Jeffrey W. Miller

Harvard School of Public Health	*Email:* jwmiller@hsph.harvard.edu
Department of Biostatistics	*Website:* http://jwmi.github.io/
655 Huntington Ave, 1-419	*Citizenship:* USA
Boston, MA 02115	*Last updated:* May 23, 2018

Professional Appointments

Harvard University, Boston, MA — 2016-present
Assistant Professor, Department of Biostatistics
Affiliate Member, Broad Institute

Duke University, Durham, NC — 2014-2016
Postdoctoral Associate, Department of Statistical Science

Brown University, Providence, RI — Summer 2014
Postdoctoral Associate, Division of Applied Mathematics

Draper Laboratory, Cambridge, MA — 2005-2008
Member of Technical Staff, Cognitive Robotics Group

United States Air Force, Air Force Research Lab, Tyndall AFB, FL — 2002-2005
Highest rank: Captain. Project Manager, Robotics Research Group

Education

Brown University, Providence, RI
PhD, Applied Mathematics. Advisers: Matthew T. Harrison and Stuart Geman — 2008-2014
Master of Science, Mathematics — 2008-2010

PhD Dissertation: *Nonparametric and Variable-Dimension Bayesian Mixture Models: Analysis, Comparison, and New Methods.* Brown University, Division of Applied Mathematics, 2014.

Stanford University, Stanford, CA — 2001-2002
Master of Science, Mechanical Engineering

Georgia Institute of Technology, Atlanta, GA — 1997-2001
Bachelor of Science, Mechanical Engineering

Current Research Interests

Studying longevity and aging at the cellular/molecular level using high-throughput sequencing data. Cancer genome analysis using phylogenetic models and mutational signatures. Cell-free DNA genome sequencing for early cancer detection and monitoring. Nonparametric methods and robustness to model misspecification. Inference, Design, and Experimentation in an Automated Loop (IDEAL).

Publications

J. W. Miller and D. B. Dunson. *Robust Bayesian inference via coarsening.* **Journal of the American Statistical Association (JASA)**, (In press), 2018.

A. Aguirre, J. Nowak, N. Camarda, and 58 others, including J. Miller. *Real-time genomic characterization of advanced pancreatic cancer to enable precision medicine.* **Cancer Discovery**, (Accepted), 2018.

J. W. Miller and M. T. Harrison. *Mixture models with a prior on the number of components.* **Journal of the American Statistical Association (JASA)**, Vol. 113, 2018, pp. 340-356.

B. Betancourt, G. Zanella, J. W. Miller, H. Wallach, A. Zaidi, and B. Steorts. *Flexible models for microclustering with application to entity resolution.* **Advances in Neural Information Processing Systems (NIPS)**, Vol. 29, 2016, pp. 1417-1425.

J. W. Miller, B. Betancourt, A. Zaidi, H. Wallach, and R. C. Steorts. *Microclustering: When the cluster sizes grow sublinearly with the size of the data set.* Bayesian Nonparametrics: The Next Generation workshop, NIPS 2015.

J. W. Miller and M. T. Harrison. *Inconsistency of Pitman–Yor process mixtures for the number of components.* **Journal of Machine Learning Research**, Vol. 15, 2014, pp. 3333-3370.

J. W. Miller and M. T. Harrison. *A simple example of Dirichlet process mixture inconsistency for the number of components.* **Advances in Neural Information Processing Systems (NIPS)**, Full oral presentation, Vol. 26, 2013.

J. W. Miller and M. T. Harrison. *Exact sampling and counting for fixed-margin matrices.* **The Annals of Statistics**, Vol. 41, No. 3, 2013, pp. 1569-1592.

J. W. Miller. *Reduced criteria for degree sequences.* **Discrete Mathematics**, Vol. 313, Issue 4, 2013, pp. 550-562.

Preprints

J. W. Miller. *Fast and accurate approximation of the full conditional for gamma shape parameters.* 2018, arXiv:1802.01610.

J. W. Miller. *A detailed treatment of Doob's theorem.* 2018, arXiv:1801.03122.

J. W. Miller. *An elementary derivation of the Chinese restaurant process from Sethuraman's stick-breaking process.* 2018, arXiv:1801.00513.

M. T. Harrison and J. W. Miller. *Importance sampling for weighted binary random matrices with specified margins.* arXiv:1301.3928.

