Andrew M. Thomas, PhD

andrew-thomas@uiowa.edu

612-281-6386

My Google Scholar

Education

June 2021 *Ph.D. Statistics*,
Purdue University.
Advisor: Takashi Owada.

August 2017 *M.S. Mathematical Statistics*,
Purdue University.

May 2014 *B.S. Statistics; Minor in Mathematics*,
B.A. Linguistics,
University of Minnesota, Twin Cities.

Employment

August 2023– *Assistant Professor*,
Department of Statistics and Actuarial Science,
University of Iowa.

July 2021–July 2023 *Postdoctoral Associate*,
Center for Applied Mathematics &
Department of Statistics and Data Science, Cornell University.
Advisors: David S. Matteson & Gennady Samorodnitsky.

2015–2021 *Graduate Research & Teaching Assistant*,
Department of Statistics, Purdue University.

Awards/Grants

- *Institute for Mathematics and its Applications (IMA) Travel award* **2022**.
- *I.W. Burr Award*, Department of Statistics, Purdue University. **2021**.
  This award recognizes "(a) promise of contribution to the profession as evidenced by academic excellence in courses and exams, and by the quality of the thesis research, (b) and excellence in teaching or consulting as a graduate student at Purdue."
- *Bilsland Dissertation Fellowship*. **Spring semester 2021**.
- *StatCom Community Service Award*, Department of Statistics, Purdue University. **2019**.

Research Interests

Topological data science, stochastic topology, machine learning in science, stochastic process limits, *U*-statistics, extreme value theory, functionals of point processes, changepoint detection.

Publications

Central limit theorems and asymptotic independence for local *U*-statistics on diverging halfspaces
Andrew M. Thomas
Bernoulli **29**(4), 3280–3306, (2023).

Feature detection and hypothesis testing for extremely noisy nanoparticle images using topological data analysis
Andrew M. Thomas, Peter A. Crozier, Yuchen Xu, and David S. Matteson.
Technometrics, DOI:10.1080/00401706.2023.2203744, (2023).

The VC-dimension of a class of multiples of the primes, and a connection to AdaBoost
Andrew M. Thomas
Online Journal of Analytic Combinatorics **17** (2022). Link

Functional strong laws of large numbers for Euler characteristic processes of extreme sample clouds
Andrew M. Thomas and Takashi Owada
Extremes, **24**(4), 699–724 (2021).

Functional limit theorems for the Euler characteristic process in the critical regime
Andrew M. Thomas and Takashi Owada
Advances in Applied Probability, **53**(1), 57–80 (2021).

Limit theorems for process-level Betti numbers for sparse and critical regimes
Takashi Owada and **Andrew M. Thomas**
Advances in Applied Probability **52**(1), 1–31 (2020).

Submissions

Bayesian changepoint detection via logistic regression and the topological analysis of image series
Andrew M. Thomas, Michael Jauch, David S. Matteson.
arXiv preprint. arXiv:2401.02917, (2024).

Novel integration of topological data analysis into a multimodal machine learning model to predict follicular carcinoma on ultrasound
Andrew M. Thomas, Ann C. Lin, Grace Deng, Yuchen Xu, Gustavo Fernandez-Ranvier, Aida Taye, Randall Owen, David S. Matteson, Denise Lee.
medRxiv preprint. medrxiv:10.1101, (2023).

Dynamic atomic column detection in transmission electron microscopy videos via ridge estimation
Yuchen Xu, **Andrew M. Thomas**, Peter A. Crozier, David S. Matteson.
arXiv preprint. arXiv:2302.00816, (2023).

Other Publications

StatCom at Purdue University: How We Do Good
Andrew M. Thomas
Amstat News, September 2020. Link

Software

Python package: `detectda`
Detection and hypothesis testing of features in greyscale videos using cubical persistent homology. Tailored for use with noisy nanoparticle videos.
Download from PyPI: https://pypi.org/project/detectda/

Presentations

Talks

September 2023 *Novel integration of topological data analysis into a multimodal machine learning model to predict follicular carcinoma on ultrasound.*
Oral abstracts #7: "Technology in Thyroid" .
American Thyroid Association Annual Meeting 2023.
Washington, DC, USA.

