Andrew Paul Bray

Assistant Professor of Statistics
Reed College
Department of Mathematics
3203 SE Woodstock Blvd
Portland, OR 97202

Phone: (760) 519-5979
Email: abray@reed.edu
GitHub: andrewpbray

Education

Ph.D. Statistics, University of California, Los Angeles, 2013.

Research Area: Spatial-temporal models, point processes, model diagnostics

Advisor: Frederic Paik Schoenberg

M.S. Statistics, University of California, Los Angeles, 2010.

B.S. Earth & Ocean Sciences and History, minor in Political Science, Duke University, 2005.

Publications

(* indicates an undergraduate student)

Manuscripts in Preparation

Baumer B., Bray A., Çetinkaya-Rundle M., and Hardin J. *Teaching Statistics with DataCamp*. Submitted to the Journal of Statistics Education June 26, 2020.

Bray A. *Tidy statistical inference with infer*. To be submitted to the Journal of Statistical Software.

Refereed Articles

Till A.*, Rypel A., Bray A. & Fey S. (2019) *Thermal Extremes Drive Fish Die-offs in North Temperate Lakes*. Nature Climate Change, (currently in proofing stage) DOI: 10.1038/s41558-019-0520-y

Couch S.*, Kazan Z.*, Shi K.*, Bray A., & Groce A (2019). *A Differentially Private Wilcoxon Signed-Rank Test*. 26th ACM Conference on Computer and Communications Security. London, United Kingdom.

Swanberg M.*, Globus-Harris I.*, Griffith I.*, Groce A., & Bray A. (2019). *Improved Differentially Private Analysis of Variance*. Proceedings of 19th Privacy Enhancing Technologies Symposium. Stockholm, Sweden.

Campbell Z.*, Bray A., Ritz A., & Groce A. (2017). *Differentially Private ANOVA Testing*. First International Conference on Data Intelligence and Security. South Padre Island, TX.

Deveaux D., Bray A., et al. (2017). *Curriculum Guidelines for Undergraduate Programs in Data Science*. Annual Review of Statistics and Its Applications, 4:2.1-2.16.

Lavine M., Bray A., & Hodges J. (2015). *Approximately Exact Calculations for Linear Mixed Models*. Electronic Journal of Statistics, 9(2), 2293-2323.

Bray A., Wong K., Barr C., & Schoenberg F.P. (2014). *Using the Voronoi tessellation to calculate residuals for spatial point process models*. Annals of Applied Statistics.

Baumer B., Çetinkaya-Rundel M., Bray A., Loi L., & Horton N. (2014). *R Markdown: Integrating a reproducible analysis tool into introductory statistics*. Technology Innovations in Statistics Education, 8(1).

Bray A., Schoenberg F.P. (2013). *Assessing point process models for earthquake forecasting*. Statistical Science, 28(4), 510-520.

Nichols K., Schoenberg F.P., Keeley J.E., Bray A., & Diez D. (2011). *The application of prototype point processes for the summary and description of California wildfires*. Journal of Time Series Analysis, 32, 420-429.

Bray A.P., Mullin P., Schoenberg F.P., Mac Gibbon K., Romero R., Goodwin T.M., & Fejzo M.S. (2011). *Prenatal exposure to hyperemesis gravidarum linked to increased risk of psychological/behavioral disorders in adulthood*. Journal of Developmental Origins of Health and Disease, 2(4), 200-204.

Khan H.D., Rosati J.A., & Bray A.P. (2011). *Statistical evaluation of data from multi-laboratory testing of a measurement method intended to indicate the presence of dust resulting from the collapse of the world trade center*. Environmental Monitoring and Assessment.

Other Publications

Bray A. (2019). *Inference for Categorical Data*. DataCamp online course, component of *Intro to Statistics with R* track.

Bray A. (2017). *Exploratory Data Analysis*. DataCamp online course, component of *Intro to Statistics with R* track.

Bray A. (2018). *Review: Modern Data Science with R*. The American Statistician: Reviews, 72(3), 295-299.

Bray A. (2014). *A festival of data: student perspectives*, President's Corner, Amstat News, September Issue.

Bray A., Çetinkaya-Rundel M., & Stangl D. (2014). *Five concrete reasons your students should be learning to analyze data in the reproducible paradigm*. Taking a Chance in the Classroom, Chance Magazine, September Issue.

Gould R., Baumer B., Çetinkaya-Rundel M., and Bray A. (2014). *Big data goes to college*, Amstat News, June issue.

Bray, A., Baumer, B. (2014). *MM Data Labs Data Science Development Program Prospectus*. MassMutual Life Insurance Company.

