Andrew M. Raim

Baltimore, MD, U.S.A.
andrew.raim@gmail.com
http://andrewraim.github.io

Last revised: 2023-08-10 07:03

Education

- Ph.D. Statistics, UMBC[†], Spring 2014.
  Dissertation: *Computational Methods for Finite Mixtures using Approximate Information and Regression Linked to the Mixture Mean.*
  Advisor: Nagaraj K. Neerchal.
- M.S. Statistics, UMBC, Fall 2011.
- M.S. Computer Science, UMBC, Fall 2007.
- B.S. Computer Science, UMBC, Spring 2002.

Work Experience

- June 2014 – Present. Research Mathematical Statistician, Center for Statistical Research and Methodology, U.S. Census Bureau.
- September 2008 – May 2014. Graduate Assistant, UMBC Dept of Mathematics and Statistics.
- September 2002 – August 2008. Software Engineer, Advertising.com.
- February 1999 – November 2001. Helpdesk Consultant, UMBC Office of Information Technology.

Graduate Assistantships

- Graduate assistant to Dr. Matthias K. Gobbert at High Performance Computing Facility, Fall 2009 – Spring 2014. Acitivities included developing center web site, participation in *Interdisciplinary Program in High Performance Computing* REU in summers 2010–2013, coauthoring technical reports demonstrating cluster performance and statistical computing, and providing support to user community.
- Instructor for STAT 350 *Statistics with Applications in the Biological Sciences* (Winter 2013, Hybrid Course). Teaching Assistant for STAT 350 *Statistics with Applications in the Biological Sciences* (Spring 2009) and STAT 351 *Applied Statistics for Business and Economics* (Fall 2008).

Journal Articles

[1] Andrew M. Raim, Thomas Mathew, Kimberly F. Sellers, Renee Ellis, and Mikelyn Meyers. Design and sample size determination for experiments on nonresponse followup using a sequential regression model. *Journal of Official Statistics*, 39(2):173–202, 2023.

[2] Andrew M. Raim. Direct sampling with a step function. *Statistics and Computing*, 33(1), 2023.

[3] Andrew M. Raim, Elizabeth Nichols, and Thomas Mathew. A statistical comparison of call volume uniformity due to mailing strategy. *Journal of Official Statistics*, 39(1):103–121, 2023.

[4] Ryan Janicki, Andrew M. Raim, Scott H. Holan, and Jerry J. Maples. Bayesian nonparametric multivariate spatial mixture mixed effects models with application to American Community Survey special tabulations. *The Annals of Applied Statistics*, 16(1):144–168, 2022.

[5] Andrew M. Raim, Scott H. Holan, Jonathan R. Bradley, and Christopher K. Wikle. Spatio-temporal change of support modeling with R. *Computational Statistics*, 36(1):749–780, 2021.

[6] Darcy Steeg Morris, Andrew M. Raim, and Kimberly F. Sellers. A Conway-Maxwell-multinomial distribution for flexible modeling of clustered categorical data. *Journal of Multivariate Analysis*, 179:104651, 2020.

[7] Sean Martin, Andrew Raim, Wen Huang, and Kofi Adragni. ManifoldOptim: An R interface to the ROPTLIB library for Riemannian manifold optimization. *Journal of Statistical Software*, 93(1):1–32, 2020.

[8] Sai K. Popuri, Andrew M. Raim, Nagaraj K. Neerchal, and Matthias K. Gobbert. Parallelizing computation of expected values in recombinant binomial trees. *Journal of Statistical Computation and Simulation*, 88(4):657–674, 2018.

[†]UMBC: University of Maryland, Baltimore County.

[9] Andrew M. Raim, Nagaraj K. Neerchal, and Jorge G. Morel. An extension of generalized linear models to finite mixture outcome distributions. *Journal of Computational and Graphical Statistics*, 27(3):587–601, 2018.

