Daniel J. McDonald

Contact information

University of British Columbia
3106 Earth Sciences Building
2207 Main Mall
Vancouver, BC Canada V6T 1Z4

phone: 604 822 0570
www: https://dajmcdon.github.io/
email: daniel@stat.ubc.ca

Research interests

Machine learning; risk estimation; computational approximations; time series; applications in economics, biology, chemistry, finance and music

Education

2012 Doctor of Philosophy in Statistics
Carnegie Mellon University, Pittsburgh PA, USA
Dissertation: "Generalization error bounds for state space models"
Advisors: Cosma R. Shalizi and Mark Schervish

2008 Master of Science in Statistics
Carnegie Mellon University, Pittsburgh PA, USA

2006 Bachelor of Arts in Economics
Indiana University, Bloomington IN, USA
Summa cum laude

2006 Bachelor of Science in Music and an Outside Field
Indiana University, Bloomington IN, USA
Majors: Cello Performance and Mathematics
Magna cum laude

Academic appointments

2020– **Associate Professor of Statistics**
University of British Columbia

2018–2020 **Associate Professor of Statistics**
Indiana University, Bloomington
Core faculty, Department of Statistics, Program on Data Science
Adjunct Faculty, Department of Computer Science
Affiliate, Center for Algorithms and Machine Learning

2018–2019 **Visiting Associate Professor of Econometrics and Statistics**
The University of Chicago Booth School of Business

2012–2018 **Assistant Professor of Statistics**
Indiana University, Bloomington

2007–2012 **Graduate Research and Teaching Assistant**
Carnegie Mellon University, Pittsburgh

Awards and honors

2018 National Science Foundation CAREER Award

2017 Indiana University Trustees' Teaching Award
2012 Umesh Gavasakar Memorial Thesis Award
2012 American Statistical Association, Pittsburgh Chapter Student of the Year
2006 Phi Beta Kappa
2006 Carroll Christenson Award in Economics
2005 Stadelmann Memorial Scholarship Award in Economics
2004 Mr. & Mrs. Harold E. Strow Award in Economics
2004 Hutton Honors College International Experiences Program Award
2004 Jacobs School of Music Summer Festival Scholarship
2003 James A. Moffat Award for Scholarship in Economics
2002 Alice Freese Honors College Scholarship
2002 Indiana Young Economist Award, Indiana Council of Economic Education
2002 The Reserve Officers Association White River Chapter Scholarship
2002 National Merit Scholarship

Peer-reviewed publications

[17] Policastro, R.A., McDonald, D.J., Brendel, V.P., and Zentner, G.E. (2021), "Flexible analysis of TSS mapping data and detection of TSS shifts with TSRexploreR," *NAR Genomics and Bioinformatics*, 3(2), 1–10.

[16] McDonald, D.J., McBride, M., Gu, Y., and Raphael, C. (2021), "Markov-switching state space models for uncovering musical interpretation," *Annals of Applied Statistics*, forthcoming.

[15] Homrighausen, D., and McDonald, D.J. (2020), "Compressed and penalized linear regression," *Journal of Computational and Graphical Statistics*, 29(2), 309–322.

[14] Khodadadi, A., and McDonald, D.J. (2019), "Algorithms for estimating trends in global temperature volatility," in *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI-19)*, eds. P. V. Hentenryck and Z.-H. Zhou, vol. 33, pp. 614–621, Association for the Advancement of Artificial Intelligence.

[13] Homrighausen, D., and McDonald, D.J. (2018), "A study on tuning parameter selection for the high-dimensional lasso," *Journal of Statistical Computation and Simulation*, 88, 2865–2892.

[12] Ding, L., and McDonald, D.J. (2017), "Predicting phenotypes from microarrays using amplified, initially marginal, eigenvector regression," *Bioinformatics*, 33(14), i350–i358.

