

DAVID B. DAHL

Professor
Department of Statistics
Brigham Young University

dahl@stat.byu.edu
<https://dahl.byu.edu>

July 13, 2017

EDUCATION

- University of Wisconsin – Madison, Ph.D., Statistics (2004)
Advisor: Michael A. Newton
Conjugate Dirichlet Process Mixture Models: Gene Expression, Clustering, and Efficient Sampling
- Brigham Young University, M.S., Statistics (1998)
Advisor: Scott D. Grimshaw
Comparing Classification Methods with Respect to Sampling Proportion and Model Specification using Evaluation Measures Incorporating Cost
- Brigham Young University, B.S., Statistics (1997)
Summa cum laude, Speaker at college graduation

PROFESSIONAL EXPERIENCE

- Professor, Department of Statistics, Brigham Young University Sept. 2015 - present
- Associate Professor, Department of Statistics, Brigham Young University July 2012 - Aug. 2015
- Associate Professor, Department of Statistics, Texas A&M University Sept. 2010 - June 2012
- Adjunct Assoc. Prof., Dept. of Biostat., U.T. MD Anderson Cancer Center Sept. 2010 - Aug. 2012
- Assistant Professor, Department of Statistics, Texas A&M University Aug. 2004 - Aug. 2010
- Adjunct Asst. Prof., Dept. of Biostat., U.T. MD Anderson Cancer Center Feb. 2007 - Aug. 2010
- Consultant, Department of Ophthalmology, UW-Madison Aug. 1999 - July 2004
- Consultant, College of Agricultural & Life Sciences, UW-Madison Jan. 1999 - July 1999
- Consultant, Department of Statistics, Brigham Young University Jan. 1997 - June 1998
- Intern, 3M Corporation May 1997 - July 1997
- Intern, McBride Biostatistical Research May 1996 - Aug. 1996

RESEARCH INTERESTS: Bayesian nonparametrics, model-based clustering, random partition models, protein structure prediction, bioinformatics, and statistical computing.

PAPERS

- **D. B. Dahl**, R. Day, J. Tsai (2017), Random Partition Distribution Indexed by Pairwise Information, *Journal of the American Statistical Association*, 112, 721-732. Associated software package on CRAN: shallot.
- Q. Li, **D. B. Dahl**, M. Vannucci, H. Joo, J. W. Tsai (2016), KScons: A Bayesian Approach for Protein Residue Contact Prediction using the Knob-socket Model of Protein Tertiary Structure, *Bioinformatics*, 32(24): 3774-3781. Associated software package on CRAN: kscons.
- Q. Li, **D. B. Dahl**, M. Vannucci, H. Joo, J. W. Tsai (2014), Bayesian Model of Protein Primary Sequence for Secondary Structure Prediction, *PLOS ONE*, 9(10), e109832. Associated software package on CRAN: bamboo.
- R. Day, H. Joo, A. Chavan, K. P. Lennox, Y. A. Chen, **D. B. Dahl**, M. Vannucci, J. Tsai (2013), Understanding the General Packing Rearrangements Required for Successful Template Based Modeling of Protein Structure from a CASP Experiment, *Computational Biology and Chemistry*, 42, 40-48.

