IAN TAYLOR

IAN.TAYLOR@COLOSTATE.EDU • (803) 606-6136 • HTTPS://WWW.IANTAYLOR.ORG/

RESEARCH TOPICS AND INTERESTS

Record Linkage, Network Statistics, Statistical Computing, Bayesian Modeling, Markov Chain Monte Carlo, Hierarchical Models, Statistical Consulting, Statistics Education, Data Visualization and Graphics

EDUCATION

2020–Present **Ph.D. Statistics**, *Colorado State University*, Fort Collins, CO.
"Bayesian Models and Streaming Samplers for Complex Data with Application to Network Regression and Record Linkage"
Expected graduation: May 2023
Ph.D. Advisors: Bailey K. Fosdick and Andee Kaplan

2017–2019 **M.S. Statistics**, *Colorado State University*, Fort Collins, CO.

2009–2013 **B.S. Mathematical Sciences**, *Clemson University*, Clemson, SC.
Concentration: Computer Science
GPA: 3.89 Manga Cum Laude

EMPLOYMENT

ACADEMIC

2022-Present **Graduate Statistical Consultant**, *Graybill Statistical Laboratory, CSU*, Fort Collins, CO.
- Consult with junior and senior researchers on statistical methods.
- Provide consultation on a variety of topics in walk-in lab.

2020–2022 **Graduate Research Assistant**, *Colorado State University*, Fort Collins, CO.
- Develop model for Bayesian record linkage of many files in a streaming context.
- Create MCMC algorithm for efficient Bayesian updating in streaming data contexts.

2017–2019 **Graduate Teaching Assistant**, *Colorado State University*, Fort Collins, CO.
- Teach recitation and lecture sections ranging from 25–60 students in size.
- Develop curriculum for STAT 301 and STAT 307 courses.

2010–2013 **Undergraduate Research**, *Intelligent River, Clemson University*, Clemson, SC.
- Research local correlation scores as statistical quality control test for real-time data.
- Implement score in distributed computing environment on Palmetto high performance cluster.

INDUSTRY

2016–2017 **Senior Analyst**, *CarMax Auto Finance*, Kennesaw, GA.
Extended Service Plans (ESP)

- Manage reserve forecast for ESP returns and claims, explore improvements to actuarial models. With team, refine models creating $2M reserve surplus.
- Research and implement Markov network clustering algorithm as record linkage solution for ESP claims data. Consolidate 350,000+ records into groups for each distinct repair facility.
- Identify independent repair facilities for vehicle service partnerships in markets with higher ESP claims risk.
- Supervise analyst intern for summer term. Guide research into improvements to risk segmentation by vehicle make/model.

2013–2016 **Strategy Analyst**, *CarMax Auto Finance*, Kennesaw, GA.
Servicing Analytics/Collections Strategy

- Design and evaluate A/B strategy testing in loan servicing
- Research and create behavioral scorecard to predict account delinquency. Reduced 1 to 30-day collections costs by 15%.
- Implement strategy for TCPA compliance.
- Automate collection of data from post-call customer surveys, create customer service evaluation metric.

2012 **Application Developer Intern**, *SCANA Corporation*, Cayce, SC.
Financial Reporting/Peoplesoft

- Develop business-critical financial reports in SAP
- Maintain financial report archives

PUBLICATIONS

- **Taylor, I.**, Kaplan, A., Betancourt, B. Fast Bayesian Record Linkage for Streaming Data Contexts. *Submitted* (2022+).
- Hughes, S., Rondeau, M., Shannon, S., Sharp, J., Ivins, G., Lee, J., **Taylor, I.**, Bendixson, B. A Holistic Self-learning Approach for Young Adult Depression and Anxiety Compared to Medication-Based Treatment-As-Usual. *Community Ment Health J* (2020). https://doi.org/10.1007/s10597-020-00666-9

COURSES TAUGHT

FULL COURSES

- STAT 307 – Introduction to Biostatistics (Spring–Fall 2019)
- STAT 301 – Introduction to Statistical Methods (Fall 2018)
- STAT 204 – General Statistics recitation (Spring 2018)
- STAT 201 – General Statistics recitation (Fall 2017)

GUEST LECTURES

- DSCI 445 – Statistical Machine Learning (Fall 2020)

- STAT 730 – Advanced Theory of Statistics (Spring 2020)
- STAT 440 – Undergraduate Bayes (Spring 2019)

CONFERENCES

PRESENTATIONS

- JSM 2022 – A Non-degrading Streaming Sampler for Recursive Bayesian Inference
- CMStatistics 2021 (Invited) – Fast Bayesian Record Linkage for Streaming Data Contexts
- JSM 2021 – Streaming Record Linkage for Online Data Deduplication
- ISBA 2021 – Restricted Regression in Networks

POSTERS

- CSU Statistics Student Posters 2022 – Fast Bayesian Record Linkage for Streaming Data Contexts
- ISBA 2021 – Restricted Regression in Networks
- CSU Statistics Student Posters 2021 – Sampling Methods for Streaming Record Linkage Models
- CSU Statistics Student Posters 2020 – Record Linkage: Basics and Streaming RL
- JSM 2019 – The Impact of Prior Choice on Latent Variable Network Models
- CSU Statistics Student Posters 2019 – Impact of Prior Choice on Network Models with Random Effects

ORGANIZATIONAL ROLES

- JSM 2022 Session Chair – Computing in Large and Complex Data Analysis
- IMS NRC 2019 at Colorado State University planning team

SEMINARS AND WORKSHOPS

- Coding and Cookies: Version Control Using Git (Spring 2022)
- Coding and Cookies: Reproducible Reports using Rmarkdown (Fall 2021)
- Coding and Cookies: Version Control Using Git (Spring 2021)

SOFTWARE

AUTHOR

bstrl	R package for Bayesian Streaming Record Linkage • GitHub: https://github.com/ianmtaylor1/bstrl • CRAN: https://cran.r-project.org/package=bstrl
gammacount	R package providing distributions related to renewal processes with gamma-distributed interarrival times • GitHub: https://github.com/ianmtaylor1/gammacount

CONTRIBUTIONS

wdman	R package that allows the user to manage the downloading/running of third party binaries relating to the webdriver/selenium projects.

- Project page: https://docs.ropensci.org/wdman/
- GitHub: https://github.com/ropensci/wdman
- CRAN: https://cran.r-project.org/package=wdman

Fixed a bug causing selenium binaries to not exit properly (December 2021)

adaptMCMC — R package that provides an implementation of the generic adaptive Monte Carlo Markov chain sampler proposed by Vihola (2011).

- GitHub: https://github.com/scheidan/adaptMCMC
- CRAN: https://cran.r-project.org/package=adaptMCMC

Incorporated third-party library for efficient rank one Cholesky update/downdate (March 2021)

AWARDS AND ACHIEVEMENTS

- Stanford Open Datathon 2021 finalist – data modeling track
- Phi Beta Kappa 2013
- Clemson University Honors College – General Honors
- National Society of Collegiate Scholars
- William Lowell Putnam Competition, 90th percentile

COMPUTING EXPERIENCE

R, Python, git, slurm, Rmarkdown, LaTeX, Stan, Bash, Linux