Bray, A. (2012). *Power analysis for residual testing of spatial point process models based on a fine regular grid*. UCLA Statistics Preprint Series.

Bray, A.P. (2010). *Investigation of the effect of conductivity on the number of mayfly genera: Poisson regression and parameter stability*. Office of Research and Development, US EPA.

Bray, A.P. (2010). *Spatial correlation of residuals from a regression model of mayfly genera in West Virginia streams*. Office of Research and Development, US EPA.

Bray, A.P. (2010). *Analysis of oil dispersant application rates and VOC levels*. Office of Research and Development, US EPA.

Presentations

Workshops

Bray A., Couch S.*, Groce A., Kazan Z.*, and Shi K.* *Private Hypothesis Testing Using Rank-based Methods*. Theory and Practice of Differential Privacy Workshop, Toronto, Canada, October 2018.

Conferences, invited

Bray A. *Salt, Fat, Acid, Heat: An alternative to cookbook statistics*, Statistics and Data Science Symposium, Bellevue, WA, June 2019.

Bray A., Ismay C. *Aligning Inference with the Tidyverse*, Joint Statistics Meetings, Vancouver, BC, August 2018.

Bray A., Ismay C. *infer: A package for tidy statistical inference*, Electronic Conference on Teaching Statistics, May 2018.

Bray A. *infer: A package for tidy statistical inference*, rstudio::conf, San Diego, CA, January 2018.

Lavine M., Bray A., Hodges J. *Approximately Exact Calculations for Linear Mixed Models*. IMS New Researchers Conference, Seattle, WA, August 2015.

Bray A., Gould R. *Developing Data Science Curricula*. US Conference on Teaching Statistics, State College, PA, May 2015.

Bray A.P. *Putting the Data in DataFest*. Panel discussion, Joint Statistics Meetings, Boston, August 2014.

Bray A. *R Markdown: Bringing Reproducible Analysis into the Classroom*. Panel discussion, New England Statistics Symposium, Boston, April 2014.

Bray A., Wong K., Schoenberg F.P., & Barr C. *Using the Voronoi tessellation to calculate residuals for spatial point process models*. Young Researchers Presentation at Southern California American Statistical Association Annual Meeting, November 2012.

Conferences, contributed

Bray A. Çetinkaya-Rundel M. *How to Integrate Open-Access and Open-Source Educational Materials*, Joint Statistics Meetings, Chicago, IL, August 2016.

Bray A., Ismay C. *Lowering the barriers to inclusive, collaborative, reproducible analyses*, Electronic Conference on Teaching Statistics, May 2016.

Lavine M., Bray A., Hodges J. *Approximately Exact Calculations for Linear Mixed Models*. UseR! Conference, Aalborg, Denmark, June 2015.

Çetinkaya-Rundel M., Bray A. *Planting seeds of reproducibility in the introductory statistics course with R markdown*, Electronic Conference on Teaching Statistics, May 2014.

Çetinkaya-Rundel M., Bray A. *Teaching data analysis in R through the lens of reproducibility*, useR! Conference, June 2014.

Bray A., Wong K., Schoenberg F.P., & Barr C. *Residuals for spatial point processes based on Voronoi tessellations*. Joint Statistical Meetings, San Diego, July 2012.

Çetinkaya M., Bray A.P. *Integrating R into introductory statistics*. Contributed paper to the Joint Statistical Meeting, San Diego, July 2012.

Potter R., Musgrave D., Bray A., D'Aoust T., & Weingartner T. *Using HF radar to map surface currents in the beaufort sea*. Alaska Marine Sci. Symposium, Anchorage, AK, Jan 2007.

Potter R., Bray A., D'Aoust T., Musgrave D., & Weingartner T. *An application of HF radar in the coastal Beaufort Sea*. 7th International Conference on Global Climate Change: Connections to the Arctic, Fairbanks, AK, Feb 2007.

Invited Seminars

Differentially Private Hypothesis Testing. Statistics Department Seminar Series. Portland State University, April 2019.

Statistics under Differential Privacy. Statistics Department Seminar Series. Oregon State University, November 2018.

infer: A package for tidy statistical inference. Oregon Chapter of the ASA Spring Meeting. Beaverton, OR, May 2018.

infer: A package for tidy statistical inference. Mathematics-Statistics Seminar Series. Washington State University, Vancouver, January 2018.

Alternative Institutional and Educational Mechanisms for Data Science Education. National Academies of Sciences, Engineering, and MedicineâĂŹs Roundtable on Data Science Education. Northwestern University, October 2017.

