

Ana Maria Kenney

Office: 2062 Donald Bren Hall,
Irvine CA 92617

Email: anamaria.kenney@uci.edu

Academic Positions

Assistant Professor 2023-Present
UC Irvine Department of Statistics

Postdoctoral Scholar 2021-2023
UC Berkeley Department of Statistics

Mentor: Professor Bin Yu

Education

Ph.D. in Statistics and Operations Research 2021
Pennsylvania State University

Advisor: Prof. Francesca Chiaromonte (Dep. of Statistics)

Co-advisor: Prof. Matthew Reimherr (Dep. of Statistics)

Thesis: Mixed Integer Programming, Whitening, and Functional Data Analysis: Improving Feature Selection in “Omics” Research

M.S. in Applied Mathematics 2015
University of Colorado Denver

Advisor: Prof. Alexander Engau (Dep. of Mathematical and Statistical Sciences)

Thesis: Optimizing the Control of Hybrid Vehicles Using Model Predictive Control

B.S. in Mathematics (Summa Cum Laude) 2012
California State University Stanislaus
Concurrent with HS Diploma

Research Interests

I work at the interface of statistics, interpretable machine learning, and large-scale optimization mainly applied to biomedical research for clinical decision making. These works span several areas including cardiovascular genetics, Omics contributions to early infant growth, and early cancer screening. Working with an interdisciplinary group of statisticians, biologists, bioinformaticians, clinicians and health professionals has instilled a nuanced and pragmatic edge to my work. I emphasize interpretability, stability, robustness, computational efficiency and practical relevance when approaching each project and technique. I aim to support methodology with useful, informative theoretical evidence. Additionally, having been trained in the NIH funded Biomedical Big Data to Knowledge program (B2D2K), I received specialized training to tackle challenges posed by working with large amounts of sensitive biomedical data - including scalability and data privacy for ethical use.

Topics: Statistical machine learning and interpretability; tree-based modeling; large-scale (non-convex) optimization; robust statistics; functional data analysis; differential privacy; statistical genomics; metabolomics, and microbiome research for clinical use; large language models for proteomics; image analysis and segmentation of bioimages

Publications

(*co-first author, † advised graduate student, ‡ UCI student)

8. Wang Q., Tang T.M., Youlton N., Weldy C., **Kenney A.M.**[†], Ronen O., Hughes W., ... & Yu B., Ashley E., (2025) Epistasis regulates genetic control of cardiac hypertrophy. (To Appear in Nature Cardiology) medRxiv preprint <https://doi.org/10.1101/2023.11.06.23297858>
7. Tang, T.M., Zhang Y., **Kenney, A.M.**, Xie C., Xiao L., ..., & Chinnaiyan A., Yu B., (2025). A simplified MyProstateScore2.0 for High-Grade Prostate Cancer. (To appear in Cancer Biomarkers) Manuscript available upon request.
 - MyProstateScore2.0 resulted in a patent for which I am a co-inventor
6. Irajizad E., **Kenney A.M.**, Tang T.M., Vykoukal J., Wu R., Murage E., Dennison J.B., Escofet M.S., Long J.P., Maitra A., Do K., Yu B., Wolpin B., Hanash S., Fahrman J.F. (2023). A blood-based metabolomic signature predictive of risk for pancreatic cancer. **Cell Reports Medicine**, 4(9).
 - Selected for comments, see: Oldfield, L., & Costello, E. (2023). Where the metabolome meets the microbiome for pancreatic cancer detection. **Cell Reports Medicine**, 4(9).
5. Craig, S. J.*^{*}, **Kenney, A.M.***, Lin, J., Paul, I. M., Birch, L. L., Savage, J. S., ... & Makova, K. D. (2023). Constructing a polygenic risk score for childhood obesity using functional data analysis. **Econometrics and Statistics Special Issue in Biostatistics**.
4. Insolia, L.*^{*}, **Kenney, A.M.***, Chiaromonte, F., & Felici, G. (2022). Simultaneous feature selection and outlier detection with optimality guarantees. **Biometrics**.
 - Open material badge awardee
3. Insolia, L., **Kenney, A.M.***, Calovi, M., & Chiaromonte, F. (2021). Robust Variable Selection with Optimality Guarantees for High-Dimensional Logistic Regression. **Stats Special Issue: Robust Statistics in Action**, 4(3), 665-681.
2. **Kenney, A.M.**, Chiaromonte, F., & Felici, G. (2021). MIP-BOOST: Efficient and Effective L_0 Feature Selection for Linear Regression. **Journal of Computational and Graphical Statistics**, 1-12.
1. Awan J.*^{*}, **Kenney A.M.***, Reimherr M., Slavkovi A. (2019). Benefits and Pitfalls of the Exponential Mechanism with Applications to Hilbert Spaces and Functional PCA. **International Conference on Machine Learning (ICML)**: 375-384. (22% acceptance rate)

