Kayla Schroeder

kaylaaschroeder@gmail.com

LinkedIn

Education

Northwestern University   **Expected Graduation: Spring 2025**
PhD Candidate in Statistics

Northwestern University   **September 2022**
Masters of Science, Statistics

University of California, Los Angeles   **June 2020**
Bachelor of Science, Statistics
Minor in Mathematics
Cum Laude

Relevant Coursework
- **Statistics**: Statistical Theory and Methodology, Experimental Design and Analysis, Sampling Theory, Probability, Linear Models, Multivariate Analysis, Regression Analysis, Computation and Optimization, Qualitative Analysis, Time Series, Causal Inference, Bayesian Statistics, Monte Carlo Methods, Meta-Analysis, Human Rights Statistics
- **Mathematics:** Probability Theory, Proof-based Numeric Analysis, Proof-based Linear Algebra, Differential Equations, Multivariable Calculus, Optimization, Machine Learning
- **Programming:** Computational Statistics Programming, Statistical Programming with R, C++ Programming (data structures and object oriented programming), Programming in Python

Consulting and Research

Statistical Machine Learning Graduate Researcher, *Department of Statistics, Northwestern University*   January 2023 - Present
- Developed a reliability measure for generative large language model judgment, investigating downstream implications of reliability and connections between reliability and LLM judgment quality.
- Pioneered a novel causally sufficient dimension reduction technique for text data, significantly outperforming existing methods and drastically reducing prediction error by effectively addressing confounding variables while maintaining causal structure.
- Developed a groundbreaking suite of three metrics for evaluating topic model reliability, surpassing the state-of-the-art. Established a definitive, optimal metric, demonstrating its critical role in downstream applications.

- Developed codebases in both R and Python.
- Subsets of this work have been submitted to NAACL and will be submitted to ACL.

Experimental Design Graduate Research Assistant, *Statistics for Evidence-Based Policy and Practice (STEPP) Center, Northwestern University*   January 2022 - January 2023
- Working with Dr. Larry Hedges, developed a methodology to calculate effect sizes for time-series single case research designs. Developed, in R, a simulation to generate time-series single case data, validate the theoretical results and estimate both autocorrelation and between-person variance.
- Created understandable, complex graphics using the Trellis package's lattice plots to convey the results in an understandable manner. Worked to tell a story with the data using comprehensible graphics.
- Findings published in Contemporary Educational Psychology; others in preparation for submission.

Statistical Assistant and Consultant, *NSF Grant: Improving Evaluations of R&D in STEM Education Summer Institute, Northwestern University*   July 2022

Statistical Assistant and Consultant, *Meta-Analysis Training Institute (MATI), Georgia State University*   July 2022

Statistical Researcher, *UCLA Brain Mapping Center*   January 2018 - June 2020
- Developed, in R, linear mixed modeling capabilities to analyze output from automated MRI imaging software. Also significantly shortened analysis time and simplified the package for the user.
- Used R to add data visualizations to the region of interest and volumetric statistical analysis output.
- Made the statistical analysis output more understandable and straightforward to reproduce.
- Preprint submitted to Journal of Neuroscience Methods.

Honors and Awards

University Fellowship, Northwestern University   2020 - 2021
Dean's Honors List, UCLA   Awarded for 7 quarters

Papers

Schroeder, K. and Wood-Doughty, Z., 2024. Can You Trust LLM Judgments? Reliability of LLM-as-a-Judge. arXiv preprint arXiv:2412.12509.

Schroeder, K. and Wood-Doughty, Z., 2024. Reliability of Topic Modeling. arXiv preprint arXiv:2410.23186.

Mammadov, S. and **Schroeder, K**., 2023. A meta-analytic review of the relationships between autonomy support and positive learning outcomes. Contemporary Educational Psychology, p.102235.

Joshi, S., Kim, Y., **Schroeder, K.**, Leahy, A., Shattuck, D. *bstr - The BrainSuite Statistics Toolbox in R* ; preprint submitted to Journal of Neuroscience Methods

Posters and Presentations

Joshi, S., Kim, Y., **Schroeder, K.**, Joshi, A., Leahy, R., Shattuck, D. (2023, July 23). *MRI Volumetric and Surface Data Analysis with the BrainSuite Statistics Toolbox in R (bstr)* [Poster presentation]. Organization for Human Brain Mapping 2023 Conference, Montreal, Canada.

Schroeder, K. (2019, August 7). *Weather Uncertainty Total Deposition Novelty Detection* [Poster presentation]. Lawrence Livermore National Laboratory 2019 Poster Symposium, Livermore, CA, United States.

Schroeder, K., Pollack, T. (2018, August 6). *Nuclear Detonation Cloud Rise Analysis and Impact* [Poster presentation]. Lawrence Livermore National Laboratory 2018 Poster Symposium, Livermore, CA, United States.

Work History

Pinterest | Data Science Intern
Jun 2024 - Sept 2024
- Optimized experimentation processes by developing and implementing statistical methods to improve pre-experiment significance testing, significantly reducing costs and influencing product decisions.
- Enhanced fairness in feature launches by implementing a fairness index metric to identify and address potential biases, ensuring equitable user experiences.
- Developed production-ready data solutions using Python and SQL.

Palo Alto Networks | Business Intelligence Analyst Intern
June 2020 - September 2020
- In Tableau, developed visualizations of cloud security data for use by both business and technical teams to make product decisions. Worked to customize dashboards to explain the data effectively and provide insight across the customer success team. Leveraged visualization best practices to develop engaging and interactive Tableau dashboards.

- Worked with SQL, BigQuery and the company data lake to analyze cloud security accounts and develop data views. Used strong analytical and problem solving skills to understand new business processes and complex relationships between data systems.
- Created a mapping of the cloud security data lake including a visualization in LucidChart. Worked to create documentation that was understandable and useful to all throughout the customer success team.

Lawrence Livermore National Laboratory | Data Science Intern

June 2019-September 2019

- One of 30 chosen for this program out of 1,800 applicants.
- Developed a statistical model in Python for determining which models were most effective at forecasting the movement of hazardous waste. This work helped drastically reduce the time and computational intensity required to produce an accurate prediction for the hazardous waste movement.
- Developed a model in Python to detect novel observations and search for structure within hazardous material release weather data.
- Used Python to predict target values from weather uncertainty data predictor variables.

Lawrence Livermore National Laboratory | Nuclear Physics Data Analytics Intern

June 2018 - September 2018

- Working with a project team, I developed parts of a model that predicts cloud movement and final height. This model is being implemented into the National Atmospheric Release Advisory Committee's fallout model to more accurately predict potential nuclear detonation effects on surrounding areas.
- In order to model safety zones in a nuclear blast, I analyzed cloud rise from nuclear detonation films using a Microsoft Visual Basic tool.
- Developed key components of a program that reads in data and determines final cloud height while relying on external data as little as possible to improve the accuracy of the model's output.

Skills

- Programming: Python, R, SQL, SAS and C++
- Data Analytics: Tableau
- BigQuery

Teaching

Northwestern University, Department of Statistics

Curriculum Development Graduate Assistant   Summer 2022

- STAT 202: Introduction to Statistics and Data Science in R

Teaching Assistant

- STAT 202: Introduction to Statistics and Data Science in R   Fall 2021-Spring 2023
- STAT 301: Data Science with R sequence (301-1, 301-2, 301-3)   Fall 2023-Spring 2024

Community Service

Statistics without Borders
Delivery and Quality Assurance   March 2023-Present