Tectonics and Tessellations. Mathematics Seminar Series. University of Portland, November 2015.

Tectonics and Tessellations. Seminar Series. USGS Cascade Volcano Observatory, November 2015.

Approximately Exact Calculations for Linear Mixed Models. Mathematics Seminar Series. Reed College, April 2015.

Approximately Exact Calculations for Linear Mixed Models. Statistics Seminar Series. University of Connecticut, April 2015.

Bray A., Reich N. *Update on the cutting edge of R and RStudio*. Western Mass Data Science, Stats, and R Meetup Group, April 2014.

Bray A., Reich N. *Open resources for teaching statistics*. Statistics Seminar Series, University of Massachusetts, Amherst, April 2014.

Smith M., Bray A., & Johnston C. *Statistical sample design for coalbed methane industry survery: projects vs wells*. Presentation to the Federal Committee on Stats. Methodology, Washington, DC, November 2009.

Funding and Awards

External Funding

Submitted 3/26/19: *S-Stem Track II: Building the STEM Workforce through Networks at a Small Liberal Arts College*. NSF Division of Undergraduate Education. Co-PI with PI Suzy Renn and Co-PI Kjersten Whittington. 2020 - 2024, #193044 ($999,998).

RUI: Differentially Private Hypothesis Testing. NSF Directorate for Computer & Information Science & Engineering. Co-PI with PI Adam Groce and Co-PI Anna Ritz. June 2018 - June 2021, #1817245 ($344,684).

SBIR: Pre-clinical Evaluation of a Novel Immune Modulator, Alpha-TEA in Combination with Immune Checkpoint Blockade. NIH National Cancer Institute. Consulting statistician with PI Emmanuel Akporiaye. September 2018 - August 2019, #R43CA228721 ($299,607).

Differentially Private Hypothesis Testing. Facebook Research: Secure the Internet Grants. Co-PI with Adam Groce and Anna Ritz. (Submitted June 2018, not funded).

Signature Verification Statistical Sampling Advisor: A proposal with Albyn Jones to evaluate methods for the statistical sampling of signatures for Oregon's initiative, referendum, recall, minor party formation and candidate nominating processes (State of Oregon RFP 165-1208-16: submitted Spring 2017, not funded).

Awards

Collegium of University Teaching Fellow, competitive university-wide fellowship through UCLA Office of Instructional Development (2012).

Teaching Assistant Consultant, competitive position funded through UCLA Office of Instructional Development (2011-2012).

Teaching Assistant of the Year, UCLA Department of Statistics (2011).

Academic and Professional Experience

Expert Witness (Statistician), US Federal Court, Portland, OR May 2018

Provided expert testimony in a case related to the occupation of the Malheur National Wildlife refuge. Assessed the design and analysis in a 2008 ballistics study that established methodology used by an expert witness for the prosecution.

NSF Postdoctoral Fellow, Five Colleges Consortium Fall 2013 - Summer 2015

Conducted original research under the supervision of Michael Lavine at the University of Massachusetts while teaching one class per semester at one of the four nearby liberal arts colleges. At the University of Massachusetts my research focused on point process data in ecological applications and algorithms for fitting statistical models.

Board of Advisors, MassMutual Data Labs Summer 2014 - Fall 2015

Provide academic oversight and advising for the students in the data science development program at MassMutual in Amherst.

Graduate Research Assistant, UtopiaCompression Corporation Summer 2011 - Summer 2013

Worked with a team of engineers to improve range detection of unmanned aerial vehicles using passive sensing. Improved algorithm performance by residual analysis and modeling.

Statistician Intern, Office of Research and Development, US EPA Summer 2010

Worked on several projects in the Quantitative Risk Methods Group including assessing the effect of oil dispersants in the Gulf of Mexico, modeling child soil ingestion, and classifying dust from the World Trade Center collapse.

Statistician Intern, Office of Water, US EPA Summer 2009

Collaborated with a team of engineers, economists, and statisticians to set water pollution guidelines. Created the sample design used to collect information related to the pollutants generated by the coalbed methane industry.

Research Technician, SALMON Project, University of Alaska Feb. 2006 - Jul. 2007

Set-up and maintained 6 High Frequency Radars as part of a research project to measure Alaskan ocean surface currents for oil spill risk assessments with implications for arctic ice. Developed a software suite to analyze oceanic and meteorological data using MATLAB.