Under Revision and Preprints

(*co-first author, † advised graduate student, ‡ UCI student)

2. Huang M.*^{*}, **Kenney, A.M.***, Tang T.M.*^{*}, (2025+). Distilling Causal Effects: Stable subgroup estimation via distillation trees in causal inference (To be submitted to JASA: Special Issue on Statistical Science in Artificial Intelligence, December 2024). arXiv preprint arXiv:2502.07275
 - Tom Ten Have Poster Award Honorable Mention for Earlier Career Researchers (7 selected at the 2024 American Causal Inference Conference)
1. Agarwal, A.*^{*}, **Kenney, A.M.***, Tan, Y. S.*^{*}, Tang, T. M.*^{*}, & Yu, B. (2025+). MDI+: A Flexible Random Forest-Based Feature Importance Framework. (Under revision at JASA) arXiv preprint arXiv:2307.01932.

Funding and Fellowships

2023-2024	CORCL Award \$3883 in research expenses
2019-2021	Alfred P. Sloan Minority PhD Scholar \$20,000 in professional and research expenses
2019	Bunton-Waller Fellowship For promising students from underrepresented backgrounds, covers stipend and tuition
Summer 2018	Visiting graduate researcher Sant'Anna School of Advanced Studies (Pisa, IT), covered travel and housing
2017/18/20	B2D2K Stipend, tuition, and professional travel assistance provided by the NIH-funded Biomedical Big Data to Knowledge predoctoral training program
2015/2016	University Graduate Fellowship The highest tier of university fellowships at Penn State, provides a stipend and tuition assistance to incoming graduate students displaying high research potential
2012-2015	Lynn Bateman Fellow Stipend and tuition assistance

Select Honors and Awards

2024	Tom Ten Hive Poster Award Honorable Mention American Causal Inference Conference
2024-25	Cultural, Competence in Computing (3C) Fellow Accepted as a team with Professors Mine Dogucu (UCI), Analisa Flores (UCR), and Marcela Alfaro-Cordoba (UCSC)
2021	Open Material Badge Biometrics Journal
2019	SREB travel award Institute on Teaching and Mentoring hosted by the Southern Regional Education Board
2019	ASA student/early career travel award Women in Statistics and Data Science Conference (WSDS)
2018	ISNPS young investigator travel award 4th Conference of the International Society for Nonparametric Statistics (ISNPS)
2017	CGSI travel and housing award Computational Genomics Summer Institute (CGSI)
2016	Robert Graham Endowed Fellowship
2016	Jack and Eleanor Pettit Scholarship in Science
2015	Lynn Bateman Memorial Teaching Award For excellence in teaching and contribution to the UC Denver Math department
2012	Inducted into Pi Mu Epsilon Mathematical Honor Society

Service and Outreach

Department Service

Fall 2024	Statistics Seminar organizer
2023-24	PhD admission committee

Campus-wide Service

Winter 2024	UCI Data Science Initiative Statistical and Machine Learning Applications in Biomedical Sciences workshop organizer
-------------	---

Professional Service

2024-28	COPSS Florence Nightingale David Award Committee Member, ASA
2024	NSF Panel Member
Organizer/chair	
2024	Interpretable Machine Learning for Genomics and Biomedical Problems, JSM
2024	Veridical Data Science in Medicine, Berkeley-Stanford Workshop on VDS
2023	Optimization for Statistical Learning, CMStatistics
2023	Thinking deeply about sequencing and omics data, WSDS
2019	The annual Big Data to Knowledge Retreat held at the Penn State University
2017	The 2017 Rao Prize Conference held in the Statistics Department at Penn State