- S. Schwartz, I. Friedberg, I. V. Ivanov, L. A. Davidson, J. S. Goldsby, **D. B. Dahl**, D. Herman, M. Wang, S. M. Donovan, R. S. Chapkin (2012), A Metagenomic Study of Diet-Dependent Interaction Between Gut Microbiota and Host in Infants Reveals Differences in Immune Response, *Genome Biology*, 2012, 13:r32.
- A. G. Chavan, H. Joo, R. Day, K. P. Lennox, P. Sukhanov, **D. B. Dahl**, M. Vannucci, and J. W. Tsai (2011), Near-Native Protein Loop Modeling using Nonparametric Density Estimation Accommodating Sparsity. *PLoS Computational Biology*, 7(10): e1002234.
- R. Day, K. P. Lennox, **D. B. Dahl**, M. Vannucci, J. W. Tsai (2010), Characterizing the regularity of tetrahedral packing motifs in protein tertiary structure, *Bioinformatics*, 26, 3059-3066.
- K. P. Lennox, **D. B. Dahl**, M. Vannucci, R. Day, J. W. Tsai (2010), A Dirichlet Process Mixture of Hidden Markov Models for Protein Structure Prediction, *Annals of Applied Statistics*, 4, 916-942.
- S. Kim, **D. B. Dahl**, M. Vannucci (2009), Spiked Dirichlet Process Prior for Bayesian Multiple Hypothesis Testing in Random Effects Models, *Bayesian Analysis*, 4, 707-732.
- L. Tao, **D. B. Dahl**, L. M. Pérez, D. H. Russell (2009), The Contributions of Molecular Framework to IMS Collision Cross-sections of Gas-phase Peptide Ions, *Journal of the American Society for Mass Spectrometry*, 20, 1593-1602.
- K. P. Lennox, **D. B. Dahl**, M. Vannucci, J. W. Tsai (2009), Density Estimation for Protein Conformation Angles Using a Bivariate von Mises Distribution and Bayesian Nonparametrics, *Journal of the American Statistical Association*, 104, 586-596.
- **D. B. Dahl** (2009), Modal Clustering in a Class of Product Partition Models, *Bayesian Analysis*, 4, 243-264.
- **D. B. Dahl**, S. Crawford (2009), RinRuby: Accessing the R Interpreter from Pure Ruby, *Journal of Statistical Software*, 29(4), 1-18.
- **D. B. Dahl**, Z. Bohannon, Q. Mo, M. Vannucci, J. W. Tsai (2008), Assessing Side-Chain Perturbations of the Protein Backbone: A Knowledge Based Classification of Residue Ramachandran Space, *Journal of Molecular Biology*, 378, 749-758.
- **D. B. Dahl**, Q. Mo, M. Vannucci (2008), Simultaneous Inference for Multiple Testing and Clustering via a Dirichlet Process Mixture Model, *Statistical Modelling*, 8, 23-39.
- **D. B. Dahl**, M. A. Newton (2007), Multiple Hypothesis Testing by Clustering Treatment Effects, *Journal of the American Statistical Association*, 102, 517-526.
- **D. B. Dahl** (2006), Model-Based Clustering for Expression Data via a Dirichlet Process Mixture Model, in *Bayesian Inference for Gene Expression and Proteomics*, Kim-Anh Do, Peter Müller, Marina Vannucci (Eds.), Cambridge University Press.
- D. Kwon, S. Kim, **D. B. Dahl**, M. Swartz, M. Tadesse and M. Vannucci (2006), Identification of DNA Regulatory Motifs and Regulators by Integrating Gene Expression and Sequence Data, in *Bayesian Inference for Gene Expression and Proteomics*, Kim-Anh Do, Peter Müller, Marina Vannucci (Eds.), Cambridge University Press.
- S. Sengupta, J. A. den Boon, I.-H. Chen, M. A. Newton, **D. B. Dahl**, M. Chen, Y.-J. Cheng, W. H. Westra, C.-J. Chen, A. Hildesheim, B. Sugden, P. Ahlquist (2006), Genome-Wide Expression Profiling Reveals EBV-Associated Inhibition of MHC Class I Expression in Nasopharyngeal Carcinoma, *Cancer Research*, 66, 7999-8006.
- M. A. Croft, A. Glasser, G. Heatley, J. McDonald, T. Ebbert, **D. B. Dahl**, N. V. Nadkarni, P. L. Kaufman (2006), Accommodative Ciliary Body and Lens Function in Rhesus Monkeys, I: Normal Lens, Zonule and Ciliary Process Configuration in the Iridectomized Eye, *Investigative Ophthalmology and Visual Science*, 47, 1076-1086.
- J. A. Mattison, M. A. Croft, **D. B. Dahl**, G. S. Roth, M. A. Lane, D. K. Ingram, P. L. Kaufman (2005), Accommodative Function in Rhesus Monkeys: Effects of Aging and Calorie Restriction, *Journal of the American Aging Association*, 27, 59-67.
- S. Wamsley, B. T. Gabelt, **D. B. Dahl**, G. L. Case, R. W. Sherwood, C. A. May, M. R. Hernandez, P. L. Kaufman (2005), Vitreous Glutamate Concentration and Axon Loss in Monkeys with Experimental

Glaucoma, *Archives of Ophthalmology*, 123, 64-70.