Grants and Funding

DoD award LC170378, "Ultrasensitive Detection of Subclinical Lung Cancer by Statistical Analysis of Plasma cfDNA-Derived Whole-Genome Sequencing Data", Co-Investigator, 2018-2020. (PI: Scott Carter)

Harvard Data Science Initiative (DSI) Competitive Research Fund Award, "Inference, Design of experiments, and Experimentation in an Automated Loop (IDEAL) for Aging Research", Co-Principal Investigator, 2017-2019. (Co-PI: William Mair)

McLennan Dean's Challenge Grant Program Award, "Studying mechanisms of aging via closed-loop analysis and experimentation", Co-Principal Investigator, 2017-2018. (Co-PI: William Mair)

Cure Alzheimer's Fund award, "Analytical and Statistical Tools for Sequence Analysis for Alzheimer's Disease", Co-Investigator, 2016-2018. (PI: Christoph Lange)

Funding awarded to organize the Radcliffe exploratory seminar, "Statistics When the Model is Wrong," 2017-2018. (Co-organizer: Pierre Jacob)

Graduate School Travel Grant, 2011-2013, Brown University. (Supplemental funding for travel.)

Supported in part by NSF, NIH, and DARPA grants, under advisor supervision, 2011-2015.

National Defense Science and Engineering Graduate (NDSEG) Fellowship, 2001-2002, 2008-2011.

Air Force ROTC Full Scholarship, 1997-2001.

Awards

ISBA New Researchers Travel Award, June 17, 2016 (International Society for Bayesian Analysis 2016 World Meeting). One of two recipients.

Travel award for "Bayesian Nonparametrics: The Next Generation" workshop at NIPS 2015.

Brown University Outstanding Dissertation Award in the Physical Sciences, generously sponsored by the Joukowsky Family Foundation (2014).

1st Prize for Poster by a Young Participant, June 14, 2013, BNP9 (9th Conference on Bayesian Nonparametrics, Amsterdam).

Sigma Xi Outstanding Graduate Student Award, May 9, 2013, Brown University chapter of Sigma Xi. "For excellence in research and high potential for further contributions to science." One of three recipients.

Presidential Award for Excellence in Teaching, May 7, 2012, Brown University. One graduate student is selected for the award each year, out of approximately 400 with teaching positions.

IBM Thomas J. Watson Research Center Student Research Award, April 16, 2011, New England Statistics Symposium (NESS). One of four winners of this award for "outstanding research in the field of Statistics and Probability."

Teaching

Applied Machine Learning (HSPH, BST 263), Spring 2018. Primary instructor.

Advanced Stochastic Modeling (Duke, STA 531), Spring 2016. Primary instructor.

Bayesian and Modern Statistics (Duke, STA 360/601), Spring 2015. Primary instructor.

Information Theory (Brown, APMA 1710), Fall 2011. Primary instructor.

Introduction to Machine Learning (Brown, CSCI 1950-F), Summer 2011. Primary instructor.

Bayesian and Modern Statistics (Duke, STA 360/601), Fall 2014 & 2015. Alternate lecturer.

Recent Applications of Probability and Statistics (Brown, APMA 2610), Spring 2013. Teaching assistant.

Computational Probability and Statistics (Brown, APMA 1690), Fall 2012. Teaching assistant.

Essential Statistics (Brown, APMA 0650), Spring 2012. Teaching assistant.

Math Resource Center (MRC), Fall 2009. Tutor. (Tutoring undergraduates for one evening each week.)

Created over 250 short video lectures (each around 10-15 minutes), freely available online at the *mathematicalmonk* YouTube channel. Topics covered include introductory probability, machine learning, and information theory. To date, the total number of views exceeds 6 million.

Peer-review Activity

Served as referee for:

- Journal of the American Statistical Association

- Journal of the Royal Statistical Society: Series B
- The Annals of Statistics
- Journal of Computational and Graphical Statistics
- Bayesian Analysis
- Statistica Sinica
- Canadian Journal of Statistics
- Statistics and Computing
- Electronic Journal of Statistics
- Neural Information Processing Systems (NIPS)
- International Conference on Machine Learning (ICML)
- Machine Learning
- Advances in Data Analysis and Classification
- Australasian Combinatorics
- SIAM Journal on Discrete Mathematics (SIDMA)

Professional Development and Memberships

Membership in professional societies: ASA, ISBA, AMS, Sigma Xi, MAA, SIAM.

Sheridan Teaching Certificate I, 2012-2013.

