Jared S. Murray

The University of Texas
McCombs School of Business
2110 Speedway Stop B6500
Austin, TX 78712

Phone: (512) 471-4130
Office: CBA 6.482
Email: jared.murray@mccombs.utexas.edu
Homepage: https://jaredsmurray.github.io/

Employment

Assistant Professor of Statistics, University of Texas at Austin, Department of Information, Risk, and Operations Management and Department of Statistics and Data Science, 07/2017–present.

Visiting Assistant Professor, Carnegie Mellon University, Department of Statistics, 07/2014–07/2017.

Postdoctoral Associate, Duke University, Department of Statistical Science, 01/2014–06/2014.

Summer Associate, RAND Corporation, 06/2012 - 08/2012.

Education

Ph.D. Statistical Science, Duke University, December 2013.

M.S. Statistical Science, Duke University, December 2011.

B.S. Interdisciplinary Mathematics (Statistics), University of New Hampshire, May 2009.

Research

Grants and Funded Research

National Science Foundation (MMS), "Improving Probabilistic Record Linkage and Subsequent Inference". 10/2016 – 09/2019, $225,000. Principal Investigator.

Patient-Centered Outcomes Research Institute, "Analysis in Distributed Data Networks Workshop". 11/2014–1/2015, $5,200. Workshop Co-organizer.

Preprints/Submitted

(*) denotes student author

1. Bolfarine*, H., C. M. Carvalho, H. F. Lopes, and J. S. Murray (2020). Decoupling Shrinkage and Selection in Gaussian Linear Factor Analysis. arXiv: 2006.11908 [stat.ME].

2. Li*, Y., A. R. Linero, and J. S. Murray (2020). Adaptive Conditional Distribution Estimation with Bayesian Decision Tree Ensembles. arXiv: 2005.02490 [stat.ME].

3. Woody*, S., C. M. Carvalho, P. R. Hahn, and J. S. Murray (2020). Estimating heterogeneous effects of continuous exposures using Bayesian tree ensembles: revisiting the impact of abortion rates on crime. arXiv: 2007.09845 [stat.AP].

4. Woody*, S., C. M. Carvalho, and J. S. Murray (2020a). Bayesian inference for treatment effects under nested subsets of controls. arXiv: 2001.07256 [stat.ME].

5. McVeigh*, B. S., B. T. Spahn, and J. S. Murray (2019). Scaling Bayesian Probabilistic Record Linkage with Post-Hoc Blocking: An Application to the California Great Registers. arXiv: 1905.05337 [stat.ME].

6. Starling*, J. E., C. E. Aiken, J. S. Murray, A. Nakimuli, and J. G. Scott (2019). Monotone function estimation in the presence of extreme data coarsening: Analysis of preeclampsia and birth weight in urban Uganda. arXiv: 1912.06946 [stat.AP].

7. Starling*, J. E., J. S. Murray, P. A. Lohr, A. R. A. Aiken, C. M. Carvalho, and J. G. Scott (2019). Targeted Smooth Bayesian Causal Forests: An analysis of heterogeneous treatment effects for simultaneous versus interval medical abortion regimens over gestation. arXiv: 1905.09405 [stat.AP].

Publications

(*) denotes student author

1. Murray, J. S. (2020+). Log-Linear Bayesian Additive Regression Trees for Multinomial Logistic and Count Regression Models. *Journal of the American Statistical Association*, (to appear). arXiv: 1701. 01503 [stat.ME].

2. Spencer*, N. A. and J. S. Murray (2020+). A Bayesian Hierarchical Model for Evaluating Forensic Footwear Evidence. *Annals of Applied Statistics*, (to appear). arXiv: 1906.05244 [stat.AP].

3. Hahn, P. R., J. S. Murray, and C. M. Carvalho (2020). Bayesian Regression Tree Models for Causal Inference: Regularization, Confounding, and Heterogeneous Effects (with discussion). *Bayesian Analysis*. Advance publication.