Outreach Service

2024-Current	Faculty mentor for Math Alliance
2021-Current	Alumni mentor to members of the Sloan Minority PhD program
2022 & 2023	Judged submissions for the ENVISION Research Competition held by the Women in STEM
2022	DEIBlueprint Committee member for the UC Berkeley Statistics Department
2020-2021	Graduate Educational Equity Programs Advisory group member
2020-2021	Student Life Committee member for the Faculty Senate at Penn State
2020	Mentor for prospective Penn State Millennium Scholars in Math and Statistics

Review Service

Annals of Applied Statistics (2)
Australian & New Zealand Journal of Statistics (1)
Computational Statistics and Data Analysis (2)
Genome Research (2)
Frontiers in Oncology (1)
The Journal of the American Statistical Association (2)
Journal of Econometrics & Statistics (1)
Nature Communications (1)
Statistics (1)
Pediatric Obesity (1)

Student Advising at UC Irvine

Research advising

Paul Rogelio Gaona Partida (PhD)	Expected graduation 2028
Rasul Ibragimov (PhD)	Expected graduation 2028
Asha Pantula (MS)	Expected graduation 2025
Nick Zuniga (BS)	Expected graduation 2026
Shuyi Chen (BS)	Expected graduation 2025
Veronica Dahlkamp (BA)	Expected graduation 2024

Academic advising

Noah Benjamin (PhD)	Expected graduation 2029
Yiqun Quan (MS)	Expected graduation 2026
Yiqun Liu (MS)	Expected graduation 2025

Teaching Experience

Instructor (Full responsibility)

STAT 211/201, **Statistical Methods for Data Analysis 2** (Winter 2024, two sections), UC Irvine, Introduction to statistical methods for analyzing data from surveys or experiments. Emphasizes application and understanding of methods for categorical data including contingency tables, logistic and Poisson regression, loglinear models.

STAT 299, **Independent Study** (Fall 2023, Spring 2024), UC Irvine, Feature importance and screening methods for interpretable machine learning (Fall 2023). Longitudinal analysis methods for wearable devices to improve model predictions of women's menstrual cycles (Spring 2024).

MATH 2830, **Introductory Statistics** (Fall 2014, Spring 2015), UC Denver, Basic statistical concepts, summarizing data, probability concepts, distributions, confidence intervals, hypothesis testing.

MATH 1110, **College Algebra** (Summer 2014), UC Denver, Topics in algebra designed for students who intend to take the calculus sequence. Functions, domains, ranges, graphs, data scatter plots and curve fitting, solving equations and systems of equations, polynomial functions, rational functions, and selected other topics.

MATH 1080, **Polynomial Calculus** (Spring 2014), UC Denver, A one-semester course in single-variable calculus. Topics include limits, derivatives, differentiation rules, integration and integration rules. Emphasis is on applications to business and social sciences.

MATH 1070, **Algebra for Social Sciences and Business** (Fall 2013, Summer 2013), UC Denver, Topics in algebra designed for students who intend to take business calculus. Functions, graphs, scatter plots, curve-fitting, solving systems of equations, polynomial and rational functions, and selected other topics.

Guest Lecturer

MCIBS 556, **Computation, Bioinformatics**, and Statistics Practicum (Fall 2019), Pennsylvania State University, A required graduate course for PhD in Bioinformatics & Genomics. Taught an introduction to modern feature selection approaches including pros/cons and available software.

MCIBS 556, **Computation, Bioinformatics, and Statistics Practicum** (Fall 2018), Pennsylvania State University, A required graduate course for PhD in Bioinformatics & Genomics. Taught an introduction to feature selection using the L_0 norm and mixed integer programming.

Teaching Assistant

STAT 500, **Applied Statistics** (Fall 2016), Pennsylvania State University, Descriptive statistics, hypothesis testing, power, estimation, confidence intervals, regression, one- and 2-way ANOVA, Chi-square tests, diagnostics (graduate course).

STAT 200, **Elementary Statistics** (Spring 2017), Pennsylvania State University, Descriptive statistics, frequency distributions, probability, binomial and normal distributions, statistical inference, linear regression, and correlation (undergraduate).

MATH 1110, **College Algebra** (Fall 2012/Spring 2013), UC Denver, Topics in algebra designed for students who intend to take the calculus sequence. Functions, domains, ranges, graphs, data scatter plots and curve fitting, solving equations and systems of equations, polynomial functions, rational functions, and selected other topics.

