Rachel Longjohn

Website: rlongjohn.github.io **Email**: rlongjoh@gmail.com **LinkedIn**: @rachel-longjohn

Summary

Research scientist at the intersection of statistics and AI/ML, passionate about improving how both are applied in practical, real-world settings. Interests include AI/ML testing and evaluation, uncertainty quantification, model selection, and Bayesian statistics.

Education

Ph.D. in Statistics Expected 2025
University of California, Irvine (Advisor: Padhraic Smyth)

M.S. in Statistics Jun 2021
University of California, Irvine *(GPA: 4.00/4.00)*

B.S. in Applied and Computational Mathematics May 2019
University of Southern California *(GPA: 3.96/4.00)*

Work Experience

Graduate Student Researcher Jun 2020 - Present
University of California, Irvine

- Investigating statistical approaches for LLM testing and evaluation that leverage techniques from Bayesian statistics for uncertainty quantification.
- Developed and evaluated probabilistic models for categorical, sequential, and text data in the context of forensic science applications.

Statistical Sciences Intern Jun 2023 - Sep 2023
Los Alamos National Laboratory

- Investigated statistical approaches for the testing and evaluation of foundation models for multimodal data, including image, text, and seismic data.
- Developed methods of uncertainty quantification for aggregate ML evaluation metrics that support a better understanding of model performance.

Machine Learning Intern May 2018 - Aug 2019
Obsidian Security

- Constructed generalizable data representations for modeling cybersecurity insights in enterprise SaaS environments; interfaced with data from a variety of APIs, including Box, DropBox, GSuite, AWS, Office 365, Slack, and Salesforce.
- Engineered a property-based code testing suite in Python for testing ETL pipelines.

Publications

Longjohn, R., Nelson, K., & Smyth, P. (2025). Score-based likelihood ratios using authorship embeddings. (Under Review).

Longjohn*, R., Kelly*, M., Singh, S., & Smyth, P. (2024). Benchmark data repositories for better benchmarking. *NeurIPS*. https://arxiv.org/abs/2410.24100

Longjohn*, R., Gopalan*, G., & Casleton, E. (2024). Statistical uncertainty quantification for aggregate task-performance metrics in ML benchmarks. *NeurIPS Workshop on Statistical Frontiers in LLMs and Foundation Models*. https://arxiv.org/abs/2501.04234

Longjohn, R., & Smyth, P. (2024). Likelihood ratios for changepoints in categorical event data with applications in digital forensics. *Journal of Forensic Sciences*. https://doi.org/10.1111/1556-

4029.15512

Longjohn, R., Smyth, P., & Stern, H. S. (2022). Likelihood ratios for categorical count data with applications in digital forensics. *Law, Probability and Risk*. https://doi.org/10.1093/lpr/mgac016

Skills

Languages: Python, R, SQL

Software packages: PyTorch, scikit-learn, pandas, matplotlib, Tidyverse, Stan

Tools: git, quarto

Service and Teaching

Workshop Organizer 2025
The Future of Machine Learning Data Practices and Repositories, ICLR

Data Curator and Librarian 2020-25
UCI Machine Learning Repository

Reviewer 2022-24
NeurIPS, Datasets and Benchmarks

Teaching Assistant 2018-24
Inferential Statistics (2019, 2024), Data Analysis (2018-19)

Editor-in-Chief, Viterbi Conversations in Ethics 2018-19
University of Southern California

Awards and Fellowships

University of California, Irvine

- Rose Hill Foundation Science and Engineering Fellowship 2024
- Robert Newcomb Graduate Award in Statistics, Honorable Mention 2020

University of Southern California

- Summa Cum Laude 2019
- Phi Beta Kappa Honor Society 2017-19
- Dean's List 2015-19

Talks and Presentations

Longjohn*, R., Gopalan*, G., & Casleton, E. (2024). Statistical uncertainty quantification for aggregate task-performance metrics in ML benchmarks. *NeurIPS Workshop on Statistical Frontiers in LLMs and Foundation Models*. https://arxiv.org/abs/2501.04234

Longjohn, R., & Smyth, P. (2023a). Bayes factors for the existence of changepoints in categorical sequences within digital forensics. *Joint Statistical Meetings*.

Longjohn, R., & Smyth, P. (2023b). A likelihood ratio approach for detecting behavioral changes in device usage over time. *Annual Meeting of the American Academy of Forensic Sciences*.

Longjohn, R., & Smyth, P. (2022b). Tutorial on likelihood ratios with applications in digital forensics. *NIST Center for Statistics and Applications in Forensic Evidence Summer Webinar Series*.

Longjohn, R., & Smyth, P. (2022a). Likelihood ratios for categorical evidence with applications to digital forensics. *Joint Statistical Meetings*.

Longjohn, R., Smyth, P., & Stern, H. (2022). Likelihood ratios for categorical evidence with applications in digital evidence. *Annual Meeting of the American Academy of Forensic Sciences*.