Alex Hayes

alexpghayes@gmail.com · 775 338 8842 · Madison, WI, USA · www.alexpghayes.com · github

Last updated on 2024-11-07

RESEARCH INTERESTS

Statistical network models, causal inference, spillover effects, multivariate analysis, applications in social science

EDUCATION

University of Wisconsin-Madison 2024
Ph.D. Statistics, advised by *Keith Levin* and *Karl Rohe*

University of Wisconsin-Madison 2020
M.S. Statistics

Rice University 2018
B.A. Statistics, with *Distinction in Research and Creative Work*

PROFESSIONAL EXPERIENCE

University of Wisconsin-Madison **August 2018 – Present**
PhD Candidate, Department of Statistics

Facebook **Summer 2020 & Summer 2021**
Research Intern, Core Data Science

RStudio **Summer 2018**
Intern, tidymodels team

Rice University **Fall 2017**
Undergraduate researcher with *Genevera Allen*

Fred Hutchinson Cancer Research Center **Summer 2017**
Undergraduate researcher *with Elizabeth Brown*

Houston Parks and Recreation Department **Spring 2016**
Undergraduate researcher

WORKING PAPERS

1. **Alex Hayes** and Kevin Levin. Peer effects in the linear-in-means model may be inestimable even when identified. *arXiv, in preparation for submission to Biometrika*. 2024.
2. **Alex Hayes**, Mark M. Fredrickson, and Keith Levin. Estimating network-mediated causal effects via principal components network regression. *Journal of Machine Learning Research (accepted with minor revisions)*. 2024+.

PUBLICATIONS

1. **Alex Hayes** and Karl Rohe. Co-factor analysis of citation networks. *Journal of Computational and Graphical Statistics*. 2024.
2. Hadley Wickham, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, **Alex Hayes**, Lionel Henry, Jim Hester, Max Kuhn, Thomas Lin Pedersen, Evan Miller, Kirill Müller, David Robinson, Dana Paige Seidel, Vitalie Spinu, Kohske Takahashi, Davis Vaughan, Claus Wilke, Kara Woo, Hiroaki Yutani. Welcome to the Tidyverse. *Journal of Open Source Software*. 2019.

RESEARCH & STATISTICAL SOFTWARE

1. fastRG (CRAN, Github): Sample generalized random dot product graphs linearly in edge counts. Useful to investigate properties of network models and spectral estimators. Easily control expected degree of sampled networks, and efficiently compute population eigendecompositions for large networks.
2. vsp (CRAN, Github): Vintage sparse PCA for semi-parametric network analysis. Estimate latent positions in random dot product graphs via spectral embeddings and interpret them via varimax rotation. Easily regularize networks to handle noise.
3. aPPR (Github): Approximate personalized pageRank. Locally clusters networks based on degree-regularized PageRank estimates. Designed specifically for large networks only available via an API.
4. gdim (CRAN, Github): Estimate graph dimension using cross-validated eigenvalues. Determine the number of communities in stochastic blockmodels and variants.
5. fastadi (CRAN, Github): Self-tuning matrix imputation. Estimating singular subspaces of sparsely observed matrices. Includes specialized methods for upper triangular data.
6. broom (CRAN, Github): Convert statistical objects into tidy tibbles. Part of the tidyverse. Puts hundreds of types of statistical estimates into a consistent format to make programming easier.
7. distributions3 (CRAN, Github): Probability distributions as S3 Objects. An object-oriented interface to probability computations, with emphasis on careful documentation, beginner friendliness and classroom applicability.

In addition to writing code, I collaborated with ROpenSci to design software development standards for statistical software, I review scientific software for ROpenSci, the R Journal, and the Journal of Open Source Software, and I helped organize the Chicago R Unconference in 2019.