[10] Derek S. Young, Andrew M. Raim, and Nancy R. Johnson. Zero-inflated modelling for characterizing coverage errors of extracts from the US Census Bureau's Master Address File. *Journal of the Royal Statistical Society: Series A*, 180(1):73–97, 2017.

[11] Andrew M. Raim, Nagaraj K. Neerchal, and Jorge G. Morel. An approximation to the information matrix of exponential family finite mixtures. *Annals of the Institute of Statistical Mathematics*, 69(2):333–364, 2017.

[12] Kimberly F. Sellers and Andrew Raim. A flexible zero-inflated model to address data dispersion. *Computational Statistics and Data Analysis*, 99:68–80, 2016.

[13] Kofi P. Adragni, Elias Al-Najjar, Sean Martin, Sai K. Popuri, and Andrew M. Raim. Group-wise sufficient dimension reduction with principal fitted components. *Computational Statistics*, 31(3):923–941, 2016.

[14] Kofi Placid Adragni and Andrew M. Raim. ldr: An R software package for likelihood-based sufficient dimension reduction. *Journal of Statistical Software*, 61(3), 2014.

[15] Andrew M. Raim, Minglei Liu, Nagaraj K. Neerchal, and Jorge G. Morel. On the method of approximate Fisher scoring for finite mixtures of multinomials. *Statistical Methodology*, 18:115–130, 2014.

[16] Andrew M. Raim, Matthias K. Gobbert, Nagaraj K. Neerchal, and Jorge G. Morel. Maximum-likelihood estimation of the random-clumped multinomial model as a prototype problem for large-scale statistical computing. *Journal of Statistical Computation and Simulation*, 83(12):2178–2194, 2013.

Selected Reports & Proceedings

[1] Andrew M. Raim and Elizabeth Nichols. A comparison of map usability via bivariate ordinal analysis. Study Series: Statistics #2023-01, Center for Statistical Research and Methodology, U.S. Census Bureau, 2023.

[2] Andrew M. Raim and Kimberly F. Sellers. COMPoissonReg: Usage, the normalizing constant, and other computational details. Research Report Series: Computing #2022-01, Center for Statistical Research and Methodology, U.S. Census Bureau, 2022.

[3] Kyle M. Irimata, Andrew M. Raim, Ryan Janicki, James A. Livsey, and Scott H. Holan. Evaluation of Bayesian hierarchical models of differentially private data based on an approximate data model. Research Report Series: Statistics #2022-05, Center for Statistical Research and Methodology, U.S. Census Bureau, 2022.

[4] Andrew M. Raim, James A. Livsey, and Kyle M. Irimata. Browsing the 2010 Census SF2 summary file with R. Study Series: Computing #2022-01, Center for Statistical Research and Methodology, U.S. Census Bureau, 2022.

[5] Andrew M. Raim. Direct sampling in Bayesian regression models with additive disclosure avoidance noise. Research Report Series: Statistics #2021-01, Center for Statistical Research and Methodology, U.S. Census Bureau, 2021.

[6] Elizabeth Nichols, Erica Olmsted-Hawala, Andrew Raim, and Lin Wang. Attitudinal and behavioral differences between older and younger adults using mobile devices. In *Human Aspects of IT for the Aged Population. Technologies, Design and User Experience*. Springer Nature Switzerland AG, 2020.

[7] Elizabeth Nichols, Sarah Konya amd Rachel Horwitz, and Andrew Raim. 2020 census research and testing report: The effect of the mail delivery date on survey login rates and helpline call rates. U.S. Census Bureau, Research and Methodology Directorate, Center for Behavioral Science Methods Research Report Series (Survey Methodology) #2019-01, U.S. Census Bureau, 2019.

[8] Darcy Steeg Morris, Andrew M. Raim, and Kimberly F. Sellers. Introducing a Conway-Maxwell-multinomial distribution for flexible modeling of categorical data. In *JSM Proceedings, Biometrics Section*, pages 716–733, Alexandria, VA, 2018. American Statistical Association.