Teaching

Reed College, as Assistant Professor

Mathematical Statistics, Spring 2018

Probability, Fall 2016

Statistical Learning, Spring 2016, Fall 2017

Data Science, Spring 2017, 2018

Introduction to Probability and Statistics, Fall 2015, Spring 2016, Fall 2016, Fall 2017

Mount Holyoke College, as Visiting Assistant Professor

Applied Regression Methods, Fall 2014

Smith College, as Visiting Assistant Professor

Introduction to Probability and Statistics, Spring 2014

Introduction to the Practice of Statistics, co-taught with Katherine Halvorsen, Fall 2013

UCLA Department of Statistics, as primary instructor

Burden of Proof: Data and Scientific Reasoning, undergraduate honors seminar, Spring 2013

Statistical Content Helping to Empower Mathematicians at Two-Year Colleges, with Robert Gould, Summer 2012

Teaching College Statistics, pedagogy course for new graduate teaching assistants, Winter 2012

Mentoring

Senior Theses

Theodore Dounias (Math/PoliSci 2018): *Turnout and mail voting in Colorado, or, How I learned to stop worrying and love voter registration files*

Boyuan Li (Math/Econ 2018): *How do natural disasters affect household saving decisions?: evidence from rural china*

Seerat Jhajj (Math/Econ 2018): *Determinants and policy implications of bilateral FDI flows: evidence from emerging economics in 2001-2012*

Emily Palmer (Math/Stat 2018): *A musical stylometry study on works by Fanny Hansel and Felix Mendelssohn*

Emerson Webb (Math/Stat 2018): *Added variable plot importance and joint added plot importance measures for random forests*

Aurora Owens (Math/Stat 2017): *INFFOREST variable importance for random forests*

Sam Olson (Math/Econ 2017): *A Big Data Analysis of Pokemon Battling*

Alicia Toshima (Math/Econ 2017): *Effect of Financial Aid Received on Undergraduate GPA*

Ian Morrison (Math/Econ 2017): *The Use of Remittances in the Household Economy*

Helena Pedrotti (Math/Econ 2017): *Prenuptial Agreements and Income Inequality*

Carolyn Cole (Math/Econ 2017): *Straight to the Source: The Economics of Human Trafficking*

Will Jones (Math/Stat 2016): *Hierarchical Models for Crowdsourced Bicycle Route Ratings*

Dean Young (Math/Econ 2016): *Myopic Loss Aversion in Investment Behavior*

Philip Stallworth (Math/Stat 2016): *A Cluster Model for k-tree Sampling of Spatial Point Process data*

Michael Weiss (Bio 2016, with Suzy Renn): *Non-kinship Social Bonds in Resident Killer Whales*

Ana Moreno-Mesa (Smith College 2014): *Investigation and visualization of trends in Smith College statistics enrollment*

Mariah Mullens (Mount Holyoke College 2014): *Exploring sensitivity of linear regression assumptions through simulation*

Student Research Awards

1st place, Undergraduate Statistics Research Project Competition (Fall 2018). Couch S., Kazan Z., and Shi K. *A Differentially Private Wilcoxon Signed-Rank Test*.

1st place, Undergraduate Statistics Research Project Competition (Spring 2016). Jones, W. *Multilevel Models and Missing Data Models for Crowd-sourced Bicycle Route Ratings*.

Honorable Mention, Undergraduate Statistics Research Project Competition (Spring 2016). Young, D. *Myopic Loss Aversion in Investment Behavior*.

Class of '21 Award (2016), Weiss M. The only academic award given by Reed College each year for "creative work of notable character, involving an unusual degree of initiative and spontaneity" for his thesis, *Non-kinship Social Bonds in Resident Killer Whales*.

Student-published software

Couch S. and Rosenberg K. (2018) *gbfs*. R package with tools for accessing and wrangling Generalized Bikeshare Feed Specification data. Authored by Bray's Math 243 students.

Lee J. and Yancheff M. (2017) *rcv*. R package with tools for wrangling and visualizing data from ranked choice voting elections. Authored by Bray's summer research students.

Statistics Education Work

OpenIntro

Co-authored a book of ten labs to accompany *OpenIntro Statistics, Second Edition* (Diez et al. 2012). Students gain experience and insight into concepts in statistics by analyzing real data in the R computing language. Available free and open-source at http://www.openintro.org/stat/labs and on github. The labs have been used at many institutions including Duke, Mt. Holyoke College, University of Chicago, Oregon State University, Vanderbilt, and Reed College.

Five College DataFest

Organized (with Ben Baumer of Smith College) the first annual Five College DataFest: an undergraduate data analysis competition in which teams of up to 5 students work for 48 hours to extract insight from a rich and complex data set. The event draws in undergraduates from across the Five Colleges, with 65 students attending the 2014 event. This year's event drew 200 students.