[11] McDonald, D.J. (2017), "Minimax Density Estimation for Growing Dimension," in *Proceedings of the* 20^{th} *International Conference on Artificial Intelligence and Statistics (AISTATS)*, eds. A. Singh and J. Zhu, vol. 54, pp. 194–203, PMLR.

[10] McDonald, D.J., Shalizi, C.R., and Schervish, M. (2017), "Nonparametric risk bounds for time-series forecasting," *Journal of Machine Learning Research*, 18(32), 1–40.

[9] Homrighausen, D., and McDonald, D.J. (2017), "Risk consistency of cross-validation for lasso-type procedures," *Statistica Sinica*, 27(3), 1017–1036.

[8] Homrighausen, D., and McDonald, D.J. (2016), "On the Nyström and column-sampling methods for the approximate principal components analysis of large data sets," *Journal of Computational and Graphical Statistics*, 25(2), 344–362.

[7] McDonald, D.J., Shalizi, C.R., and Schervish, M. (2015), "Estimating beta-mixing coefficients via histograms," *Electronic Journal of Statistics*, 9, 2855–2883.

[6] Loewenstein, G., Krishnamurti, T., Kopsic, J., and McDonald, D.J. (2015), "Does increased sexual frequency enhance happiness?" *Journal of Economic Behavior and Organization*, 116, 206–218.

[5] Homrighausen, D., and McDonald, D.J. (2014), "Leave-one-out cross-validation is risk consistent for lasso," *Machine Learning*, 97(1-2), 65–78.

[4] Homrighausen, D., and McDonald, D.J. (2013), "The lasso, persistence, and cross-validation," in *Proceedings of the* 30^{th} *International Conference on Machine Learning (ICML)*, eds. S. Dasgupta and D. McAllester, vol. 28, pp. 1031–1039, PMLR.

[3] Jue, J.J.S., Press, M.J., McDonald, D.J., Volpp, K.G., Asch, D.A., Mitra, N., Stanowski, A.C., and Loewenstein, G. (2012), "The impact of price discounts and calorie messaging on beverage consumption: A multi-site field study," *Preventive Medicine*, 55, 629–633.

[2] McDonald, D.J., Shalizi, C.R., and Schervish, M. (2011), "Estimating beta-mixing coefficients," in *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics (AISTATS)*, eds. G. Gordon, D. Dunson, and M. Dudík, vol. 15, pp. 516–524, PMLR.

[1] McDonald, D.J., and Thornton, D.L. (2008), "Primer on the mortgage market and mortgage finance," *The Federal Reserve Bank of St. Louis Review*, 90(1), 31–46.

Other publications

[1] McDonald, D.J. (2020), "Book review: Sufficient dimension reduction: Methods and applications with R," *Journal of the American Statistical Association*, 115(532), 1032–1033.

Working papers

[5] McDonald, D.J., Bien, J., Green, A., Hu, A.J., DeFries, N., Hyun, S., Oliveira, N.L., Sharpnack, J., Tang, J., Tibshirani, R., Ventura, V., Wasserman, L., and Tibshirani, R.J. (2021), "Can auxiliary indicators improve COVID-19 forecasting and hotspot prediction?" *medRxiv*.

[4] Pham, D., Ding, L., McDonald, D.J., and Mejia, A. (2020+), "Leverage scrubbing: a PCA-based outlier detection method for fMRI," in preparation.

[3] Ding, L., and McDonald, D.J. (2019+), "Sufficient principal component regression for pattern discovery in transcriptomic data," arXiv:2107.02150.

[2] McDonald, D.J., Sharpnack, J., Bassett, R., and Sadhanala, V. (2019+), "Exponential family trend filtering on grids," in preparation.

[1] McDonald, D.J., and Shalizi, C.R. (2019+), "Empirical macroeconomics and DSGE modeling in statistical perspective," in preparation.

Technical reports

[5] McDonald, D.J., and Shalizi, C.R. (2017), "Rademacher complexity of stationary sequences," arXiv:1106.0730.