[9] Andrew M. Raim, Scott H. Holan, Jonathan R. Bradley, and Christopher K. Wikle. A model selection study for spatio-temporal change of support. In *JSM Proceedings, Government Statistics Section*, pages 1524–1540, Alexandria, VA, 2017. American Statistical Association.

[10] Andrew M. Raim. Informing maintenance to the U.S. Census Bureau's Master Address File with statistical decision theory. In *JSM Proceedings, Government Statistics Section*, pages 648–659, Alexandria, VA, 2016. American Statistical Association.

[11] Krista Heim and Andrew M. Raim. Predicting coverage error on the Master Address File using spatial modeling methods at the block level. In *JSM Proceedings, Survey Research Methods Section*, pages 1541–1555, Alexandria, VA, 2016. American Statistical Association.

[12] Andrew M. Raim, Marissa N. Gargano, Nagaraj K. Neerchal, and Jorge G. Morel. Bayesian analysis of overdispersed binomial data using mixture link regression. In *JSM Proceedings, Statistical Computing Section*, pages

2794–2808, Alexandria, VA, 2015. American Statistical Association.

[13] Andrew M. Raim and Marissa N. Gargano. Selection of predictors to model coverage errors in the Master Address File. Research Report Series: Statistics #2015-04, Center for Statistical Research and Methodology, U.S. Census Bureau, 2015.

[14] Andrew M. Raim, Nagaraj K. Neerchal, and Jorge G. Morel. Modeling overdispersion in R. Technical Report HPCF-2015-1, UMBC High Performance Computing Facility, University of Maryland, Baltimore County, 2015.

[15] Andrew M. Raim, Nagaraj K. Neerchal, and Jorge G. Morel. Large cluster approximation to the finite mixture information matrix with an application to meta-analysis. In *JSM Proceedings, Statistical Computing Section*, pages 4025–4037, Alexandria, VA, 2014. American Statistical Association.

[16] Andrew M. Raim and Nagaraj K. Neerchal. Modeling overdispersion in binomial data with regression linked to a finite mixture probability of success. In *JSM Proceedings, Statistical Computing Section*, pages 2760–2774, Alexandria, VA, 2013. American Statistical Association.

[17] Andrew M. Raim. Computational methods in finite mixtures using approximate information and regression linked to the mixture mean. Ph.D. Thesis, Department of Mathematics and Statistics, University of Maryland, Baltimore County, 2014.

[18] Andrew M. Raim. Introduction to distributed computing with pbdR at the UMBC High Performance Computing Facility. Technical Report HPCF-2013-2, UMBC High Performance Computing Facility, University of Maryland, Baltimore County, 2013.

[19] Andrew M. Raim, Brandon E. Fleming, and Nagaraj K. Neerchal. An analysis of categorical injury data using mixtures of multinomials. In *JSM Proceedings, Statistical Computing Section*, pages 2444–2458, Alexandria, VA, 2012. American Statistical Association.

[20] Andrew M. Raim and Matthias K. Gobbert. Parallel performance studies for an elliptic test problem on the cluster tara. Technical Report HPCF-2010-2, UMBC High Performance Computing Facility, University of Maryland, Baltimore County, 2010.

HPCF Technical Reports are available at `https://hpcf.umbc.edu/publications`, including some not listed in this document.

Software

[1] Andrew M. Raim, James A. Livsey, and Kyle M. Irimata. *sfreader: An R Package to Read Census Summary Files*, 2022. R package version 0.3.0. Available online at `https://github.com/andrewraim/sfreader`.

[2] Andrew M. Raim. *Direct Sampling*, 2021. R package version 0.1.3. Available online at `https://github.com/andrewraim/DirectSampling`.

[3] Andrew M. Raim and Darcy Steeg Morris. *COMMultReg: Conway-Maxwell Multinomial Regression*, 2020. R package version 0.1.0. Available online at `https://github.com/andrewraim/COMMultReg`.

