John W. Emerson

Curriculum Vitae

438 Humphrey Street, New Haven, CT 06511

john.emerson@yale.edu

http://www.stat.yale.edu/~jay/

203-215-3540

Education

1994-2002	Yale University, Ph.D. in Statistics *Dissertation:* Asymptotic Admissibility and Bayesian Estimation *Advisor:* John Hartigan, Eugene Higgins Professor of Statistics
1992-1994	Oxford University, M.Phil. in Economics *Topics of Study:* Micro and Macroeconomic Theory, Industrial Organization, and Econometrics *Thesis:* Finite-Sample Properties of Sample Selection Estimators: Principles and a Case Study
1988-1992	Williams College, B.A. in Economics and Mathematics, Magna Cum Laude *Honors Thesis in Mathematics:* Bayesian Categorical Data Analysis with Multivariate Qualitative Measurements

Academic appointments

2009-present	Associate Professor, Department of Statistics, Yale University
2006-present	Faculty Affiliate, Yale Center for Environment Law and Policy
2007, 2009	Guest Professor, Peking University
2003-2009	Assistant Professor, Department of Statistics, Yale University
2003	Lecturer, Department of Statistics, Yale University
2001	Graduate Instructor for special Data Analysis section for professionals offered at Pfizer, Inc., Groton, CT
1996, 1997, 1999	Graduate Instructor, Department of Statistics, Yale University

Other employment

2002-present	Statistical Consultant, FBI and the U.S. Department of Justice, District of Connecticut. Perform study design and analysis for investigations of insurance and billing fraud.
2007-2009	Advisor/Consultant, REvolution Computing. Consulted on project design, strategic planning and open-source community involvement; taught training classes at Novartis, Sybase, and conferences.
2006-2009	Consultant, Pfizer, Inc. Conducted analysis of electrocardiogram data and developed software in R for use by researchers.

1999-2003	Systems Administrator, Department of Geology and Geophysics, Yale University. Performed Solaris/Linux/Windows systems design and maintenance (extensive Perl and shell scripting), supporting research and teaching of faculty and staff (full-time).
1995-1999	UNIX Systems Administrator, Department of Statistics, Yale University. Maintained departmental computing resources (part-time).
1996-1998	Marketing Research Consultant, CBS Network Television. Developed software in C and Perl for exploration of Nielsen panel viewer data.

Academic honors and professional memberships

- John M. Chambers Statistical Software Award, 2010; with Michael Kane
- Member, Institute of Mathematical Statistics, 2010-present
- Fellow, Connecticut Academy of Arts and Sciences, 2009-present
- Member, American Statistical Association, 1998-present
- Winner, Student Paper Competition of the ASA Section on Statistical Graphics, 1998
- Dissertation Fellowship, Yale University, 1998-1999
- Graduate Fellowship, Yale University, 1994-1998
- Stirling Fellowship, Yale University, 1994-1996
- Donovan and Moody Fellowships for graduate study at Exeter College, Oxford, 1992-1994
- Class of 1960 Scholar in Economics, Williams College, 1992
- Phi Beta Kappa and Sigma Xi, Williams College, 1992

Ph.D. students

- Chandra Erdman. *Dissertation:* Bayesian Change Point Analysis with Applications to Microarray Data, 2009
- Michael Kane. *Dissertation:* Scalable Strategies for Computing with Massive Data, 2010

Patent application

Cone-beam Computed Tomography on a Multimodal Imaging System. With Ming Jiang et.al, Peking University, 2010

Grant

Federal Express R09621, "Trade and Environmental Linkage Study," Principal Investigator ($250,000, 2010-2011)

