

Kentaro Hoffman

✉ khoffm3@uw.edu

EMPLOYMENT

University of Washington, Seattle

Postdoctoral Scholar

Seattle, WA

June 1st 2023- Present

- Advised by: Tyler McCormick

Johns Hopkins University, Baltimore

Postdoctoral Scholar

Baltimore, MD

2022-2023

- Advised by: Peter Searson and Scott Zeger

AFFILIATIONS

Center for Statistics and the Social Sciences (UW)

eScience Institute (UW)

Postdoctoral Scientist

Data Science Fellow

EDUCATION

University of North Carolina, Chapel Hill

PhD in Statistics and Operations Research

Chapel Hill, NC

2017- 2022

- Dissertation: Borrowing from Your Neighbors: Three Statistical Techniques from Nontraditional Sources
- Advised by: Kai Zhang and Cynthia Rudin

Rice University

BA in Mathematics and Statistics

Houston, TX

2013- 2017

RESEARCH INTERESTS

Statistical Inference with AI, Interpretable and Responsible Machine Learning, Prediction-Powered Inference, Global Mortality Estimation, Electronic Medical Records, Active Learning, Experimental Design, Rashomon Sets

PUBLICATIONS

Peer Reviewed Publications.....

7. Stephen Salerno, Jiacheng Miao, Awan Afiaz, Kentaro Hoffman, Anna Neufeld, Qiongshi Lu, Tyler H. McCormick, Jeffrey T. Leek (2025) **ipd: An R Package for Conducting Inference on Predicted Data** *To appear in Biostatistics*
6. Kentaro Hoffman, Tyler McCormick (2024) **Bayesian Optimal Experimental Design of Streaming Data Incorporating Machine Learning Generated Synthetic Data** *NeurIPS 2024 Workshop on Bayesian Decision-making and Uncertainty*
5. Simon Dovan Nguyen, Kentaro Hoffman, and Tyler McCormick **Using Rashomon Sets for Robust Active Learning** (2024) *NeurIPS 2024 Workshop on Bayesian Decision-making and Uncertainty*

4. Adam Visokay, Trinity Fan, Kentaro Hoffman, Stephen Salerno, Jeffrey T. Leek, Li Liu, Tyler H. McCormick (2024) **From Narratives to Numbers: Valid Inference Using Language Model Predictions from Verbal Autopsy Narratives** *Presented at Conference on Language Models 2024*
3. Harsh Parikh*, Kentaro Hoffman*, Haoqi Sun*, Wendong Ge, Rajesh Amerineni, Lin Liu, Alexander Volfovsky, Sahar Zafar, Cynthia Rudin, M. Brandon Westover. (2023) **Effects of epileptiform activity on discharge outcome in critically ill patients in the USA: a retrospective cross-sectional study** *Lancet Digital Health* ,Vol. 5, Issue 8, pp. 495–e502
2. Hoffman, K. Lees, Johnathan, and Zhang, Kai. (2023) **Local Change Point Detection and Signal Cleaning using EEMD with applications to Acoustic Shockwaves** *Circuits Systems and Signal Processing* , Vol 42, Number 8, pp. 4669–4690
1. Hoffman, K. Babichev, A. and Dabaghian, Y. (2016) **A model of topological mapping of space in bat hippocampus..** *Hippocampus* 26: 1345-1353.

Peer Reviewed Short Comment.....

Ruobin Gong, Kentaro Hoffman, Yifan Cui, and Jan Hannig. **Technical Comment on “Policy impacts of statistical uncertainty and privacy”** *Science* DOI:10.1126/science.adf9724

Hoffman, K., Hannig, J. and Zhang, Kai. (2021) **Comments on “A Gibbs sampler for a class of random convex polytopes,** *Journal of the American Statistical Association* 116:535, 1206-1210

PREPRINT/UNDER REVIEW

- a) Stephen Salerno, Kentaro Hoffman, Awan Afiaz, Anna Neufeld, Tyler McCormick, and Jeffrey T. Leek **Sample Size Considerations for Post-Prediction Inference** *Under Review*
- b) Kentaro Hoffman, Stephen Salerno, Jeff T. Leek, Tyler McCormick **Some models are useful, but for how long?: A decision-theoretic approach to choosing when to refit large-scale prediction models** <https://arxiv.org/abs/2405.13926> *Under Review*
- c) Kentaro Hoffman, Qirui Zhao, Tyler McCormick **Two phase sampling with economic considerations using Post-Prediction Inference** *In preparation*
- d) Adam Visokay, Kentaro Hoffman, Stephen Salerno, Sasha Johfre, Tyler H. McCormick, **BMI can predict adiposity, but not well enough to learn about trends in obesity** *In Preperation*
- e) Kentaro Hoffman, Stephen Salerno, Awan Afiaz, Jeffrey T. Leek, Tyler H. McCormick **Do We Really Even Need Data?** *Proposal Accepted*
- f) Simon Dovan Nguyen, Kentaro Hoffman, and Tyler McCormick **Using Rashomon Sets for Robust Active Learning** *Under Review*