[4] Homrighausen, D., and McDonald, D.J. (2011), "Spectral approximations in machine learning," arXiv:1107.4340.

[3] McDonald, D.J., Shalizi, C.R., and Schervish, M. (2011), "Estimated VC dimension for risk bounds," arXiv:1111.3404.

[2] McDonald, D.J., Shalizi, C.R., and Schervish, M. (2011), "Generalization error bounds for stationary autoregressive models," arXiv:1103.0942.

[1] McDonald, D.J., Loewenstein, G.F., and Kadane, J. (2009), "The behavior of weight-loss study participants in response to incentives," technical report.

Grants awarded

2021–2025 Regularization and approximation: Statistical inference, model selection, and large data. PI. National Sciences and Engineering Research Council of Canada, Discovery Grant, RGPIN-2021-02618; 135,000 CAD

2020–2021 Detecting anomalies in COVID19 indicators. PI. Canadian Statistical Sciences Institute, Rapid Response Program; 5,000 CAD.

2019–2020 Indiana University Institute for Advanced Studies, Recently Tenured Working Group Research Stipend; 8,000 USD.

2018–2023 CAREER: Calibrating regularization for enhanced statistical inference. PI. National Science Foundation, DMS–1753171; 400,000 USD.

2014–2018 Collaborative research: Statistical and computational efficiency for massive data sets via approximation-regularization. PI. National Science Foundation, DMS–1407439; 90,000 USD.

2014–2016 High-Dimensional Statistics for Macroeconomic Forecasting. Co-PI. (with C. Shalizi as PI), Institute for New Economic Thinking, Grant # INO14–00020; 146,142 USD.

2011–2013 Model Complexity and Prediction Error in Macroeconomic Forecasting. (with C. Shalizi as PI), Institute for New Economic Thinking; 170,000 USD.

Invited presentations

2020 "Compressed and penalized linear regression," International Conference on Computational and Methodological Statistics

"Delphi's COVIDcast Project: An Ecosystem for Tracking and Forecasting the Pandemic," British Columbia COVID-19 Modeling Group

"Minimax density estimation under triangular array asymptotics," International Workshop on Applied Probability (cancelled, COVID19)

"Trend filtering in exponential families," Conference on Statistical Learning and Data Science (cancelled, COVID19)

"Trend filtering in exponential families," Simon Fraser University, Department of Statistics and Actuarial Sciences

"Trend filtering in exponential families," University of British Columbia, Department of Statistics

2019 "Markov-switching state space models for uncovering musical interpretation," Simon Fraser University, Department of Statistics and Actuarial Sciences

"Trend filtering in exponential families," University of Chicago and Toyota Technological Institute Machine Learning Seminar

"Regularization, optimization, and approximation: The benefits of a convex combination," Texas A&M University, Department of Statistics

"Regularization, optimization, and approximation: The benefits of a convex combination," University of Toronto, Department of Statistics

"Regularization, optimization, and approximation: The benefits of a convex combination," Colorado State University, Department of Statistics

"Regularization, optimization, and approximation: The benefits of a convex combination," University of Pittsburgh, Department of Statistics

2018 "Matrix sketching for alternating direction method of multipliers optimization," Symposium on Statistics and Data Science

"A Switching Kalman Filter for Modeling Classical Music Performances," Institute of Mathematical and its Applications Frontiers in Forecasting Workshop

"Statistical implications of (some) computational approximations," University of Virginia, Department of Statistics

2017 "Predicting phenotypes from microarrays using amplified, initially marginal, eigenvector regression," 25^{th} Intelligent Systems for Molecular Biology and the 16^{th} European Conference on Computational Biology

"Compressed and penalized linear regression," Toyota Technological Institute at Chicago, Machine Learning Seminar

"Compressed and penalized linear regression," Pontificia Universidad Católica del Perú, Department of Mathematics and Statistics

"Estimating β-mixing coefficients with histograms," American Mathematical Society Spring Central Sectional Meeting, Special Session on Dependence in Probability and Statistics

