genetic variants.
Available on CRAN: `https://cran.rstudio.com/web/packages/AUtests/`

Technical Skills

Adv R (tidyverse, Rcpp), SQL

Int Python (numpy, scipy, pandas, scikit-learn)

Basic TensorFlow, PyTorch