- g) Eddie Hock, Adam Visokay, Kentaro Hoffman, Tyler McCormick, **Embracing Multiplicity in Uncertain Times: Doing Social Science with Rashomon Sets** *In Preparation*
- h) Kentaro Hoffman, Tyler McCormick, and Jan Hannig. (2024) **Dempster-Shafer P-values: Thoughts on an Alternative Approach for Multinomial Inference** <https://arxiv.org/abs/2402.17070>

Press.....
 Fitzgerald, S. (2023, July 20). New study shows association between Epileptiform Activity and Poor ICU Outcomes. *Neurology Today*

HONORS AND AWARDS

eScience Institute Azure Compute Funding (\$ 6,500 Award)	2023-2024
eScience Institute Postdoctoral Fellowship (\$ 2,000 Award)	2023-2024
SAMSI Research Assistant Fellowship	2022
NIH BD2K Biomedical Graduate Fellow	2017-2022

SOFTWARE

IPD (co-owner) <https://cran.r-universe.dev/ipd/doc/manual.html>
 The Inference on Predicted Data is a package to allow one to perform valid statistical inference when some of the data is generated from a black-box AI model. *Avaiable from CRAN*

PRESENTATIONS (Invited and Given)

Meijo University Mathematics Seminar <i>Meijo University, Japan</i>	2025
A Decision-Theoretic Approach to Choosing when to Refit Large-scale Prediction Models <i>8th International Conference on Econometrics and Statistics (Invited Session)</i>	2025
Bayesian Experimental Design Incorporating Machine Learning Generated Synthetic Data <i>Joint Statistical Meeting 2025 (Invited Session)</i>	2025
Bayesian Experimental Design Incorporating Machine Learning Generated Synthetic Data <i>ENAR 2025 (Invited Session)</i>	2025
Valid Inference Using Language Model Predictions from Verbal Autopsy Narratives <i>eScience Institute (UW)</i>	2024
Bayesian Optimal Experimental Design of Streaming Data <i>NeurIPS Workshop on Bayesian Decision Making and Uncertainty</i>	2024
Inference on Predicted Data with Applications to Autopsies and Obesity <i>eScience Institute (UW)</i>	2024
Causal Estimation of Seizure-Like Brain Activity <i>UNC BIOS/STOR Joint Seminar</i>	2021
Multi-Resolution Inference for Multinomial Tests of Uniformity using Dempster-Shafer <i>Bayesian Frequentist and Fiducial 2021</i>	2021
Multi-Resolution Inference for Multinomial Tests of Uniformity using Dempster-Shafer <i>Universiy of Liverpool Institute for Risk and Uncertainty.</i>	2021
Multi-Resolution Inference for Multinomial Tests of Uniformity using Dempster-Shafer	

UNC STOR Graduate Seminar 2021
Causal Estimation of Seizure-Like Brain Activity
Joint Statistical Meeting, Seattle 2021

TEACHING EXPERIENCE

A Primer on ML and AI for Population Research (Co-Instructor)
Population Association of America 2025

CS&SS 594 A: AI and the Social Sciences (Guest Lecture): Word Embeddings Seattle, WA
University of Washington 2024

CS&SS 594 A: AI and the Social Sciences (Guest Lecture): Fitting Neural Networks with Scorchr
 Seattle, WA
University of Washington 2024

CS&SS 594 A: AI and the Social Sciences (Guest Lecture): Colinarity and Model Misspecification
 Seattle, WA
University of Washington 2024

PhD Qualifying Exam Recitation (Instructor) Chapel Hill, NC
University of North Carolina Chapel Hill 2021

STOR 320: Introduction to Data Science (Instructional Assistant) Chapel Hill, NC
University of North Carolina Chapel Hill 2021

STOR 120: Foundations of Statistics and Data Science (Primary Instructor) Chapel Hill, NC
University of North Carolina Chapel Hill 2020

STOR 155: Introduction to Data Models and Inference (Primary Instructor) Chapel Hill, NC
University of North Carolina Chapel Hill Fall 2018, Spring 2018, Fall 2019, Spring 2019

STAT 601: Neural Machine Learning (Instructional Assistant) Houston, TX
Rice University 2017

Government and Industry Experience

Adobe (Remote due to COVID)
Data Science Internship Summer 2020

Los Alamos National Laboratory Los Alamos, NM
SULI Internship Summer 2017

Argonne National Laboratory Lemont, IL
SULI Internship Summer 2016

MENTORSHIP

Troy Russo, Undergraduate, UW Statistics Directed Reading Project Co-Mentor
 Eddie Hock, PhD Candidate, UW Sociology Collaborated on: **g)**
 Simon Nugyen, PhD Candidate, UW Statistics Collaborated on: **5), f)**
 Qirui Zhao, MA Candidate, UW Statistics Collaborated on: **c)**
 Adam Visokay, PhD Candidate, UW Sociology Collaborated on: **4), d), g)**
 Trinity Fan, PhD Candidate, UW Statistics Collaborated on: **4)**
 Awan Afiaz, PhD Candidate, UW Biostatistics Collaborated on: **e)**
 Mikhail Ben-Joseph, B.A, UNC STOR Undergraduate Research Mentor

PROFESSIONAL ACTIVITIES

Programming Committee

NeurIPS workshop on BDU

NeurIPS 2024

Student Aide to Hiring Committee

UW 2023,2024

Graduate Mentor

ASA DataFest 2020

Referee Services *Management Science, NeurIPS BDU Workshop, Stat*