4. Hill, J., A. Linero, and J. Murray (2020). Bayesian Additive Regression Trees: A Review and Look Forward. *Annual Review of Statistics and Its Application* **7**(1), 251–278.

5. Starling*, J. E., J. S. Murray, C. M. Carvalho, R. K. Bukowski, and J. G. Scott (2020). BART with targeted smoothing: An analysis of patient-specific stillbirth risk. *Ann. Appl. Stat.* **14**(1), 28–50.

6. Woody*, S., C. M. Carvalho, and J. S. Murray (2020b). Model interpretation through lower-dimensional posterior summarization. *Journal of Computational and Graphical Statistics* **0**(ja), 1–34.

7. Carvalho, C., A. Feller, J. S. Murray, S. Woody*, and D. Yeager (2019). Assessing Treatment Effect Variation in Observational Studies: Results from a Data Challenge. *Observational Studies* **5**, 21–35.

8. Dalmasso*, N., R. Mejia, J. Rodu, M. Price, and J. Murray (2019). Feature Engineering for Entity Resolution with Arabic Names: Improving Estimates of Observed Casualties in the Syrian Civil War. *Artificial Intelligence for Humanitarian Assistance and Disaster Response Workshop, NeurIPS*.

9. Yeager, D. S., P. Hanselman, G. M. Walton, J. S. Murray, R. Crosnoe, C. Muller, E. Tipton, B. Schneider, C. S. Hulleman, C. P. Hinojosa, D. Paunesku, C. Romero, K. Flint, A. Roberts, J. Trott, R. Iachan, J. Buontempo, S. M. Yang, C. M. Carvalho, P. R. Hahn, M. Gopalan, P. Mhatre, R. Ferguson, A. L. Duckworth, and C. S. Dweck (2019). A national experiment reveals where a growth mindset improves achievement. *Nature* **573**(7774), 364–369.

10. Murray, J. S. (2018). Multiple Imputation: A Review of Practical and Theoretical Findings. *Statistical Science* **33**(2), 142–159.

11. McVeigh*, B. S. and J. S. Murray (2017b). Scalable Bayesian Record Linkage. *Advances in Approximate Bayesian Inference Workshop, NIPS*.

12. Hahn, P. R., J. S. Murray, and I. Manolopoulou (2016). A Bayesian Partial Identification Approach to Inferring the Prevalence of Accounting Misconduct. *Journal of the American Statistical Association* **111**(513), 14–26.

13. Muray, J. S. (2016a). Bayesian factor analysis in R: Gaussian, probit and Gaussian copula factor modeling with bfa. *ISBA Bulletin* **23**(3), 11–14.

14. Muray, J. S. (2016b). Review of "Bayesian Statistics for the Social Sciences". *Journal of the American Statistical Association* **111**(513), 440.

15. Murray, J. S. and J. P. Reiter (2016). Multiple Imputation of Missing Categorical and Continuous Values via Bayesian Mixture Models With Local Dependence. *Journal of the American Statistical Association* **111**(516), 1466–1479.

16. Murray, J. S. (2015). Probabilistic Record Linkage and Deduplication after Indexing, Blocking, and Filtering. *Journal of Privacy and Confidentiality* **7**(1).

17. Banerjee, A., J. Murray, and D. Dunson (2013). Bayesian learning of joint distributions of objects. In: *Proceedings of the Sixteenth International Conference on Artificial Intelligence and Statistics.* Ed. by C. M. Carvalho and P. Ravikumar. Vol. 31. Proceedings of Machine Learning Research. Scottsdale, Arizona, USA: PMLR, pp.1–9.

18. Henao, R., J. Murray, G. Ginsburg, L. Carin, and J. E. Lucas (2013). Patient Clustering with Uncoded Text in Electronic Medical Records. In: *AMIA Annual Symposium Proceedings.* Vol. 2013. American Medical Informatics Association, pp.592.