MassMutual Data Science Development Program

Designed (with Ben Baumer of Smith College) the prospectus for a new program at MassMutual to recruit undergraduates and train them over three years in the methods and principles of data science. The program's first cohort of seven students began summer 2014 and the program is expected to grow to about 30 students.

Active Learning Workshop

Planned and facilitated a workshop to help graduate teaching assistants incorporate active learning strategies into the classroom. The workshop was held for students from across campus as part of the UCLA TA Conference and again for students from the Department of Statistics. Funded by the UCLA Office of Instructional Development (Fall 2012).

Scientific Software - R Packages

Bray A., Ismay C., Baumer B., and Çetinkaya-Rundel, M. *infer*. Conducts randomization- and simulation-based inference procedures with expressive syntax. Forms a component of the *tidymodels* package that is maintained by RStudio (2018).

Lavine M. and Bray A. *lmmoptim*. Implements the WHIM algorithm described in Lavine, Bray, and Hodges (2015).

Bray A.P., Diez D., and Çetinkaya-Rundel M. *oidata*. Provides data sets from several sources that may be useful for teaching, practice, or other purposes. Functions have also been included to assist in the retrieval of table data from websites or in visualizing sample data.

Bray A.P. and Çetinkaya, M. *oilabs*. Provides data sets and functions to facilitate statistics education through computing.

Bray A.P. *uav*. A suite of functions to visualize and analyze navigation data from unmanned aerial vehicles.

Service

Service to the profession

Associate Reviews Editor, Journal of the American Statistical Association (2015 - present).

Associate Editor, Journal of Statistics Education (2015 - 2018).

NSF Reviewer, Division of Undergraduate Education (DUE) and Harnessing the Data Revolution (HDR).

Member, program review committee, US Conference on Teaching Statistics (2019).

Member, ASA/MAA Joint Committee on Statistics Education (2016 - present).

Member, ASA Finance Committee (2016 - present).

Member, ASA DataFest national organizing committee (2014 - 2018).

Contributor, OpenIntro Project (2010 - present).

Reviewer, Journal of Statistical Education (2013 - present) and Technological Innovations in Statistics Education (2014 - present), CRC Press (two books).

Co-organizer, Western Mass Data Science, Stats, and R meetup group (2013 - 2015).

Secretary/Treasurer, Five Colleges Statistics Program (2014-2015)

Service to Reed College

De-facto Chair: Statistics program. Coordinate searches, staffing, and scheduling; assess transfer credits; pursue long-term planning (secured a second TT line) (2015 - 2018).

Committees: Off-campus Study Committee, Radiation Safety Committee.

Search Committees: statistics tenure-track, statistics visitor (twice).

Ad-hoc Group on Distribution Requirements: along with four other junior faculty members, organized a revision of the Reed College general education requirements to emphasize both breadth and depth of study.

Institutional Grants: worked with STEM faculty and adminstrators to prepare proposals to HHMI (2017) and the Keck Foundation (2018).

Organizer: R Film Series weekly screenings of RStudio::conf talks for students, staff, and faculty from across the college (spring 2019). Replicated at several institutions across the country.

Co-organizer: Pacific Northwest eCOTS 2016 one day conference at Reed for statistics educators in the region.

Organizer: visit by Dr. Andrew Nobel, Dept. of Biostatistics, UNC Chapel Hill including meeting with students, research seminar, and graduate school panel (fall 2018).

Skills

Programming Languages

R, MATLAB, markdown, LaTeX, git, Python, SQL, SAS, Unix shell, C++

Languages

English (native), Spanish (intermediate), Norwegian (basic).

Membership

Fellow, OpenIntro, a non-profit supporting free and open-source education, (2011 - present).

Member, American Statistical Association, Oregon Chapter.

Member, section on Statistics and the Environment; section on Statistical Education section; section on Statistical Learning and Data Science; Joint Statistical Computing and Graphics Section.

Member, American Mathematical Society.

References

Frederic Paik Schoenberg
Professor
Department of Statistics
University of California, Los Angeles
frederic@stat.ucla.edu

Robert Gould
Undergraduate Vice-Chair
Department of Statistics
University of California, Los Angeles
rgould@stat.ucla.edu

Michael Lavine
Professor
Department of Mathematics and Statistics
University of Massachusetts, Amherst
lavine@math.umass.edu

Katherine Halvorsen
Professor
Department of Mathematics and Statistics
Smith College
khalvors@smith.edu

Mine Çetinkaya-Rundel
Associate Professor of the Practice
Department of Statistical Science
Duke University
mine@stat.duke.edu