Select recent publications

- Emerson, John W. and Taylor Arnold (2011). "Statistical Sleuthing by Leveraging Human Nature: A Study of Olympic Figure Skating." Accepted, *The American Statistician.*
- Emerson, John W., Daniel C. Esty, Tanja Srebotnak, and Diana Connett (2011). "Exploring Trade & the Environment: An Empirical Examination of Trade Openness and National Environmental Performance." New Haven: Yale Center for Environmental Law and Policy.
- Emerson, John W., Daniel C. Esty, Mark A. Levy, and William Dornbos (2011). "The Environmental Performance Index." *Berkshire Encyclopedia of Sustainability: Measurements, Indicators, and Research Methods Sustainability*.
- Emerson, John W. and Silas Meredith (2011). "Nationalistic Judging Bias in the 2000 Olympic Diving Competition." *Math Horizons*, 18(3), 8-11.
- Crimin, Kimberly S., John W. Emerson, and Robb J. Muirhead (2010). "Reference regions for beat-to-beat ECG data." *Pharmaceutical Statistics* DOI: 10.1002/pst.444.
- Emerson, John W., D.C. Esty, M.A. Levy, C.H. Kim, V. Mara, A. de Sherbinin, and T. Srebotnjak (2010). *2010 Environmental Performance Index.* New Haven: Yale Center for Environmental Law and Policy.
- Emerson, John W., Miki Seltzer, and David Lin (2009). "Assessing Judging Bias: an Example from the 2000 Olympic Games." *The American Statistician* 63(2), 124-131.
- Emerson, John W., Marisa Dolled-Filhart, Lyndsay Harris, David L. Rimm, and David P. Tuck (2009). "Quantitative Assessment of Tissue Biomarkers and Construction of a Model to Predict Outcome in Breast Cancer Using Multiple Imputation." *Cancer Informatics* 2009:7, 29-40.
- Erdman, Chandra and John W. Emerson (2008). "A fast Bayesian change point analysis for the segmentation of microarray data." *Bioinformatics* 24(19) 2143-2148; doi: 10.1093/bioinformatics/btn404.
- Erdman, Chandra, and John W. Emerson (2007). "bcp: A Bayesian Analysis for Change Point Problems Using R." *Journal of Statistical Software* 23(3).
- Emerson, John W. (2007). "Chance, On and Off the Ice." *Chance* 20(2).
- Levin, Hillel Y. and John W. Emerson (2006). "Biases in the Jury Selection Process: A Study in Connecticut." *Connecticut Law Review* 38(3), 325-353.
- Shachar, Ron and John W. Emerson (2000). "Cast Demographics, Unobserved Segments, and Heterogeneous Switching Costs in a TV Viewing Choice Model." *Journal of Marketing Research* 37(2), 173-186.

Other publications

- "Towards Rational Molecular Design: Derivation of Property Guidelines for Reduced Acute Aquatic Toxicity." Voutchkova, A. M., Emerson, J. W., Osimitz, T., Zimmerman, J. B., and Anastas, P. T., accepted, *Green Chemistry*.
- Klein, A., Evan Beach, John W. Emerson, and Julie Zimmerman (2010). "Accelerated solvent extraction of lignin from Aleurites moluccana (candlenut) nutshells." *Journal of Agricultural and Food Chemistry* 58 (18), 10045-10048, DOI 10.1021/jf1019856.
- Simpson, Andrew K., J. Sabino, P. Whang, John W. Emerson, and Jonathan Grauer (2009). "The Assessment of Cervical Foramina with Oblique Radiographs: the Effect of Film Angle on Foraminal Area." *Journal of Spinal Disorders and Techniques* 22(1), 21-25.