19. Murray, J. S., D. B. Dunson, L. Carin, and J. E. Lucas (2013). Bayesian Gaussian copula factor models for mixed data. *Journal of the American Statistical Association* **108**(502), 656–665.

Technical Reports

(*) denotes student author

1. Hahn, P. R., V. Dorie, and J. S. Murray (2019). Atlantic Causal Inference Conference (ACIC) Data Analysis Challenge 2017. arXiv: `1905.09515 [stat.ME]`.

2. McVeigh*, B. S. and J. S. Murray (2017a). Practical Bayesian Inference for Record Linkage. arXiv: `1710.10558 [stat.ME]`.

Dissertation

"Some Recent Advances in Non- and Semiparametric Bayesian Modeling with Copulas, Mixtures and Latent Variables". Advisor: Jerome P. Reiter

Seminar and Conference Presentations

Invited

1. "Scaling Bayesian Probabilistic Record Linkage with Post-Hoc Blocking: An Application to the California Great Registers". Computational and Methodological Statistics (CMStatistics), London, UK. 12/2019.

2. "Bayesian Nonparametric Models for Treatment Effect Heterogeneity: Parameterization, Prior Choice, and Posterior Summarization". SAMSI, Durham, NC. 12/2019.

3. "Scaling Bayesian Probabilistic Record Linkage with Post-Hoc Blocking: An Application to the California Great Registers". School of Mathematical and Statistical Sciences Department Seminar, Arizona State University, Tempe, AZ. 09/2019.

4. "Bayesian Tree Models for Continuous Treatment Effects". Joint Statistical Meetings, Denver, CO. 07/2019.

5. "Recent developments in model specification, regularization, and summarization for nonparametric Bayesian models of heterogeneous treatment effects". 12th International Conference on Bayesian Nonparametrics, Oxford, UK. 06/2019.

6. "Bayesian nonparametric models for treatment effect heterogeneity: model parameterization, prior choice, and posterior summarization". Bayesian Causal Inference Workshop, Mathematical Biosciences Institue, the Ohio State University, Columbus, OH. 06/2019.

7. "Nonparametric regression models of multilevel treatment effect moderation". Nowrthwestern Depratment of Statistics Seminar Series, Evanston, IL. 05/2019.

8. "When there can be only one: Power from modeling constraints in historical record linkage". Putting the Pieces Together: Promise, Programs, and Pitfalls in Linking Historical and Contemporary Records (Northwestern University's Center for Economic History Workshop), Evanston, IL. 05/2019.

9. "Using Bayesian Causal Forest Models to Examine Treatment Effect Heterogeneity". Society for Research on Educational Effectiveness Spring 2019 Conference, Washington, DC. 03/2019.

10. "Bayesian regression tree models for causal inference: regularization, confounding and heterogeneity". UCLA Department of Statistics Seminar Series, Los Angeles, CA. 10/2018.

11. "Record Linkage". Board Meeting for the LIFE-M Project at the University of Michigan, Ann Arbor MI. 10/2018.

12. "Improving probabilistic record linkage: Accurate links and measures of uncertainty". American Political Science Association Annual Meeting, Boston, MA 8/2018.

13. "Bayesian regression tree models for causal inference: regularization, confounding and heterogeneity". Causal Workshop at UAI 2018, Monterey, CA. 8/2018.

14. "Nonparametric regression models of multilevel treatment effect moderation: The National Study of Learning Mindsets". Joint Statistical Meetings, Vancouver, BC. 7/2018.

15. "Improving probabilistic record linkage: Accurate links and measures of uncertainty". Polmeth, Provo, UT. 6/2018.

16. "Bayesian regression tree models for causal inference: regularization, confounding and heterogeneity". International Society for Bayesian Analysis: 14th World Meeting, Edinburgh, Scotland. 6/2018.

17. "Interpreting complex models: Efficient, valid posterior inference for meaningful quantities". International Society for Bayesian Analysis: 14th World Meeting, Edinburgh, Scotland. 6/2018.