- N. Binkley, **D. B. Dahl**, T. Kawahara-Baccus, D. Krueger, R. J. Colman (2003), Bone Loss Assessment in Ovariectomized Rats Utilizing Two Dual-Energy X-Ray Absorptiometers, *Journal of Bone and Mineral Research*, 18, 370-375.
- N. T. Taylor, G. M. Burlingame, K. B. Kristensen, A. Fuhriman, J. Johansen, **D. B. Dahl** (2001), A Survey of Mental Health Care Provider's and Managed Care Organization Attitudes Toward, Familiarity with, and Use of Group Interventions, *International Journal of Group Psychotherapy*, 51(2), 243-263.
- M. Lopez, N. L. Van Zeeland, **D. B. Dahl**, R. Weindruch, J. M. Aiken (2000), Cellular Phenotypes of Age-Associated Skeletal Muscle Mitochondrial Abnormalities in Rhesus Monkeys, *Mutation Research*, 452, 123-138.

NONREFEREED ITEMS

- **D. B. Dahl** (2014), Book review of "Clustering: A Data Recovery Approach" by Boris Mirkin, *Journal of the American Statistical Association*, 109, 866.
- **D. B. Dahl**, Q. Li, M. Vannucci, H. Joo, J. W. Tsai (2013), A Bayesian Model for Protein Secondary Structure Prediction, *Proceedings 59th ISI World Statistics Congress*, 25-30 August 2013, Hong Kong (Session IPS015), 133-138.
- B. Hartman, **D. B. Dahl** (2010), Bayesian Nonparametric Regression for Diabetes Deaths, *Actuarial Research Clearing House*, 10.1.
- **D. B. Dahl** (2009), Bayesian Methods for Protein Structure Prediction, *International Society for Bayesian Analysis Bulletin*, 16(2), 6-8.
- **D. B. Dahl** (2007), Discussion of Jain and Neal's "Splitting and Merging Components of a Nonconjugate Dirichlet Process Mixture Model," *Bayesian Analysis*, 2, 473-478.

GRANTS

- **PI**: *Nonparametric Bayesian Approaches to Modeling Protein Structure*, Joint NSF/NIGMS Mathematical Biology Program, NIH NIGMS R01 GM104972 (Co-PI Jerry Tsai, Dept. of Chemistry, Univ. of the Pacific & Co-PI Marina Vannucci, Dept. of Statistics, Rice University). 07/01/2012 – 04/30/2017. Total award: \$1,403,580.
- **Co-PI**: *Training Grant: Nutrition, Biostatistics and Bioinformatics*, NIH NCI R25 CA090301 (PI Raymond Carroll, Dept. of Statistics, Texas A&M Univ.). 07/01/2006 – 06/30/2012. Effort: 1.96-3.75%.
- **Subcontract-PI**: *Side Chain Driven Refinement of Protein Structure*, NIH NIGMS R01 GM081631 (PI Jerry Tsai, Dept. of Chemistry, Univ. of the Pacific). 07/01/2007 – 06/30/2010. Subcontract award: ~ \$175,000.

SOFTWARE

- **xtable** — An R package to export tabular data into LaTeX and HTML tables.
- **shallot** — An R package implementing the distributions described in **D. B. Dahl**, R. Day, J. Tsai (2017), Random Partition Distribution Indexed by Pairwise Information, *Journal of the American Statistical Association*, accepted. An R package described in: **D. B. Dahl** (2017), rscala: Integration of R and Scala.
- **rscala** — An R package described in: **D. B. Dahl** (2017), rscala: Integration of R and Scala.
- **bamboo** — An R package implementing the methods described in Q. Li, **D. B. Dahl**, M. Vannucci, H. Joo, and J. W. Tsai (2014), Bayesian Model of Protein Primary Sequence for Secondary Structure Prediction, *PLOS ONE*, 9(10), e109832.
- **modalclust** — Software described in: **D. B. Dahl**. (2009), Modal Clustering in a Class of Product Partition Models, *Bayesian Analysis*, 4, 243-264.
- **RinRuby** — Software described in: **D. B. Dahl**, S. Crawford (2009), RinRuby: Accessing the R

Interpreter from Pure Ruby, *Journal of Statistical Software*, 29(4), 1-18.

- **Isclust** — A program implementing the least squares clustering described in: **D. B. Dahl** (2006), Model-Based Clustering for Expression Data via a Dirichlet Process Mixture Model, in *Bayesian Inference for Gene Expression and Proteomics*, Kim-Anh Do, Peter Müller, Marina Vannucci (Eds.), Cambridge University Press.