- Bruscia, Emanuela M., Ping-Xia Zhang, Elisa Ferreira, Christina Caputo, John W. Emerson, David Tuck, Diane Krause, and Marie Egan (2008). "Microphages directly contribute to the exaggerated inflammatory response in Cftr -/- mice." *American Journal of Respiratory Cell and Molecular Biology* 40, 295-304.
- Emerson, John W. (2008). "Review of Interactive and Dynamic Graphics for Data Analysis." *Biometrics*, 64 (4), 1301-1302.
- Goldstein, Robin, Johan Almenbert, Anna Dreber, John W. Emerson, Alexis Herschkowitsch, and Jacob Katz (2008). "Do More Expensive Wines Taste Better? Evidence from a Large Sample of Blind Tastings." *Journal of Wine Economics*, 3(1), 2008.
- George, Mathew K., John W. Emerson, Sameer A. Cheema, Robert McGlynn, Bryce Ford, James F. Martone, Milton Bruce Shields, and Martin Wand (2008). "Evaluation of a modified protocol for Selective Laser Trabeculoplasty." *Journal of Glaucoma* 17(3), 197-202.
- Bible, Jesse E., Andrew K. Simpson, John W. Emerson, Debdut Biswas, and Jonathan N. Grauer (2008). "Quantifying the Effects of Age, Gender, Degeneration, and Adjacent Level Degeneration on Cervical Spine Range of Motion Using Multivariate Analyses." *Spine* 33(2), 1793-1799.
- Arya, Adarsh V., John W. Emerson, Michael Engelbert, Curtis Hagedom, and Ron A. Adelman (2006). "Surgical management of Pseudophakic Retinal Detachments." *Opthalmology* 113(10), 1724-1733.
- George, Mathew K, Thomas Kuriakose, Brian M. DeBroff and John W. Emerson (2006). "The effect of Gonioscopy on keratometry and corneal surface topography." *BMC Ophthalmology* 6:26; doi:10.1186/1471-2415-6-26.
- Bertagnoli, Rudolf, James J. Yue, Trace Kershaw, Rahul V. Shah, Frank Pfeiffer, Andrea Fenk-Mayer, Regina Nanieva, Amin Karg, Daniel Husted, and John W. Emerson (2006). "Lumbar total disc arthroplasty utilizing the ProDisc prosthesis in smokers versus nonsmokers: a prospective study with 2-year minimum follow-up." *Spine* 31(9), 992-997.
- Bertagnoli, Rudolf, James J. Yue, Regina Nanieva, Andrea Fenk-Mayer, Daniel Husted, Rahul V. Shah, and John W. Emerson (2006). "Lumbar total disc arthroplasty in patients older than 60 years of age: a prospective study of the ProDisc prosthesis with 2-year minimum follow-up period." *Journal of Neurosurgery: Spine* 4(2).
- Bertagnoli, Rudolf, James J. Yue, Andrea Fenk-Mayer, Jonathan Eerulkar, and John W. Emerson (2006). "Treatment of symptomatic adjacent-segment degeneration after lumbar fusion with total disc arthroplasty by using the prodisc prosthesis: a prospective study with 2-year minimum follow-up." *Journal of Neurosurgery: Spine* 4(2).
- Emerson, John W. and Walton Green (2006). "The Spark Mat: a graphical method of exploring spatially distributed time series." JASA poster competition, 2006.
- Emerson, John W. (1998). "Mosaic Displays in S-PLUS: A General Implementation and a Case Study." *Statistical Computing and Graphics Newsletter* 9(1), 17-23.
- "Mosaic Displays in S-PLUS." Winner, 1998 Student Paper Competition of the ASA Section on Statistical Graphics.

Work in progress

- "Scalable Strategies for Computing with Massive Data." Michael J. Kane, John W. Emerson, and Steve Weston, in submission, *Journal of Computational and Graphical Statistics.*

- "Goodness-of-Fit Tests for Discontinuous Distributions." Taylor B. Arnold and John W. Emerson, in submission, *The R Journal.*
- "Evaluation of steady-state visual evoked potentials in diabetic retinopathy." Mathew George, George Hu, John W. Emerson, Veena Rao, Ron Adelman, James Tsai, and John Huang, in submission.
- "A Comprehensive Rehabilitation Protocol for Lumbar Disc Replacement: The Straubing Rehabilitation Index." With James Yue *et.al,* in submission.
- "Cone-beam reconstruction for the 2-circles plus line trajectory." Yanbin Lu, Jiansheng Yang, John W. Emerson, Heng Mao, Yuanzheng Si, and Ming Jiang, in submission.
- "The generalized pairs plot." John W. Emerson, Walton A. Green, Barret Schloerke, Di Cook, Heike Hofmann, and Hadley Wickham, in submission.
- "Minimalist Sufficient Statistics." In preparation.
- "Distance-based ANOVA with functional magnetic resonance imaging." With Zarrar Shehzad, in preparation.
- "Multivariate Bayesian Change Point Analysis." With Chandra Erdman, in progress.

Committees and service

- Associate Editor, *Journal of Statistical Software*, 2010-present
- Secretary/Treasurer, American Statistical Association Section on Statistical Graphics, 2010-present
- Program Chair Elect, American Statistical Association Section on Statistical Computing, 2011
- Volunteer Instructor, Yale Center for Analytical Sciences Summer Young Scholars Program (teaching New Haven area high school students), 2010-present
- Quan 199 Guest Lecture Module: Quantitative Methods across the Disciplines
- MCDB 900 Guest Lecture: Statistics and Statistics at Yale
- Department of Statistics Statistical Consulting Clinic Faculty Lead, 2003-present
- Quantitative Reasoning Committee, 2008-present
- Bates Fellowship Committee 2007-present
- Secretary of the Fellows, Jonathan Edwards College, Yale University, 2007-present
- Jonathan Edwards Senior Essay Awards Committee, 2007-2008, 2011
- Department of Statistics Director of Graduate Studies, 2006-2007
- Department of Statistics Computer Committee, Chair, 2003-2009
- Yale University Teaching & Learning Committee, 2005-2007
- Yale University Teaching & Learning Portal Advisory Committee, 2005-2007
- Library Hiring Committee, Social Science Data Librarian, 2005-2006
- Department of Statistics Hiring Committee, 2004-2006
- Department of Statistics Library Coordinator, 2003-2006
- Bioinformatics Admissions Committee, 2003-2005
- Reviewer, *Theoretical Biology and Medical Modeling*
- Reviewer, *Bioinformatics*
- Reviewer, *The R Journal*
- Reviewer, *The American Statistician*
- Reviewer, *Journal of Industrial Ecology*
- Reviewer, *Computational Statistics*