18. "Recent developments in model specification, regularization, and summarization for heterogeneous treatment effects." Bayesian Econometrics Workshop at the International Society for Bayesian Analysis: 14th World Meeting, Edinburgh, Scotland. 6/2018.

19. "Bayesian regression tree models for causal inference: regularization, confounding and heterogeneity". Seminar on Bayesian Inference in Econometrics and Statistics. 5/2018.

20. "Nonparametric regression models of multilevel treatment effect moderation: The National Study of Learning Mindsets". Atlantic Causal Inference Conference, Pittsburgh PA. 5/2018.

21. "Bayesian regression tree models for causal inference: regularization, confounding and heterogeneity". Florida State University Statistics Seminar, Tallahassee, FL. 2/2018.

22. "Interpreting complex models: Efficient, valid posterior inference for meaningful quantities". Computational and Methodological Statistics (CMStatistics), London, UK. 12/2017.

23. "Semiparametric Approaches to Bayesian Inference in Binary Instrumental Variable Models". Joint Statistical Meetings, Baltimore, MD. 07/2017.

24. "Estimating the Prevalence of Accounting Misconduct: A Semiparametric Bayesian Approach". International Chinese Statistical Association Meeting, Chicago, IL. 06/2017.

25. "Probabilistic Record Linkage after Indexing, Blocking, and Filtering". NSF-Census Research Network Spring Meeting, Washington, D.C. 04/2017.

26. "Log-linear Bayesian Additive Regression Trees". New York University PRIISM Seminar Series, New York, NY. 04/2017.

27. "Probabilistic Record Linkage after Indexing, Blocking, and Filtering". Computational and Methodological Statistics (CMStatistics), Seville, Spain. 12/2016.

28. "Log-linear Bayesian Additive Regression Trees". Medical College of Wisconsin, Milwaukee, WI. 11/2016.

29. "Estimating the Prevalence of Accounting Misconduct with Imperfect Data". Center for Accounting Research and Education Conference, Leesburg, VA. 08/2016.

30. "Advances in Bayesian Regression Tree Modeling". Joint Statistical Meetings, Chicago, IL. 08/2016

31. "Probabilistic Record Linkage after Indexing, Blocking, and Filtering". Isaac Newton Institute, Cambridge, UK. 07/2016.

32. "Density Regression with Bayesian Additive Regression Trees". Computational Finance and Econometrics. London, UK. 12/2015.

33. "Multiple Imputation of Missing Categorical and Continuous Values via Bayesian Mixture Models with Local Dependence". Center for Statistical Research and Methodology Seminar Series, U.S. Census Bureau. Suitland, MD. 11/2015.

34. "Multiple Imputation of Missing Categorical and Continuous Values via Bayesian Mixture Models with Local Dependence". Food and Drug Administration Statistics Public Workshop - Planning for and Analysis of Randomized Controlled Clinical Trials and Observational Studies with Missing Data for Tobacco Product Regulatory Submissions. Silver Spring, MD. 11/2015.

35. "Distributed Statistical Model Fitting In Federated Networks: A User Guide". (with Daniella Meeker). AcademyHealth EDM Forum Methods Workgroup, San Diego, CA. 06/2014.

36. "Hierarchically Coupled Mixture Models with Local Dependence as Imputation Engines". SAMSI Computational Methods for Censuses and Surveys Workshop, Washington, DC. 01/2014.

37. "Flexible Bayesian Modeling for Multiple Imputation in the SIPP". 83nd Annual Conference of the Southern Economic Association, Tampa, FL. 11/2013.

38. "Synthesizing Bipartite Graphs: Challenges and New Approaches". Joint Statistical Meetings, San Diego, CA. 08/2012.

Contributed

1. "When there can be only one: The Highlander probability model for historical record linkage". Joint Statistical Meetings, Vancouver, BC. 8/2018.

2. "Semiparametric Approaches to Principal Stratification in Binary Instrumental Variable Models". International Society for Bayesian Analysis: 13th World Meeting, Sardina, Italy. 06/2016.