[4] Andrew M. Raim, Scott H. Holan, Jonathan R. Bradley, and Christopher K. Wikle. *stcos: Space-Time Change of Support*, 2019. R package version 0.3.0. Available online at `https://cran.r-project.org/package=stcos`.

[5] Andrew M. Raim and Tommy Wright. *allocation: Exact Optimal Allocation Algorithms for Stratified Sampling*, 2019. R package version 0.0.1. Available online at `https://github.com/andrewraim/allocation`.

[6] Kofi P. Adragni, Sean R. Martin, Andrew M. Raim, and Wen Huang. *ManifoldOptim: An R Interface to the ROPTLIB Library for Riemannian Manifold Optimization*. R package version 1.0.0. Available online at `https://cran.r-project.org/package=ManifoldOptim`.

[7] Andrew M. Raim. *mixlink: Mixture Link Regression*. R package version 0.1.5. Available online at `https://cran.r-project.org/package=mixlink`.

[8] Kimberly Sellers, Thomas Lotze, and Andrew Raim. *COMPoissonReg: Conway-Maxwell Poisson (COM-Poisson) Regression*. R package version 0.8.0. Available online at `https://cran.r-project.org/package=COMPoissonReg`.

[9] Kofi Placid Adragni and Andrew Raim. *ldr: Methods for Likelihood-Based Dimension Reduction in Regression*. R package version 1.3.3. Available online at `https://cran.r-project.org/package=ldr`.

[10] Andrew M. Raim. *Overdispersion Models In R*. Available online at `http://github.com/andrewraim/OverdispersionModelsInR`.

Presentations

[1] Statistical modeling of the Mobile Questionnaire Assistance operation. Talk presented at 2023 Joint Statistical Meetings in Toronto, Canada on 8/8/2023.

[2] Rejection sampling for weighted densities by majorization. Talk presented at U.S. Census Bureau, Center for Statistical Research & Methodology Seminar (virtual seminar) on 6/8/2023.

[3] Rejection sampling for weighted densities by majorization. Poster with lighting talk presented at 2023 Symposium on Data Science and Statistics in St. Louis, MO on 5/24/2023–5/25/2023.

[4] Direct sampling in Bayesian hierarchical models for privacy protected data. Talk presented at 2021 International Conference on Advances in Interdisciplinary Statistics and Combinatorics (virtual meeting) on 10/8/2021.

[5] Direct sampling in Bayesian regression models with additive disclosure avoidance noise. Talk presented at U.S. Census Bureau, Center for Statistical Research & Methodology Seminar (virtual seminar) on 4/6/2021.

[6] Direct sampling in Bayesian regression models with additive disclosure avoidance noise. University of Kentucky Statistics Seminar Series (virtual seminar) on 4/2/2021.

[7] Sample size selection in continuation-ratio logit models. Poster presented at 2020 Joint Statistical Meetings (virtual meeting) on 8/3/2020.

[8] Statistical assessment of bovine body weight via functional gait data. Speed talk/poster presented at 2019 Joint Statistical Meetings in Denver, CO on 7/29/2019.

[9] A statistical comparison of call volume uniformity due to mailing strategy. Talk presented at 2018 Joint Statistical Meetings in Vancouver, CA on 7/31/2018.

[10] An R package for spatio-temporal change of support. Talk presented at ICSA 2018: International Conference on Recent Advances in Statistical Methodologies with Applications in Clinical & Official Statistics in Kerala, India on 1/3/2018.

[11] An R package for spatio-temporal change of support. Talk presented at 2017 Joint Statistical Meetings in Baltimore, MD on 8/2/2017.

[12] A flexible zero-inflated model to address data dispersion. Talk presented at International Conference on Statistical Distributions and Applications in Niagara Falls, Canada on 10/15/2016.

[13] Informing maintenance to the U.S. Census Bureau's Master Address File with statistical decision theory. Poster presented at 2016 Joint Statistical Meetings in Chicago, IL on 8/2/2016.

