William G. Underwood

Centre for Mathematical Sciences, Wilberforce Road, Cambridge, CB3 0WB, UK
wgu21@cam.ac.uk wgunderwood.github.io

Employment

Postdoctoral Research Associate in Statistics Jul 2024 – Jul 2026
University of Cambridge

- Advisor: Richard Samworth, Department of Pure Mathematics and Mathematical Statistics
- Funding: European Research Council Advanced Grant 101019498
- Consultant, Cambridge Statistics Clinic

Supervisor and Drop-in Session Leader Jan 2025 – Jul 2026
University of Cambridge

- Part III Statistical Learning in Practice, Lent 2026
- Part IB Statistics, Lent 2025
- Part III Essay on Inference with Random Forests, Lent 2025
- Part III Modern Statistical Methods, Lent 2025
- Part III Topics in Statistical Theory, Lent 2025
- Part III Concentration Inequalities, Lent 2025

Assistant in Instruction Sep 2020 – May 2024
Princeton University

- ORF 499: Senior Thesis, Spring 2024
- ORF 498: Senior Independent Research Foundations, Fall 2023
- SML 201: Introduction to Data Science, Fall 2023
- ORF 363: Computing and Optimization, Spring 2023, Fall 2020
- ORF 524: Statistical Theory and Methods, Fall 2022, Fall 2021
- ORF 526: Probability Theory, Fall 2022
- ORF 245: Fundamentals of Statistics, Spring 2021

Education

PhD in Operations Research & Financial Engineering Sep 2019 – May 2024
Princeton University

- Dissertation: Estimation and Inference in Modern Nonparametric Statistics
- Advisor: Matias Cattaneo, Department of Operations Research & Financial Engineering

MA in Operations Research & Financial Engineering Sep 2019 – Sep 2021
Princeton University

MMath in Mathematics & Statistics Oct 2015 – Jun 2019
University of Oxford

- Dissertation: Motif-Based Spectral Clustering of Weighted Directed Networks
- Supervisor: Mihai Cucuringu, Department of Statistics

Research & publications

Articles

- Yurinskii's coupling for martingales, with M. D. Cattaneo and R. P. Masini. *Annals of Statistics*, forthcoming, 2025. arXiv:2210.00362.
- Uniform inference for kernel density estimators with dyadic data, with M. D. Cattaneo and Y. Feng. *Journal of the American Statistical Association*, *119*(548), 2695–2708, 2024. arXiv:2201.05967.
- Motif-based spectral clustering of weighted directed networks, with A. Elliott and M. Cucuringu. *Applied Network Science*, *5*(62), 2020. arXiv:2004.01293.
- Simple Poisson PCA: an algorithm for (sparse) feature extraction with simultaneous dimension determination, with L. Smallman and A. Artemiou. *Computational Statistics*, 35:559–577, 2019.

Preprints

- Upgrading survival models with CARE, with H. W. J. Reeve, Oliver Y. Feng, Samuel A. Lambert, Bhramar Mukherjee and Richard J. Samworth. *Submitted*, 2025. arXiv:2506.23870.
- Sharp anti-concentration inequalities for extremum statistics via copulas, with M. D. Cattaneo and R. P. Masini. *Submitted*, 2025. arXiv:2502.07699.
- Inference with Mondrian random forests, with M. D. Cattaneo and J. M. Klusowski. *Submitted*, 2025. arXiv:2310.09702.

Presentations & conferences

- London Symposium on Information Theory, Cambridge, May 2025.
- International Conference on Statistics and Data Science, Nice, December 2024.
- Statistics Seminar, University of Pittsburgh, February 2024.
- Statistics Seminar, University of Illinois Urbana-Champaign, January 2024.
- Statistics Seminar, University of Michigan, January 2024.
- Eighth Princeton Day of Statistics, Princeton University, November 2023.
- PhD Poster Session, Two Sigma Investments, New York, July 2023.
- Statistical Foundations of Data Science and their Applications, Princeton University, May 2023.
- Research Symposium, Two Sigma Investments, New York, June 2022.
- Seventh Princeton Day of Statistics, Princeton University, October 2021.
- Statistics Laboratory, Princeton University, September 2021.

Software

- care-survival: upgrading survival models with CARE in Python, 2025.
  GitHub: wgunderwood/care-survival
- tex-fmt: LaTeX formatter written in Rust, 2024. GitHub: wgunderwood/tex-fmt
- MondrianForests: Mondrian random forests in Julia, 2023.
  GitHub: wgunderwood/MondrianForests.jl
- DyadicKDE: dyadic kernel density estimation in Julia, 2022.
  GitHub: wgunderwood/DyadicKDE.jl
- motifcluster: motif-based spectral clustering in R, Python and Julia, 2020.
  GitHub: wgunderwood/motifcluster

Awards & funding

- School of Engineering and Applied Science Award for Excellence, Princeton University 2022
- Francis Robbins Upton Fellowship in Engineering, Princeton University 2019
- Royal Statistical Society Prize, Royal Statistical Society & University of Oxford 2019
- Gibbs Statistics Prize, University of Oxford 2019
- James Fund for Mathematics Research Grant, St John's College, University of Oxford 2017
- Casberd Scholarship, St John's College, University of Oxford 2016

Professional experience

Quantitative Research Intern Jun 2023 – Aug 2023
Two Sigma Investments

Machine Learning Consultant Oct 2018 – Nov 2018
Mercury Digital Assets

Educational Consultant Feb 2018 – Sep 2018
Polaris & Dawn

Premium Tutor Feb 2016 – Oct 2018
MyTutor

Statistics & Machine Learning Researcher Aug 2017 – Sep 2017
Cardiff University

Data Science Intern Jun 2017 – Aug 2017
Rolls-Royce

Service

Peer review
Annals of Statistics; Bernoulli; Econometric Theory; Information Theory, Probability and Statistical Learning; Journal of the American Statistical Association; Journal of Business & Economic Statistics; Journal of Causal Inference; Journal of Econometrics; Journal of Machine Learning Research; Journal of Nonparametric Statistics; Journal of the Royal Statistical Society, Series B; Operations Research.

References

- Richard Samworth, Professor, Statistical Laboratory, University of Cambridge
- Matias Cattaneo, Professor, ORFE, Princeton University
- Jianqing Fan, Professor, ORFE, Princeton University