3. "Semiparametric Bayesian Inference for Partially Identified Conditional Treatment Effects". Atlantic Causal Inference Conference, New York, NY. 05/2016.

4. "Log-linear Bayesian Additive Regression Trees for Classification, Counts and Heteroskedastic Regression". Bayesian Nonparametrics 10, Raleigh, NC. 06/2015.

5. "Tensor Factorization Transformation Priors for Density Regression". Joint Statistical Meetings, Boston, MA. 08/2014.

6. "Density Regression with Bayesian Additive Regression Trees". International Society for Bayesian Analysis: 12th World Meeting, Cancun, Mexico. 07/2014.

7. "Flexible Bayesian Density Regression without Discrete Mixtures". EFaB@Bayes250, Durham, NC. 12/2013. **OUP-EFaB Research Prize**

8. "Joint Stochastic Blockmodeling of Attributed Random Graphs". International Society for Bayesian Analysis: 11th World Meeting, Kyoto, Japan. 06/2012.

Teaching

University of Texas at Austin

Statistics and Modeling - Honors (STA-371H) (Spring 2019, 2020)

Statistics and Modeling (STA-371G) (Fall 2017, Spring 2019)

Carnegie Mellon University

Sampling, Survey and Society (36-303) (Spring 2015, 2016, 2017)

Applied Linear Models (36-617) (Fall 2015, 2016)

Advising

Ph.D. Advisor

Spencer Woody (UT Austin, Department of Statistics and Data Science) (2020, Postdoc at UT Austin)

Brendan McVeigh (Carnegie Mellon University, Department of Statistics and Data Science) (2020, Waymo)

Ph.D. Dissertation Committee Member

Melanie Gonzalez (UT Austin, Department of Psychology) (2020)

Jennifer Starling (UT Austin, Department of Statistics and Data Science) (2020)

Neil Spencer (Carnegie Mellon University, Department of Statistics and Data Science and Machine Learning Department) (2020)

Jared Fisher (2019) (UT Austin, Department of Information, Risk, and Operations Management)

Kirstin Early (2017) (Carnegie Mellon University, Machine Learning Department)

Service/Professional Activities

Department/College/University Service

Canfield Business Honors Program Committee, 2018+

Masters of Statistical Practice Committee (CMU), 2016-2017

Professional Service

Secretary/Treasurer of the American Statistical Association's Statistical Computing Section, 2018-2019

National Academy of Science Committee Member: Standing Committee for the American Opportunity Study – Phase 1, 2016-2017

Reviewer for: *Journal of the American Statistical Association, Annals of Statistics, Biometrika, Annals of Applied Statistics, American Political Science Review, Journal of Computational and Graphical Statistics, Journal of the Royal Statistical Society (Series A), Journal of Business and Economic Statistics, Review of Accounting Studies, Statistics in Medicine, Biostatistics, Biometrics, Journal of Machine Learning Research, Journal of Educational and Behavioral Statistics, Statistics and Computing, Computational Statistics and Data Analysis, Journal of Privacy and Confidentiality, Journal of Statistical Planning and Inference, Statistica Neerlandica, AISTATS 2013*

Software Developed

`bcf`: R package for Bayesian causal forests

`MixedDataImpute`: R package for imputing mixed continuous and ordered/unordered categorical data using nonparametric Bayesian modeling

`bfa`: R package for Bayesian factor analysis. Accommodates mixed continuous and ordered discrete variables with Gaussian, probit or Gaussian copula models.

`medpolish`: Python scripts and standalone GUI implementing median polish algorithms for anomaly detection.

Awards & Honors

ISBA Young Researcher Travel Grant (2014)

Oxford University Press OUP-EFaB Research Prize (2013)

AISTATS 2013 Notable Paper Award

ISBA Young Researcher Travel Grant (2012)

Miscellaneous

Security Clearance: Special Sworn Status, U.S. Census Bureau.

Last updated: August 25, 2020