INVITED PRESENTATIONS

- International Society for Bayesian Analysis 2016 World Meeting, Sardinia, Italy (June 2016)
- International Biometric Conference, Florence, Italy (July 2014)
- University of Wisconsin, Madison, WI (March 2014)
- Marquette University, Milwaukee, WI (March 2014)
- ISI 59th World Statistics Congress, Hong Kong (August 2013)
- ICERM Conference on Bayesian Nonparametrics, Brown University, Providence, RI (Sept. 2012)
- Joint Statistical Meetings, San Diego, CA (July 2012) [Roundtable Discussion]
- Interface 2012, Rice University, Houston, TX (May 2012)
- Statistical Science Seminar Series, Duke University, Durham, NC (January 2012)
- Joint Statistical Meetings, Miami, FL (August 2011)
- Eighth Workshop on Bayesian Nonparametrics, Veracruz, Mexico (June 2011)
- ENAR of the International Biometric Society, Miami, FL (March 2011)
- Joint Statistical Meetings, Vancouver, BC (August 2010)
- SAMSI Program on Semiparametric Bayesian Inference: Applications in Pharmacokinetics and Pharmacodynamics, Research Triangle Park, NC (July 2010)
- Joint Statistical Meetings, Washington, DC (August 2009)
- WNAR of the International Biometric Society, Portland, OR (June 2009)
- ENAR of the International Biometric Society, San Antonio, TX (March 2009)
- Brigham Young University, Provo, UT (December 2008)
- Classification Society Meeting, St. Louis, MO (June 2008)
- MD Anderson Cancer Center, Department of Biostatistics, Houston, TX (April 2008)
- ENAR of the International Biometric Society, Arlington, VA (March 2008)
- Joint Statistical Meetings, Salt Lake City, UT (August 2007)
- Third Erich L. Lehmann Symposium, Houston, TX (May 2007)
- First Annual Bioinformatics Workshop, College Station, TX (October 2006)
- Joint Research Conference on Statistics in Quality, Industry and Technology, Knoxville, TN (June 2006)
- University of Pennsylvania, Dept. of Biostatistics and Epidemiology, Philadelphia, PA (October 2005)
- Houston Area Chapter of the American Statistical Association, Houston, TX (September 2005)
- MD Anderson Cancer Center, Department of Biostatistics, Houston, TX (February 2005)
- “Second Workshop on Monte Carlo Method,” Cambridge, MA (August 2004)
- “Nucleic Acid Technologies in Animal Health,” College Station, TX (September 2004)

CONTRIBUTED PRESENTATIONS

- Joint Statistical Meetings, Chicago, IL (August 2016)
- Joint Statistical Meetings, Seattle, WA (August 2015)
- 10th Conference on Bayesian Nonparametrics, Raleigh, NC (June 22 2015)
- Joint Statistical Meetings, Boston, MA (August 2014)
- Joint Statistical Meetings, Montréal, QC (August 2013)
- Joint Statistical Meetings, San Diego, CA (July 2012)
- ISBA 11th World Meeting on Bayesian Statistics, Kyoto, Japan (June 2012)
- Valencia / ISBA Ninth International Meeting on Bayesian Statistics, Benidorm, Spain (June 2010)
- “Frontiers of Statistical Decision Making and Bayesian Analysis,” San Antonio, TX (March 2010)
- Joint Statistical Meetings, Denver, CO (August 2008)

- ENAR of the International Biometric Society, Atlanta, GA (March 2007)
- Joint Statistical Meetings, Seattle, WA (August 2006)
- IMS New Researchers Conference, Seattle, WA (August 2006)
- Valencia / ISBA Eighth International Meeting on Bayesian Statistics, Benidorm, Spain (June 2006)
- ENAR of the International Biometric Society, Tampa, Florida (March 2006)
- Joint Statistical Meetings, Minneapolis, MN (August 2005)
- UW-Madison, Department of Biostatistics and Medical Informatics, Madison, WI (January 2004)
- “Statistical Methods for Gene Expression: Microarrays and Proteomics”, Minneapolis, MN (September 2003)
- “International Workshop on Bayesian Data Analysis,” Santa Cruz, CA (August 2003)
- Joint Statistical Meetings, San Francisco, CA (August 2003)
- UW-Madison, Department of Biostatistics and Medical Informatics, Madison, WI (April 2003)
- Joint Statistical Meetings, New York, NY (August 2002)

