Eugene Katsevich

Assistant Professor of Statistics and Data Science

Academic Research Building 311
265 South 37th Street
Philadelphia, PA 19104-1686

Phone: (650) 250 3331
Email: ekatsevi@wharton.upenn.edu
Homepage: https://ekatsevi.github.io/

Employment

University of Pennsylvania, Department of Statistics and Data Science, Assistant Professor, 2020-present.

Carnegie Mellon University, Department of Statistics and Data Science, Postdoctoral Researcher, 2019-2020.
Advisors: Kathryn Roeder and Aaditya Ramdas.

Education

Stanford University, Ph.D. in Statistics, 2019.
Thesis Advisor: Chiara Sabatti.

Princeton University, A.B. in Mathematics (with Highest Honors), 2014.
Thesis Advisor: Amit Singer.

Funding

- NSF DMS-2113072, awarded for "Testing and estimation for multi-modality single cell genomics." $184,999 (2021-2024).
- Wharton Analytics Fund, awarded for "Statistical Software for Single Cell CRISPR Screens." $75,000 (2021-2023).

Awards

- Wharton Teaching Excellence award (2021)
- Jerome H. Friedman Applied Statistics Dissertation Award (2019).
- Statistics Department Teaching Assistant Award (2016).
- Hertz Fellowship (2014-2019).
- National Defense Science and Engineering Fellowship (2014-2017).
- George B. Covington Thesis Prize in Mathematics (2014).
- Early election to Phi Beta Kappa (2013).
- Barry Goldwater Scholarship (2012).
- Shapiro Prize for Academic Excellence (2011, 2012).
- Freshman First Honor Prize (2011).

Teaching

Courses Taught at University of Pennsylvania
- STAT 961: Statistical Methodology, Fall '21 – PhD Level.
- STAT 471: Modern Data Mining, Spring '21, Fall '21 – Undergraduate Level.

Courses Taught at Stanford
- STATS 302: Qualifying Exams Workshop, Summer '17 – Graduate Level.

Training

Doctoral Student Supervision
- Timothy Barry; jointly with Kathryn Roeder (Spring 2020 – present)
- Sophia Lu (Summer 2021)
- Abhinav Chakraborty (Fall 2021 – present)

Master's Student Supervision
- Ziang Niu (Spring 2022 – present)

Professional Service Activities

- *Reviewer*, Annals of Statistics (4), Journal of the Royal Statistical Society (Series B) (4), Journal of the American Statistical Association (4), Annals of Applied Statistics (1), Statistical Science (1), Proceedings of the National Academy of Sciences (1), Statistics and Probability Letters (1), Journal of Computational and Graphical Statistics (1), Journal of Multivariate Analysis (1), Electronic Journal of Statistics (1), Genetics (1), Genome Medicine (1).
- *Organizer*, Stanford statistics department orientation program for PhD students (2018).
- *Organizer*, Hertz West Coast Retreat (2017).
- *Academic chair*, Princeton Math Club (2012).
- *Head problem writer*, Princeton University Math Competition (2011).

Presentations

Invited talks
- *Statistical Analysis of Single Cell CRISPR Screens.*
  Biostatistics seminar at Yale University,
  February 2, 2021, held virtually.
- *Statistical Analysis of Single Cell CRISPR Screens.*
  Open Insights in Biomedical Data Science Seminar at the University of Pennsylvania,
  November 13, 2020, held virtually.

- *Finite-sample optimality and large-sample power analysis of the conditional randomization test.*
  Young Data Science Researcher Seminar at ETH Zurich, November 6, 2020, held virtually.

- *Statistical analysis of single cell CRISPR screens.*
  Stanford Biostatistics Workshop, October 8, 2020, held virtually.

- *The conditional randomization test in theory and in practice.*
  International Seminar on Selective Inference, Aug. 20, 2020, held virtually.

- *A theoretical treatment of conditional independence testing under Model-X.*
  Joint Statistical Meetings, Aug. 2–6, 2020, held virtually.

- *Multiple testing for modern data: structure, curation, and replicability.*
  U.C. San Diego Halicioglu Data Science Institute Special Seminar, Feb. 12, 2019.

- *Multiple testing for modern data: structure, curation, and replicability.*
  Rutgers University Statistics Seminar, Feb. 5, 2019.

- *Multiple testing for modern data: structure, curation, and replicability.*
  Wharton Statistics Seminar, University of Pennsylvania, Jan. 30, 2019.

- *Multiple testing for modern data: structure, curation, and replicability.*
  University of Chicago Statistics Colloquium, Jan. 24, 2019.

- *Controlling FDR while highlighting distinct discoveries, with applications to GO enrichment analysis.*
  Stanford University Biostatistics Workshop, Oct. 11, 2018.

- *Controlling FDR while highlighting distinct discoveries, with applications to GO enrichment analysis.*
  U.C. Berkeley Statistics and Genomics Seminar, Sep. 27, 2018.

Contributed talks

- *A theoretical treatment of conditional independence testing under Model-X.*
  Bernoulli-IMS One World Symposium, Aug. 24–28, 2020, held virtually.

- *Simultaneous High-Probability Bounds on the False Discovery Proportion
  for Nested Sequences of Rejection Sets.*
  International Conference on Multiple Comparison Procedures, Dec. 13–15, 2019, in Taipei, Taiwan.