2016 "Approximation-regularization for analysis of large data sets," University of Louisville, Department of Bioinformatics and Biostatistics

"Approximation-regularization for analysis of large data sets," University of California, Davis, Department of Statistics

"Approximation-regularization for analysis of large data sets," Indiana University-Purdue University Indianapolis, Department of Biostatistics

"Approximation-regularization for analysis of large data sets," National Center for Atmospheric Research

"Risk estimation for high-dimensional lasso regression," Joint Statistical Meetings

2015 "Approximate principal components analysis of large data sets," Yale University, Department of Statistics

2014 "Approximate principal components analysis of large data sets," Joint Statistical Meetings

"Clustering classical music performance," Université Laval, Department of Mathematics and Statistics

2013 "Clustering classical music performance," 15^{th} IMS New Researchers Conference

2012 "Nonparametric risk bounds for time-series prediction," Yahoo! Research

"Nonparametric risk bounds for time-series prediction," George Mason University, Department of Statistics

"Nonparametric risk bounds for time-series prediction," Bocconi University, Department of Decision Sciences

"Nonparametric risk bounds for time-series prediction," Indiana University, Department of Statistics

"Estimating beta-mixing coefficients," 14^{th} International Conference on Artificial Intelligence and Statistics

"Estimating beta-mixing coefficients," American Statistical Association, Pittsburgh Chapter Annual Meeting

"Spectral approximation methods: performance evaluations in clustering and classification" The Classification Society Annual Meeting

2010 "Generalization error bounds for state-space models with an application to economic forecasting," Joint Statistical Meetings

Invited short courses

2015 "Short course on the predictive viewpoint," Institute for New Economic Thinking Young Scholar Workshop, New York

2013 "Short course on the predictive viewpoint," Institute for New Economic Thinking Young Scholar Workshop, Hong Kong

Contributed talks

2019 "Sufficient principal component regression for genomics," Joint Statistical Meetings

2017 "Compressed and penalized linear regression," Indiana University, Department of Statistics

"Predicting phenotypes from microarrays using amplified, initially marginal, eigenvector regression," Joint Statistical Meetings

2015 "Approximate principal components analysis of large data sets," Indiana University SOIC, Intelligent & Interactive Systems

2014 "Statistical machine learning with structured data," Indiana University, Department of Statistics

2012 "Nonparametric risk bounds for time-series prediction," Indiana University, Department of Economics

Course instruction

University of British Columbia, *2020–*

Stat 535a Computational Statistics - Convex optimization (Winter21)

Stat 406 Methods for Statistical Learning (Winter20)

Indiana University Department of Statistics, *2012–2020*

S301 Introduction to Business Statistics (Fa13, Sp14, Fa14)
S432 Applied Linear Models II (Sp16, Sp17, Sp18)
S682 Topics in Statistical Machine Learning (Sp14)
S721 Advanced Statistical Theory I (Fa12, Fa14)
S722 Advanced Statistical Theory II (Sp13)
S771 Advanced Data Analysis I (Fa16, Sp17, Fa19)
S772 Advanced Data Analysis II (Fa16, Sp17, Fa19)
S782 Topics in Statistical Learning Theory (Fa17)
S785 Seminar on Statistical Theory (Fa17, Sp17, Fa19)

University of Chicago Booth School of Business, *2018*

41911 Advanced Econometrics (Fa18)

Carnegie Mellon University, *2010–2011*

36–226 Introduction to Probability and Statistics II (Su10, Su11)

Student advising

(year indicates actual or anticipated completion)

PhD advisor

2024 Jiaping Liu (UBC Statistics)
2020 Lei Ding (IU Statistics)

PhD dissertation committee

2021 Sanna Wager (PhD Informatics)
2020 Yucong Jiang (IU Computer Science)
2020 Chao Tao (IU Computer Science)
2020 Shay Liu (IU Geology)
2019 Zikun Yang (IU Statistics)
2019 Xuefu Wang (IU Statistics)
2018 Robert Lunde (Carnegie Mellon Statistics)
2017 Rong Jin (IU Informatics)
2015 Yupeng Gu (IU Informatics)