Service

Seminar organizer

- Co-organizer of the Biostatistics – Biomedical Informatics Big Data (B3D) Seminar, 2016-2018. This is a seminar series on statistical, computational, and machine learning methods for analyzing large complex data sets, with a focus on applications in biomedical science and public health. (`https://www.hsph.harvard.edu/biostatistics/b3d-seminar/`)
- Co-organizer of the Radcliffe exploratory seminar, "Statistics When the Model is Wrong," May 31 – June 1, 2018. Small workshop with a diverse group of researchers, focused on robustness to model misspecification. (`https://projects.iq.harvard.edu/stat_mod`)
- Co-organizer of the Practical Bayesian Nonparametrics workshop at Neural Information Processing Systems (NIPS) 2016. (`https://sites.google.com/site/nipsbnp2016/`)
- Founder and organizer of the Graduate Student Seminar for Applied Mathematics, 2012-2014. This is a forum for graduate students from Applied Math and other departments to present and discuss their research.
- Co-organizer of the Pattern Theory Group Seminar, 2012-2013. Researchers from Brown and other universities speak about statistics-related topics at this pizza seminar. (`http://www.dam.brown.edu/ptg/seminar.html`)

Departmental committee member

- 2017-2018: Master's of Health Data Science committee, PhD admissions committee, Colloquium committee, PQG travel award review committee.
- 2016-2017: Master's of Health Data Science committee, Colloquium committee.

Dissertation/Thesis committee member

- Michele Zemplenyi (PhD student), 2017-2018. (Co-advisor)

- William Townes (PhD student), 2017-2018.
- Sheila Gaynor (PhD student), 2017-2018.
- Divy Kangeyan (PhD student), 2017-2018.
- Fangyuan Hong (CBQG Masters student), 2017-2018. (Advisor)
- Kelly Mosesso (PhD student), 2016-2018.
- Zachary McCaw (PhD student), 2016-2018.
- Siyuan Ma (PhD student), 2016-2018.
- Wei Yu (CBQG Masters student), 2017.

Other

- Summer Program in Biostatistics & Computational Biology – Application review committee, 2018.
- Panelist in "Academic Careers in Biostatistics" discussion for HSPH students, Nov 29, 2017.
- Chair of the "Personalized Medicine" Contributed Session at JSM 2017.
- Panelist in "Advice on Getting Jobs in (Bio)statistics" discussion for HSPH students, Nov 1, 2016.

Talks

BU Probability and Statistics Seminar, Mar 15, 2018, Boston University. *Robust Bayesian inference via coarsening.* (Invited talk)

WPI Statistics Seminar, Mar 12, 2018, Worcester Polytechnic Institute. *Robust Bayesian inference via coarsening.* (Invited talk)

MyAgeGroup meeting, Mar 3, 2018, Austin, TX. *Statistics and machine learning for the biology of aging.* (Invited talk)

International Workshop on Objective Bayes Methodology (O-Bayes), Dec 11, 2017, University of Texas, Austin. *Inference for cancer phylogenetics.* (Poster, Discussant)

Joint Statistical Meetings (JSM), Jul 31, 2017, Baltimore, MD. *Cancer phylogenies and nonparametric clustering.* (Invited talk)

19th Meeting of New Researchers in Statistics and Probability, Jul 28, 2017, Johns Hopkins University. *Cancer phylogenetic inference.* (Poster)

11th Conference on Bayesian Nonparametrics (BNP11), June 29, 2017, Paris, France. *Several interpretations of the power posterior.* (Invited talk)

ICERM Probabilistic Scientific Computing Workshop, June 8, 2017, Providence, RI. *Inference using partial information.* (Invited talk)

MIT Machine Learning Colloquium, April 26, 2017, Massachusetts Institute of Technology. *Robust Bayesian inference via coarsening.* (Invited talk)

9th International Conference of the ERCIM WG on Computational and Methodological Statistics (CM-Statistics 2016), Dec 11, 2016, Seville, Spain. *Non-standard approaches to nonparametric Bayes.* (Invited talk)

FocuStat L^η Research Kitchen, Oct 12, 2016, University of Oslo, Norway. *Robust Bayesian inference via coarsening.* (Invited talk)

Pattern Theory Seminar, Oct 5, 2016, Brown University. *Robust Bayesian inference via coarsening.* (Invited talk)

Joint Statistical Meetings (JSM), Aug 3, 2016, Chicago, IL. *Robust Bayesian inference via coarsening.* (Invited talk)