[14] An extension of generalized linear models to finite mixture outcomes. Talk presented at UMBC Statistics seminar on 2/19/2016.

[15] Selection of predictors to model coverage errors in the Master Address File. Talk presented at "What's Up: Updates on Three Decennial-Related Projects Edition" at Census Bureau, hosted by Research & Methodology Directorate on 9/30/2015.

[16] Selection of predictors to model coverage errors in the Master Address File. Talk presented at 2015 International Total Survey Error Conference in Baltimore, MD on 9/20/2015.

[17] Mixture link models for binomial data with overdispersion. Talk presented at 2015 Joint Statistical Meetings in Seattle, WA on 8/9/2015.

[18] Mixture link models for binomial data with overdispersion. Talk presented at 9th Annual Probability and Statistics Day at UMBC, Special Alumni Session, on 4/18/2015.

[19] Large cluster approximation to the information matrix using complete data. Talk presented at 2014 Joint Statistical Meetings in Boston, MA on 8/7/2014.

[20] Zero-inflated regression modeling for coverage errors of the Master Address File. Talk presented at 2014 Joint Statistical Meetings in Boston, MA on 8/7/2014 as substitute speaker for Derek Young.

[21] Mixture link models for binomial data with overdispersion. Talk presented at U.S. Census Bureau, Center for Statistical Research & Methodology Seminar on 11/23/2013.

[22] Modeling overdispersion using finite mixtures with a regression linked to the mean. Poster presented at 2013 Joint Statistical Meetings in Montreal, Canada on 8/6/2013.

[23] An analysis of categorical injury data using mixtures of multinomials. Poster presented at 7th Annual Probability and Statistics Day at UMBC on 4/27/2013, won 2nd prize in student posters.

[24] An approximate Fisher scoring algorithm for finite mixtures of multinomials. Talk presented at 2012 Joint Statistical Meetings in San Diego, CA on 7/31/2012.

[25] An approximate Fisher scoring algorithm for finite mixtures of multinomials. Talk presented at 34th Annual Graduate Research Conference at UMBC on 4/27/2012.

[26] An approximate Fisher scoring algorithm for finite mixtures of multinomials. Talk presented at 6th Annual Probability and Statistics Day at UMBC on 4/21/2012, won 1st prize in student talks.

[27] Maximum likelihood estimation of multinomial mixture models using high performance computing. Talk presented

at 39th Annual Meeting of the Statistical Society of Canada, 6/14/2011 at Acadia University, Wolfville, NS, Canada.

[28] The approximate Fisher information matrix for multinomial mixture models. Poster presented at 5th Annual Probability and Statistics Day at UMBC on 4/23/2011, won 3rd prize in student posters.

[29] Maximum likelihood estimation of the random-clumped multinomial model using high performance computing. Poster presented at Fourth Annual Probability and Statistics Day at UMBC 2010 on 4/24/2010, won 3rd prize in student posters.

[30] UMBC High Performance Computing Facility. Poster presented at recruiting event "U(2) Can Do Math+Stat Beyond College" for potential graduate students on 10/9/2009 to showcase the facility.

[31] An application of parallel computing to maximum likelihood estimation. Poster presented at 3rd Annual Probability and Statistics Day at UMBC on 8/25/2009.

[32] An application of parallel computing to maximum likelihood estimation. Poster presented at 31st Annual Graduate Research Conference at UMBC on 8/24/2009.

Workshops

[1] Advanced Statistical Programming with Rcpp. A two day (3 hours + 3 hours) Rcpp workshop presented to ~15 attendees at the U.S. Census Bureau on 5/24/2018 and 5/31/2018. Joint presentation with James Livsey (Census Bureau) and Iris Gauran (UMBC). Discussed basic examples, larger statistical applications, and some additional topics such as packages with Rcpp and object-oriented programming.