August 2023 *Learning topological statistics for Bayesian changepoint analysis.*
Topic-contributed paper session
"Inference for sequence data and applications" .
JSM 2023. Toronto, Ontario, Canada.

June 2023 *Central limit theorems and asymptotic independence for local U-statistics on diverging halfspaces.*
Contributed talk in "Best student paper award I" session.
Extreme Value Analysis 2023. Bocconi University, Milan, Italy.

May 2023 *The detecTDA algorithm: feature detection and hypothesis testing for extremely noisy images.*
Statistical Methods in Imaging Conference 2023.
Minneapolis, Minnesota, USA.

The persistent homology of stationary time series.
Conference on Advances in Time Series Analysis.
Chicago Booth School of Business, Illinois, USA.

February 2023 *The detecTDA algorithm: feature detection and hypothesis testing for extremely noisy images.*
Cornell Statistics Seminar. Cornell University, New York, USA.

September 2022 *Central limit theorems and asymptotic independence for local U-statistics on diverging halfspaces.*
Purdue Probability Seminar. Purdue University, Indiana, USA. (Invited)

July 2021 Invited talk in "Long memory processes and non-standard EVT" session.
Extreme Value Analysis 2021. The University of Edinburgh, Scotland, UK.

October 2020 *Functional strong laws of large numbers for Euler characteristic processes of extreme sample clouds*.
Purdue Probability Seminar. Purdue University, Indiana, USA.

August 2020 *20 Years of StatCom at Purdue: Statistics in the Community and what drives engagement*.
JSM 2020 Virtual Conference. Session title: "Understanding Statistics For Good: Integrating Service Learning Into Statistics Education". Slides

July 2020 *Functional limit theorems for Euler characteristic processes*.
Applied Algebraic Topology Research Network (AATRN). Online.

January 2020 *Limit theorems for Betti number and Euler characteristic processes*.
Workshop and winter school on geometric and topological data analysis.
Centro de Invesigación en Matemáticas, Guanajuato, Mexico.

September 2019 *Percolation results for Poisson Boolean models*.
Purdue Student Probability Seminar. Purdue University, Indiana, USA.

September 2018 *Limit theorems for process-level Betti numbers for sparse, critical and Poisson regimes*.
Purdue Probability Seminar. Purdue University, Indiana, USA.

Posters

September 2022 *Detection and hypothesis testing for extremely noisy videos using TDA*
AATRN/APATG poster session. Online.

June 2022 *Limit theory for local U-statistics on diverging halfspaces*.
Stochastic Networks Conference. Cornell University, New York, USA.

Teaching

- **Instructor of Record (Iowa)**
  - Probability and Stochastic Processes I (STAT 6300). Fall 2023.
- **Instructor of Record (Cornell)**
  - Understanding Machine Learning (STSCI 4750/5750). Spring 2022, Spring 2023.
- **Recitation leader and teaching assistant (Purdue)**
  - Elementary Statistical Methods (Stat 301).
    Totaling 10 semesters, Fall 2015-Spring 2020.
- **Online teaching assistant (Purdue)**
  - Statistical Methods (Stat 511). Summer 2020.

Leadership & Service

- **Assistant Editor** Data Science in Science (new journal) Link.
- **Reviewer**
  - Foundations of Data Science
  - IEEE / CVF Computer Vision and Pattern Recognition Conference (2023)
  - Journal of Applied and Computational Topology
  - Journal of the American Statistical Association
  - Electronic Journal of Statistics ($\times 2$)
  - Advances in Applied Probability
- **Project Leader and Recitation Instructor,** Tripods NSF REU-Stem For All 2021.
  July–August 2021.

- **Director,** StatCom Purdue (Statistics in the Community).
  August 2018–August 2020.

  StatCom is a statistical consulting organization run by graduate students that provides pro bono services for non-profits and local governments, as well as PK-12 engagement, in the greater Lafayette area and nationwide. My role involved recruiting student volunteers and clients, coordinating meetings for projects, advertising the organization, and facilitating communication within the organization between members.

- **Treasurer,** Purdue Statistics Graduate Student Organization.
  Fall 2019–Spring 2020.

Miscellaneous

Coding Python, R, SPSS.

Languages English (Native), Spanish (Intermediate).