SERVICE

- NIH Study Section, Biostatistical Methods and Research Design (February 23-24, 2017)
- Program Committee, ENAR Spring Meeting, Washington, DC (2017)
- Chair, Mitchell Prize Committee (2017). *The Mitchell Prize is awarded in recognition of an outstanding paper that describes how a Bayesian analysis has solved an important applied problem.*
- Member, Rank and Status Committee, Dept. of Statistics, BYU (2016-present)
- Chair, Computing Committee, Dept. of Statistics, BYU (2012-present)
- Program Chair, Section on Bayesian Statistical Science, American Statistical Association (2016)
- President, Utah Chapter of American Statistical Association (2015-2016)
- Editor, *Bayesian Analysis* (2013-2015)
- Mentor to New Faculty (Matthew Heaton), Faculty Development Series, BYU (2013-2015)
- Chair, Curriculum Computing Committee, Dept. of Statistics, BYU (2013-2014)
- Associate Editor, Reviews in *Journal of the American Statistical Association* and *The American Statistician* (2011-2013)
- Associate Editor, *Bayesian Analysis* (2009-2012)
- Member, Institute of Mathematical Statistics Committee on New Researchers (2009-2012)
- Member, Student Award Selection Committee, Section on Bayesian Statistical Science, American Statistical Association (2011)
- Vice-President, Southeast Texas Chapter of the American Statistical Association (2008-2011)
- Member, Ph.D. Examination Committee, Dept. of Statistics, TAMU (2007-2012)
- Member, Bioinformatics Faculty Committee, Dept. of Statistics, TAMU (2004-2012)
- Member, Computing Faculty Committee, Dept. of Statistics, TAMU (2005-2012)
- Referee for *American Statistician*, *Annals of Applied Statistics*, *Bayesian Analysis*, *Bioinformatics*, *Biometrics*, *Biostatistics*, *Communications in Statistics - Theory and Methods*, *Electronic Journal of Statistics*, *International Conference on Artificial Intelligence and Statistics*, *Journal of Computational and Graphical Statistics*, *Journal of the Royal Statistical Society: Series B*, *Journal of Statistical Planning and Inference*, *Journal of Statistical Software*, *Journal of the American Statistical Association*, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *Statistical Modelling*, *Statistics and Computing*, and *Statistics in Medicine*

COURSE INSTRUCTION

- Brigham Young University:
 - STAT 121: Principles of Statistics Fall '12
 - STAT 123: Introduction to R Programming Fall '14, '15, Winter '15, '16
 - STAT 223: Applied R Programming '14, '15, Winter '15-'17
 - STAT 230: Analysis of Variance Fall '16
 - STAT 340: Inference Winter '13, '16, '17, Fall '13
 - STAT 624: Statistical Computations Fall '13-'16
- Texas A&M University:
 - STAT 689: Introduction to Bayesian Analysis (*New course*) Spring '11, Fall '11
 - STAT 651: Statistics in Research I Spring '05-'07, Fall '04, '07-'08
 - STAT 605: Advanced Statistical Computations Spring '09, '10
 - STAT 604: Introduction to Statistical Computing Fall '05-'07
 - STAT 302: Statistical Methods Fall '09-'10, Spring '10, '12

ADVISING

- Undergraduates mentored: Brandon Carter (2016-2017), Floid Gilbert (2015-2017), Arthur Lui (2014), Richard Payne (2013)
- M.S. committees chaired: Floid Gilbert (2018), Nathan Bean (2017), Scott Ferguson (2016), Arthur Lui (2015), Deepthi Uppalapati (2012), Bryce Little (2011), Yiyi Wang (2009), Adarsh Joshi (2007, with Marina Vannucci)
- Ph.D. committees chaired: Yiyi Wang (2013 with Faming Liang), Kristin Lennox (2010), Adarsh Joshi (2010 with Valen Johnson), Jaesik Jeong (2008 with Marina Vannucci)

AWARDS

- Faculty Heritage Fellowship in Statistical Science, Dept. of Statistics, BYU (2014-2015)
- Travel Award, "Valencia / ISBA 8th World Meeting on Bayesian Stat.," Benidorm, Spain (June 2006)
- National Eye Institute Traineeship (tuition + stipend) in Biostatistics at UW-Madison (1999-2004)
- Ellis R. Ott Scholarship (\$5000) from the American Society for Quality (1997)
- Research Award (\$1000) from the BYU's Office of Research and Creative Activities (1996)

PROFESSIONAL MEMBERSHIPS

- Institute of Mathematical Statistics (IMS)
- American Statistical Association (ASA) and Utah Chapter of the ASA
 - Sections: Biometrics Section, Biopharmaceutical Section, Business and Economic Statistics Section, Section for Statistical Programmers and Analysts, Section on Bayesian Statistical Sciences, Section on Statistical Learning and Data Sciences, Statistical Computing Section
- International Society for Bayesian Analysis (ISBA)