International Society for Bayesian Analysis (ISBA) World Meeting, June 14, 2016, Sardinia, Italy. *Robust Bayesian inference via coarsening.* (Invited talk)

Information Theory and Applications (ITA) Workshop, Feb 3, 2016, La Jolla, CA. *Robust Bayesian inference via coarsening.* (Poster)

8th International Conference of the ERCIM WG on Computational and Methodological Statistics (CM-Statistics 2015), Dec 14, 2015, London, UK. *Robust Bayesian inference via coarsening.* (Invited talk)

Bayesian Nonparametrics: The Next Generation workshop at NIPS, Dec 12, 2015, Montreal, Canada. *Non-standard approaches to nonparametric Bayes.* (Invited talk)

Harvard Statistics Departmental Colloquium, Sept 21, 2015, Harvard University. *Robust Bayesian inference via coarsening.* (Invited talk)

Joint Statistical Meetings (JSM), August 11, 2015, Seattle, WA. *Robust Bayesian inference via coarsening.* (Speed talk, poster)

Bayesian Nonparametrics: Synergies between Statistics, Probability and Mathematics, June 30, 2015, SAMSI. *Robust Bayesian inference via coarsening.* (Poster)

10th Conference on Bayesian Nonparametrics (BNP10), June 23, 2015, Raleigh, NC. *An approach to inference under misspecification.* (Talk)

G70: A Celebration of Alan Gelfand's 70th Birthday, April 20, 2015, Duke University. *The small clustering problem: What if the clusters don't grow with N?* (Poster)

Texas A&M Statistics Departmental Colloquium, October 31, 2014, Texas A&M University. *Combinatorial stochastic processes for variable-dimension models.* (Invited talk)

International Society for Bayesian Analysis (ISBA) World Meeting, July 14-18, 2014, Cancún, Mexico. *Combinatorial stochastic processes for variable-dimension models.* (Invited talk)

New England Statistics Symposium (NESS), April 25-26, 2014, Harvard School of Public Health. *Combinatorial stochastic processes for variable-dimension models.* (Talk)

Duke Statistical Science Seminar, February 7, 2014, Duke University. *Combinatorial stochastic processes for variable-dimension models.* (Invited talk)

Neural Information Processing Systems (NIPS), December 5-8, 2013, Lake Tahoe, NV. *A simple example of Dirichlet process mixture inconsistency for the number of components.* (Full oral presentation)

Pattern Theory Seminar, November 6, 2013, Brown University. *Dirichlet process mixture inconsistency for the number of components, and dimension mixture models.* (Invited talk)

REU Seminar (Research Experience for Undergraduates), June 28, 2013, Brown University, Division of Applied Mathematics. *Random matrices with fixed row and column sums.* (Invited talk)

9th Conference on Bayesian Nonparametrics (BNP9), June 10-14, 2013, Amsterdam. *Dimension mixtures of finite-dimensional models.* (Poster) Winner of 1st place in poster competition.

New England Machine Learning day (NEML), May 1, 2013, Cambridge, MA. *Posterior consistency for the number of components in a finite mixture.* (Poster)

New England Statistics Symposium (NESS), April 27, 2013, Storrs, CT. *Posterior consistency for the number of components in a finite mixture.* (Poster)

MathSlam, March 22, 2013, Brown University, Division of Applied Mathematics. *Exact sampling and counting for fixed-margin binary matrices.* (Invited talk)

Brown University Symposium for Undergraduates in the Mathematical Sciences (SUMS), March 9, 2013. *High-dimensional parameter spaces and Fisher information.* (Invited talk)

Neural Information Processing Systems (NIPS), Workshop on Modern Nonparametric Methods in Machine Learning, December 3-8, 2012, Lake Tahoe, NV. *Posterior consistency for the number of components in a finite mixture.* (Speed talk, poster)

Graduate Student Statistics Seminar (GSSS), October 12, 2012, Brown University. *Doob's remarkable theorem on posterior consistency.* (Talk)

ICERM Bayesian Nonparametrics Workshop, September 17-21, 2012, Providence, RI. *Dirichlet process mixtures are inconsistent for the number of components in a finite mixture.* (Talk)

New England Statistics Symposium (NESS), April 16, 2011, Storrs, CT. *A practical algorithm for exact inference on tables.* (Talk) One of four winners of the IBM Student Research Award.

Joint Statistical Meetings (JSM), July 31-August 5, 2010, Vancouver, BC. *A practical algorithm for exact inference on tables.* (Talk)