MS thesis advisor

2021 Xiaoxuan Liang (UBC Statistics)
2021 Wei Tang (UBC Statistics)
2020 Haoran Liu (IU Statistics)
2018 Arash Khodadadi (IU Statistics)
2017 Jia Wang (IU Statistics)

PhD exam advisor

2020 Robert Granger (IU Statistics)
2020 Aaron Cohen (IU Statistics)
2017 Lei Ding (IU Statistics)
2017 Raksha Kumaraswamy (IU Computer Science)
2015 Sanna Wager (PhD Informatics)
2014 Zikun Yang (IU Statistics)
2014 Lijiang Guo (IU Statistics)

Undergraduate research

2020 Mackenzie Turner (IU Center for Women in Technology REU)

2018 Michael McBride (IU)

Professional service

Associate Editor
Journal of the American Statistical Association
The American Statistician

Conference Area Chair
International Conference on Learning Representations (ICLR)
International Conference on Machine Learning and Statistics (ICML)

Conference Program Committee
Midwest Machine Learning Symposium (cancelled, COVID19)
Neural Information Processing Systems (NeurIPS)
Conference on Uncertainty in Artificial Intelligence (UAI)
International Conference on Artificial Intelligence and Statistics (AISTATS)
International Conference on Machine Learning and Statistics (ICML)

Referee
Electronic Journal of Statistics
Harvard Data Science Review
IEEE Transactions on Information Theory
International Conference on Learning Theory (COLT)
Journal of the American Statistical Association
Journal of Business and Economic Statistics
Journal of Computational and Graphical Statistics
Journal of Machine Learning Research
Journal of Optimization Theory and Applications
Journal of Statistical Computation and Simulation
Machine Learning Journal
Statistics and Computing
Technometrics

Grant reviewer
Division of Mathematics and Statistics, National Science Foundation
Institute for New Economic Thinking

Other
Awards committee, Student Paper Competition ASA SLDS Section; 2018–
Session organizer, CMStatistics; 2020
Session chair, Joint Statistical Meetings; 2016
Session organizer, Joint Statistical Meetings; 2016

Society memberships
American Statistical Association
Institute of Mathematical Statistics
Bernoulli Society

University service

Indiana University
2019–2020 Academic Fairness Committee, *College of Arts and Sciences*

2019–2020 Center of Excellence for Women & Technology. *REU advisor*

2017–2020 Center of Excellence for Women & Technology, *Faculty Ally*

2015–2020 CrossFit Club, *Faculty advisor*

2016–2018 Faculty Student Mentoring Initiative, *Office of the Vice President for Diversity Equity and Multicultural Affairs*

2014–2018 Data Science Curriculum Committee, *School of Computing, Informatics, and Engineering + College of Arts and Sciences*

Department service

University of British Columbia

2020– Member, Graduate Operations Committee

Indiana University

2019–2020 Chair, Executive Committee

2019–2020 Chair, Teaching-Track Promotion Committee

2019–2020 Computing Committee

2019–2020 Committee on Business Statistics

2018–2020 Teaching-track Promotion Committee

2018–2020 Faculty Hiring Committee

2018–2020 Tenure Review Committee

2016–2020 Graduate Studies Committee

2012–2018 Colloquium Committee

2015–2016 Undergraduate Studies Committee

2015–2016 Executive Committee

2012–2015 Graduate Studies Committee

2012–2015 Faculty Hiring Committee

2013–2014 Chair, Colloquium Committee

Industry experience

2020– **COVID-19 Forecasting team lead**, *Delphi Group, Carnegie Mellon University*, Pittsburgh, PA

2010 **Statistical Consultant**, *Alvarez & Marsal*, New York, New York

2006–2007 **Research Associate**, *Federal Reserve Bank of St. Louis*, St. Louis, Missouri