Andrew M. Thomas, PhD
amt269@cornell.edu
612-281-6386
https://people.cam.cornell.edu/amt269/
My Google Scholar

Education

| Year      | Degree                  | Institution                               |
|-----------|-------------------------|-------------------------------------------|
| 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

| Year      | Role                           | Institution                                      |
|-----------|--------------------------------|--------------------------------------------------|
| July 2021 | 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, multiple Monte Carlo hypothesis testing, stochastic process limits, $U$-statistics, extreme value theory, functionals of point processes, boosting and its applications.

Publications

**Central limit theorems and asymptotic independence for local $U$-statistics on diverging halfspaces**
Andrew M. Thomas
To appear in *Bernoulli*. arXiv:2207.11142, 27 pp. (2022).

**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

**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, 27 pp. (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.
Submitted to *Technometrics*. arXiv:2209.13584, 30 pp. (2022).

**The VC-dimension of a class of multiples of the primes, and a connection to AdaBoost**
Andrew M. Thomas
Submitted to *Online Journal of Analytic Combinatorics*. arXiv:2208.06442, 9 pp. (2022).
Fractal dimension, approximation and data sets
Livia Betti, Ivan Chio, Julie Fleischman, Alex Iosevich, Filippo Iulianelli, Scott Kirila, Michele Martino, Azita Mayeli, Svetlana Pack, Zhelin Sheng, Conor Taliancic, Andrew [M.] Thomas, Nathan Whybra, Emmett Wyman, Ustun Yildirim, and Kaiyuan Zhao.
Submitted to SIAM Journal on Mathematics of Data Science. arXiv:2209.12079, 22 pp. (2022).

Other Publications

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

Teaching & Grading

• Instructor of Record (Cornell) Developed and taught machine learning course at Cornell which integrated theory, practice in Python and scikit-learn, and simulation to develop a holistic understanding of machine learning. Topics covered include: probably approximately correct (PAC) learning, bias-complexity tradeoff, regression, support vector machines, kernel learning, regularized loss minimization, convex learning problems, validation and k-fold cross validation, Monte Carlo simulation and integration, AdaBoost, stochastic gradient descent, and the basics of feedforward neural networks. Course was taught at master’s level.
  - Understanding Machine Learning (STSCI 4750/5750). Spring 2022, Spring 2023.

• Online teaching assistant (Purdue)
  - Statistical Methods (Stat 511). Summer 2020.

• Recitation leader and teaching assistant (Purdue) Responsibilities: leading Friday labs and helping students with their assignments; holding weekly office hours; preparing introductions to week's material to present at start of labs; grading assignments in a timely manner; completing lab assignments and checking them for errors before students receive them; writing “perfect” essay for students as grading key for peer-graded essay assignment.
  - Elementary Statistical Methods (Stat 301). Totaling 10 semesters, Fall 2015-Spring 2020.

• Grader (Purdue) Responsibilities: writing solutions to homeworks for students to view online; holding weekly office hours; grading assignments in a timely manner; substitute lecturing.
  - Probability Theory II (Stat 539). Fall 2020.
  - Introduction to Probability (Stat 519). Spring 2017, Spring 2019, Spring 2020.
  - Elements of Stochastic Processes (Stat 532). Fall 2017, Fall 2018, Fall 2019.
  - Probability Theory I (Stat 538). Spring 2018.
- Basic Probability and Applications (Stat 516). Spring 2016, Fall 2016.
- Experimental Statistics I (Stat 501). Fall 2015.

• Substitute lecturer (Purdue)
  - Elements of Stochastic Processes (Stat 532): 2 instances (branching chains, Poisson processes, renewal reward processes, regenerative processes).
  - Introduction to Probability (Stat 519). 1 instance ($L^p$ and weak convergence).

## Presentations

### Invited

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

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

### Contributed

**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 Investigación 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.

Leadership & Service

- **Assistant Editor** Data Science in Science (new journal) [Link].
- **Project Leader and Recitation Instructor**, Tripods NSF REU-Stem For All 2021. July–August 2021.
  Lead homework problem sessions in learning theory, linear predictors, and neural networks for motivated undergraduate students. Advised participants on research project investigating how neural networks learn on functions whose graphs have various Hausdorff dimensions. Program conducted virtually at the University of Rochester.
- **Reviewer**
  - Journal of Applied and Computational Topology
  - Journal of the American Statistical Association
  - Electronic Journal of Statistics
  - Advances in Applied Probability
- **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.
  My responsibilities as treasurer pertained primarily to making sure reimbursements for expenses were filed and that the Statistics graduate student organization was able to access its funds.

Miscellaneous

**Coding**  Python, R, SPSS.

**Languages**  English (Native), Spanish, French (Intermediate).