- *Gene Ontology enrichment testing: Reconciling FDR control with filtering.*
  Joint Statistical Meetings, Jul. 28–Aug. 2, 2018, in Vancouver, Canada.

- *The multilayer knockoff filter: Controlled multi-resolution variable selection.*
  International Conference on Multiple Comparison Procedures, Jun. 20–23, 2017, in Riverside, California.

Poster Presentations

- *Conditional resampling improves sensitivity and specificity of genome-wide CRISPR regulatory screens.*
  American Society for Human Genetics, Oct. 27–30, 2020, held virtually.

- *Controlling FDR While Highlighting Selected Discoveries.*
  International Conference on Multiple Comparison Procedures, Dec. 13–15, 2019, in Taipei, Taiwan.

- *Multi-resolution association analysis for exome-wide sequencing.*
  American Society for Human Genetics, Oct. 16–20, 2018, in San Diego, California.

- *Controlling FDR while highlighting distinct discoveries.*
  Workshop on Higher-Order Asymptotics and Post-Selection Inference, Sep. 8–10 2018, in St. Louis, Missouri.

- *Multilayer FDR control for genetic association studies.*
  Graybill Conference on Statistical Genomics and Genetics, Jun. 5–7 2017, in Fort Collins, Colorado.
  Best student poster award.

- *The multilayer knockoff filter: Multilayer FDR control for association studies.*
  Probabilistic Modeling in Genomics, Sep. 12–14, 2016, in Oxford, United Kingdom.

Workshop Participation

- *Participant*, Simons Institute Workshop on Robust and High-Dimensional Statistics (October 2018).

- *Participant*, Joint UCLA and Stanford Statistical Genetics Programming Workshop (Summer 2016).

- *Participant*, UCLA Computational Genomics Summer Institute short course (Summer 2016).

Preprints

[1] **E. Katsevich** and A. Ramdas. A theoretical treatment of conditional independence testing under Model-X. Available on arXiv.

[2] J. Tian, X. Chen, **E. Katsevich**, J. Goeman, and A. Ramdas. Large-scale simultaneous inference under dependence. Available on arXiv.

[3] J. A. Morris, Z. Daniloski, J. Domingo, T. Barry, M. Ziosi, D. A. Glinos, S. Hao, E. Mimitou, P. Smibert, K. Roeder, **E. Katsevich**, T. Lappalainen, and N. E. Sanjana. Discovery of target genes and pathways of blood trait loci using pooled CRISPR screens and single cell RNA sequencing. Available on bioRxiv.

[4] T. Barry, **E. Katsevich**, and K. Roeder. Exponential family measurement error models for single-cell CRISPR screens. Available on arXiv.

Publications

[1] **E. Katsevich**, A. Katsevich, and G. Wang. Stability of the interior problem for polynomial region of interest. *Inverse Problems*, 28(6), 2012. Available at PubMed.

[2] B. Shi, **E. Katsevich**, B. Chiang, A. Katsevich, and A. Zamyatin. Image registration for motion estimation in cardiac CT. In *SPIE Medical Imaging*, San Diego, California, February 2014. Available at SPIE digital library.

[3] **E. Katsevich**, A. Katsevich, A. Singer. Covariance matrix estimation for the cryo-EM heterogeneity problem. *SIAM Journal on Imaging Sciences*, 8(1):126–185, 2015. Available at PubMed.

[4] J. Anden, **E. Katsevich**, and A. Singer. Covariance estimation using conjugate gradient for 3D classification in cryo-EM. In *IEEE Int Symp Biomed Imaging*, New York, New York, April 2015. Available at PubMed.

[5] **E. Katsevich** and C. Sabatti. Multilayer Knockoff Filter: Controlled variable selection at multiple resolutions. *Annals of Applied Statistics*, 13(1):1–33, 2019. Available at Project Euclid.

[6] J. Zhu, Q. Zhao, **E. Katsevich**, C. Sabatti. Exploratory Gene Ontology Analysis with Interactive Visualization. *Nature Scientific Reports*, 9:7793, 2019. Available at nature.com.

[7] M. Sesia, **E. Katsevich**, S. Bates, E. Candès, C. Sabatti. Multi-resolution localization of causal variants across the genome. *Nature Communications*, 11:1093, 2020. Available at nature.com.

[8] **E. Katsevich** and A. Ramdas. Simultaneous high-probability bounds on the false discovery proportion in structured, regression, and online settings. *Annals of Statistics*, 48(6):3465–3487, 2020. Available at Project Euclid.

[9] **E. Katsevich**, C. Sabatti, and M. Bogomolov. Filtering the rejection set while preserving false discovery rate control. *Journal of the American Statistical Association*, 2021. Available at Taylor and Francis Online.

[10] M. Liu, **E. Katsevich**, L. Janson, and A. Ramdas. Fast and powerful conditional randomization testing via distillation. To appear in *Biometrika*, 2021. Available at Oxford University Press.

[11] T. Barry, X. Wang, J. A. Morris, K. Roeder, and **E. Katsevich**. SCEPTRE improves calibration and sensitivity in single-cell CRISPR screen analysis. Available at Genome Biology.

Last updated: February 13, 2022