- Reviewer, *International Journal of Biostatistics*
- Reviewer, *Journal of Marketing*
- Board of Directors, Yale Co-op. 1995-1999
- Graduate Coordinator, Hall of Graduate Studies, Yale University, 1998-1999
- Dean's Committees, Yale Graduate School, 1995-1997: McDougal Committee (designing the new Graduate Student Center); Kutzinski Committee (reviewing the Teaching Fellow Program).

Teaching

- "An Intensive Course in Data Analysis with Statistical Case Studies," Peking University, summers 2007 and 2009
- Stat 662, Statistical Computing. New course, 2008, 2011
- Stat 661, Advanced Data Analysis
- Stat 627, Statistical Consulting. New course, 2003-present
- Stat 626, Practical Work
- Stat 625, Case Studies
- Stat 242/542, Theory of Statistics
- Stat 230/530, Introductory Data Analysis
- Stat 128, Real-World Statistics. New course, 2009
- Stat 100, Introductory Statistics

Conferences and presentations

- Organizer, Invited Session, Joint Statistical Meetings, Miami, FL 2011: "The Future of Statistical Computing Environments"
- Organizer and Speaker, Topic Contributed Session, Joint Statistical Meetings, Miami, FL 2011: "Statistical Analyses of Judging in Athletic Competitions"
- Invited Speaker, Statistical Methods for Very Large Data Sets Conference, Baltimore, MD, June 2011
- R in Finance, Chicago, IL, April 2011
- Invited Speaker, American Association for the Advancement of Science, Washington, DC, February 2011
- Presenter for expert review panel, Yale Center for Environmental Law and Policy, Washington, DC, November 2010
- Presenter, Chicago R-Meetup, October 2010
- Seminar series speaker, Yale Center for Analytical Sciences, October 2010
- Invited Speaker, seminar/workshop series, UNICAMP, Campinas, Brazil, September 2010
- Invited Speaker, seminar/workshop series, IME-USP, Sao Paulo, Brazil, September 2010
- Invited Speaker, IM-UFRJ, Rio do Janeiro, Brazil, September 2010
- Invited Speaker, Boston University, September 2010
- Invited Speaker, Census Bureau, Washington, DC, July 2010
- Presenter, UseR! Topic Contributed Session, Washington, D.C., July 2010
- Presenter, Interface 2010 Topic Contributed Session, Seattle, June 2010

- Presenter, New York City R-Meetup, June 2010, available online at http://www.vcasmo.com/video/drewconway/8555
- Invited Neyman Seminar Speaker, University of California, Berkeley, April 2010
- Invited Speaker, Choate Rosemary Hall, April 2010
- Invited Speaker, Bristol-Myers Squibb, April 2010
- Department Seminar Speaker, Yale University, March 2010
- Invited Seminar Speaker, Middlebury College, March 2010
- Presenter, New York City R-Meetup, March 2010
- Presenter of the 2010 Environmental Performance Index (press conference), World Economic Forum, Davos, Switzerland, January 2010
- "60 Years: Antony Unwin. Mathematisches Festkolloquium." Augsburg, Germany, January 2010
- Department Seminar Speaker, Biostatistics, Yale University, September 2009
- Workshop Instructor, "A hands-on introduction to R." Annual Meeting of the American College of Clinical Pharmacology, San Antonio, TX, September 2009
- Presenter, UseR! Focus Session, Rennes, France, July 2009
- Speaker, Rmetrics Workshop, Meielisalp, Switzerland, July 2009
- Speaker, Chinese Environmental Performance Index Workshop, Hong Kong, June 2009
- Invited Seminar Speaker, Department of Statistics and Mathematics, Vienna University of Economics and Business Administration, Vienna, Austria, June 2009
- Invited Seminar Speaker, New York City Chapter of the American Statistical Association, May, 2009
- Workshop Instructor, "R and Statistics." West Point, New York, May 2009
- Workshop Instructor, "What's the Fuss about R?" New York City, April, May, November 2009
- Invited Seminar Speaker, Iowa State University, December 2008
- Organizer and Presenter, Joint Statistical Meetings, Denver, CO, August 2008
- Presenter, UseR! Focus Session, Dortmund, Germany, August 2008
- Workshop Instructor, Pfizer, Inc., Groton, CT, July 2008
- Invited Seminar Speaker, Boston University, May 2008
- Workshop Instructor, Novartis. Boston, MA, May 2008
- Invited Seminar Speaker and Guest Lecturer, Universität Augsburg, Germany, April 2008
- Invited Seminar Speaker, Colby College, April 2008
- Invited Seminar Speaker, Middlebury College, 2007 and 2005
- Invited Panel Participant and Session Chair, Joint Statistical Meetings, Salt Lake City, Utah, 2007
- Joint Statistical Meetings Poster Presenter: "The Sparkmat: A Graphics Method of Exploring Spatially Distributed Time Series," Seattle, WA, 2006
- Invited Lecture, Pfizer, Inc., Groton, CT, 2006
- Keynote Speaker, Annual Meeting of the Connecticut Independent Teachers Association, Greenwich, CT, 2006.
- Speaker, UseR! Session: "Sparkmats and Generalized Pairs Plots," Vienna, Austria, 2006
- Joint Statistical Meetings, Minneapolis, MN, 2005
- Joint Statistical Meetings, Toronto, Canada, 2004