[2] High Performance Statistical Computing using R. A 2 day workshop on R presented to ~80 attendees 1/6/2018–1/7/2018 at St. Thomas College, Pala, Kerala, India. Presented with Nagaraj Neerchal (UMBC) and George Ostrouchov (Oak Ridge National Lab). Topics included introductory R/Rstudio, Rcpp, and parallel & distributed computing with pbdR.

[3] Advanced Statistical Programming with Rcpp. A 3.5 hour Rcpp workshop presented to ~15 attendees on 9/22/2017 at UMBC. Joint presentation with Iris Gauran (UMBC). Discussed basic examples, larger statistical applications, and some additional topics such as packages with Rcpp.

[4] Using R for Analysis of Big Data. Four day workshop (8/8/2016–8/11/2016) at King Mongkut's University of Technology Thonburi (KMUTT) in Bangkok, Thailand to ~40 attendees. Presented with Nagaraj Neerchal (UMBC) and George Ostrouchov (Oak Ridge National Lab). Topics included programming, statistical analysis, and high performance computing with R for big data applications.

[5] R supplement to *Analysis of Overdispersed Data using SAS*. Assisted workshop speakers Jorge G. Morel and Nagaraj K. Neerchal at 8th Annual Probability and Statistics Day at UMBC on 4/18/2014. Gave a short presentation on using R to achieve results comparable to SAS. Material was developed into Technical Report HPCF-2015-1 at `http://hpcf.umbc.edu/publications`.

[6] Introduction to pbdR with tara @ HPCF. Workshop on parallel computing with pbdR package in R, given at UMBC Dept. of Math and Statistics on 6/16/2013 to ~15 attendees. See Technical Report HPCF–2013–2 at `http://hpcf.umbc.edu/publications` for the handout.

[7] R workshop. Three day workshop (8/19, 8/20, & 8/24/2010) on R programming and statistical applications. Presented with Martin Klein and Nagaraj Neerchal. Material was prepared and presented at the UMBC Training Centers to a group of ~20 FDA employees. Topics included R Commander, basic programming, plotting, simulations, writing efficient code, debugging, databases, and object-oriented programming.

Organizations & Service

- Member of American Statistical Association (ASA).
- Refereed papers for American Statistician, Biometrical Journal, BMC Medical Research Methodology, Communications in Statistics: Theory and Methods, Communications in Statistics: Simulation and Computation, Computational Statistics and Data Analysis, Hacettepe Journal of Mathematics and Statistics, Heliyon, Journal of Statistical Computation and Simulation, Statistical Analysis and Data Mining, Statistical Methodology, Statistics and Computing, Statistics and Operation Research Transactions, and Mathematical Population Studies.
- Member of PhD committee for Elias Al-Najjar (UMBC 2015), Marilena Flouri (UMBC 2016), and Sai Popuri (UMBC 2017). Reader for Qing Ji (UMBC 2019), Reetam Majumder (UMBC 2021), and Nadeesri Wijekoon (UMBC 2021).

Computing Skills

- R for statistical and numerical computing. Includes associated tools such as Rstudio, Rcpp, and Rmarkdown. Experience authoring packages and preparing for CRAN. ★★★★
- Linux environment. ★★★★
- Source control with Git; also used Subversion and CVS. ★★★★
- Document preparation in LaTeX and Markdown, including papers, posters, slides, and web pages. ★★★★
- Software development in C and C++ w/ STL; also used Java. ★★★★
- Shell scripting with Bash and Python; also used Perl. ★★★
- Bayesian computing with Stan. ★★★
- Data wrangling and exploration with tidyverse, ggplot, SQL. ★★★
- Parallel computing on HPC clusters. Includes MPI with C/C++; SNOW, Rmpi, and pbdR packages for R; PBS and SLURM schedulers. ★★
- Technical computing with Julia and Python. ★★
- Data processing and analysis with SAS. ★

More ★'s indicate more frequently and recently used.

Software Engineering Experience. Documentation, diagramming, conceptual design, data design, software design, source control management, source code and database query analysis, system and code development, unit and system testing, best practices and processes, code and document reviews.