PRESENTATIONS

1. Estimating network-mediated causal effects via spectral embeddings — 2024-06-17
   NetSci 2024
2. Asymptotic identification of peer effects in linear models — 2024-04-04
   Dissertation defense
3. Peer effects are parametrically indistinguishable from baseline behaviors in the asymptotic limit — 2023-11-27
   Statistics Graduate Student Seminar, UW-Madison
4. Latent contagion in low-rank networks — 2023-10-11
   Levin Lab Meeting, UW-Madison
5. Peer diffusion over uncertain networks — 2023-09-18
   IFDS Ideas Seminar, UW-Madison
6. Estimating network-mediated causal effects via spectral embeddings — 2023-08-09
   JSM 2023
7. Estimating network-mediated causal effects via spectral embeddings — 2023-04-24
   IFDS Ideas Seminar, UW-Madison
8. Estimating network-mediated causal effects via spectral embeddings — 2022-10-14
   Statistics Graduate Student Seminar, UW-Madison
9. Estimating indirect effects induced by homophily via spectral network regression. — 2022-07-07
   Tianxi Li and Can Le Joint Lab Meeting
10. distributions3: From basic probability to probabilistic regression — 2022-06-23
    UseR 2022
11. The Low Hanging Fruit of the Twitter Following Graph — 2021-08-11
    JSM 2021
12. Solving the model representation problem with broom — 2019-01-25
    rstudio::conf(2019)

13. Solving the model representation problem with broom 2018-11-30
Statistics Graduate Student Seminar, UW-Madison.
14. Convenient data analysis with broom 2018-11-14
RStudio Webinar Series
15. Solving the model representation problem with broom 2018-09-19
Madison R User Group

POSTERS

1. Estimating peer influence: limitations of linear-in-means models 2024-11-22
Wisconsin ASA Chapter Meeting
2. Estimating peer influence: limitations of linear-in-means models 2024-11-12
American Family Funding Initiative Networking Meeting
3. Estimating network-mediated causal effects via spectral embeddings 2023-08-07
IFDS Annual Meeting
4. Estimating network-mediated causal effects via spectral embeddings 2023-05-24
ACIC 2023
5. Using data to support real-time decision making by the Hurricane Harvey crisis management team 2017-10-10
Rice Data Science Conference
6. An exploratory analysis of the effect of waiting room interactions on adherence in clinical trials 2017-08-10
Fred Hutch Intern Poster Competition

TEACHING

Co-instructor

- Applied Machine Learning Workshop 2019-01-15 & 2019-01-16
Short course on computational tools for machine learning in R *rstudio::conf(2019)*

Graduate Teaching Assistant

- STAT 340 Intro to Data Modeling II Fall 2022
Undergraduate introduction to statistics with computational focus *UW-Madison*
- STAT 324 Intro to Statistics for Engineers Spring 2019
Undergraduate introduction to statistics for engineers *UW-Madison*
- STAT 324 Intro to Statistics for Engineers Fall 2018
Undergraduate introduction to statistics for engineers *UW-Madison*
- Statistics Department Outstanding TA Award 2018-2019
UW-Madison

Undergraduate Teaching Assistant

- COMP 540 Statistical Machine Learning Spring 2018
Masters-level course on theory and algorithms for predictive modeling *Rice University*
- COMP 330 Data Science: Tools & Models Fall 2017
Undergraduate course on databases and computational infrastructure for large scale data analysis *Rice University*

Guest Lecturer

- Two sample and paired hypothesis tests. STAT 310. *UW-Madison* 2024-11-05 & 2024-11-07
- Confidence intervals. STAT 340. *UW-Madison* 2022-10-25 & 2022-10-27
- Sampling with Twitter following graph with aPPR. STAT 992. *UW-Madison* 2020-10-08

- Hypothesis testing. STAT 324. *UW-Madison* 2018-10-18

MENTORING

- Nathan Kolbow (undergraduate research assistant), currently a PhD student in Biostatistics at UW-Madison

REFERENCES

- Karl Rohe, Professor of Statistics, UW-Madison `karl.rohe@wisc.edu`
- Keith Levin, Assistant Professor of Statistics, UW-Madison `kdlevin@wisc.edu`
- Hyunseung Kang, Associate Professor of Statistics, UW-Madison `hkang84@wisc.edu`