R packages

- Bayesian change point analysis (with R and C/C++), with Chandra Erdman: package **bcp** (http://cran.r-project.org/web/packages/bcp/).
- A collection of tools for exploratory data analysis, with Walton Green: package **YaleToolkit** (http://cran.r-project.org/web/packages/YaleToolkit/), packages **barcode** (http://cran.r-project.org/web/packages/barcode/) and **gpairs** (http://cran.r-project.org/web/packages/gpairs/).
- The Bigmemory Project (http://www.bigmemory.org/), with Michael Kane: a family of R packages (with R and C++) for scalable solutions to computing with massive data (**bigalgebra** and **bigvideo** available on R-Forge, others available on CRAN):
  - **bigmemory**
  - **biganalytics**
  - **bigtabulate**
  - **synchronicity**
  - **bigalgebra**
  - **bigvideo**
- Analysis of electrocardiogram QT/TQ data, with Robb Muirhead: package **QTTQ** (used internally at Pfizer, Inc.).
- Methods for one-sample goodness-of-fit tests with discrete distributions, with Taylor Arnold: packages **ks.test** and **cvm.test** (https://r-forge.r-project.org/R/?group_id=802).
- Methods for fMRI processing and analysis, with Zarrar Shehzad: packages **niftir**, **connectir**, and **bigextensions** (https://r-forge.r-project.org/R/?group_id=966).

In popular media

- *Wall Street Journal* (10/25/2010). "In Medicaid's Data Trove, Clues to Curing Cost Crisis," contributor to methodology and analysis.
- *Newsday.com* (2/22/2010). "Statistics say skating judges still don't have it right," quoted.
- *The Statesman* (1/28/2010). "The 2010 Environmental Performance Index," quoted.
- *Wall Street Journal* (6/20/2008). "World's Greatest Athlete: How we did it," contributor to methodology and analysis.
- *Wall Street Journal* (4/24/2008). "Ford's Googol Googol Console Options," cited.
- *Wall Street Journal* (3/31/2007). "Reservations at Top Restaurants," conducted analysis and produced the graphics.
- Emerson, John W. (2006). "Stat prof assesses the Game day data," *Yale Daily News Yale-Harvard Issue*, November 17.
- *Wall Street Journal* (3/18/2006). "The Perfect Payday." Collaborated with eventual Pulitzer Prize winning authors Charles Forelle and James Bandler on the methodology (and implemented the R code for the analysis) leading to the infamous stock option backdating scandal.
- *Wall Street Journal* (3/23/2006). "Picking the Perfect NCAA Bracket," citation.
- *ABC World News Tonight* (2/17/2006). "John Berman on the new scoring system in skating." Interviewed, following my analysis of a problem with the new scoring system leading up to the Winter Olympics.
- *Wall Street Journal* (2/15/2006). "Does the New Way of Judging Skating Actually Add Up?" Author Charles Forelle wrote the first piece on my analysis, with subsequent coverage appearing in the *Scientific American* and many online news outlets.
- Emerson, John W. "On the Court, In the Office: A Game of Statistics," *Hartford Courant*, March 17, 2005.