MATH 1070, **College Algebra** (Spring 2012), CSU Stanislaus, Review of basic algebraic topics and basic analytic geometry, complex numbers, functional notation, graphs, polynomials and rational functions, exponential and logarithmic functions, systems of equations, and conic sections.

Memberships in Professional Societies and Organizations

ASA - American Statistical Association

IMS - Institute of Mathematical Statistics

INFORMS - Institute for Operations Research and the Management Sciences

Mathematical Alliance

Pi Mu Epsilon - Mathematical Honor Society

SACNAS - Society for Advancement of Chicanos/Hispanics and Native Americans in Science

SSMN - Sloan Scholars Mentoring Network

Software

imodels: Interpretable ML package for concise, transparent, and accurate predictive modeling. Available on GitHub

vdocs: An R package for seamless documentation of data analyses via lab notebooks to encourage transparent and reliable data science. Available on GitHub

Patents

MyProstateScore2.0: A simplified score for high-grade prostate cancer. Co-inventor with Dr. Tiffany Tang (Notre Dame), Dr. Yuping Zhang (University of Michigan), Professor Bin Yu (UC Berkeley), and Professor Arul Chinnaiyan (University of Michigan).

Invited Talks/Discussions

1. *Feature selection and outlier detection using Mixed Integer Programming Under Complex Data Structures*, **ECDA**, September 2024, Sopot, Poland
2. *Distilling causal effects: stable subgroup estimation via distillation trees in causal inference*, **JSM**, August 2024, Portland, OR
3. *Interpreting models in biomedical research to improve feature recommendations*, **CHOC UCI Pediatric Data Science Seminar**, August 2024, Irvine, CA
4. *Interpretable machine learning for biomedical research*, **Computational and Data Sciences Research Team at CHOC**, June 2024, Irvine, CA
5. *Statistical learning in genomics for biomarker discovery*, **OCB ASA Quarter 2 Meeting**, May 2024, Irvine, CA
6. *Discussant for Veridical Data Science in Medicine*, **Berkeley-Stanford Workshop on VDS**, May 2024, Berkeley, CA
7. *Introduction to my lab*, **Statistical and Machine Learning Applications in Biomedical Sciences UCI Data Science Initiative workshop**, February 2024, Irvine, CA
8. *Interpreting models in biomedical research to improve feature recommendations*, **BrownBag Seminar UC Irvine**, November 2023, Irvine, CA
9. *MDI+: a flexible random forest-based feature importance framework*, **CMStatistics**, December 2023, Berlin, Germany
10. *Leveraging problem structure to improve feature recommendations in biomedical research*, **Applied Research and Education Seminar (ARES)**, November 2023, University of Toronto
11. *A metabolomic based panel for pancreatic cancer*, **13th EDRN Scientific Workshop**, September 2023, Washington DC
12. *Simultaneous Feature Selection and Outlier Detection Using Mixed-Integer Programming with Optimality Guarantees*, **Data Science Institute Seminar Series**, October 2022, Virtual
13. *Dynamic Biomarker Discovery for Childhood Obesity Using Functional Data Analysis*, **WNAR**, June 2022, Virtual
14. *Dynamic Biomarker Discovery for Childhood Obesity Using Functional Data Analysis* (talk), **Bio-statistics Seminar**, March 2022, Berkeley California
15. *Finding Biomarkers for Childhood Obesity Using Functional Data Analysis*, **Women in Data Science (WiDS) Berkeley**, March 2022, Virtual
16. *Confronting Challenges in Omics Research Through Functional Data Analysis and Mixed Integer Programming*, **Yu Group at UC Berkeley**, November 2020, Virtual

17. *Finding Biomarkers for Childhood Obesity Using Functional Data Analysis*, **NIST Statistical Engineering Division Seminar**, November 2020, Virtual
18. *Finding Biomarkers for Childhood Obesity Using Functional Data Analysis*, **CMStatistics**, December 2019, London England
19. *Finding Biomarkers for Childhood Obesity Using Functional Data Analysis*, **ENAR**, March 2019, Philadelphia Pennsylvania
20. *Statistical Disclosure Control for Functional PCA*, **CMStatistics**, December 2018, Pisa Italy
21. *Computationally Viable and Effective Feature Selection in L_0 Norm*, **International Conference on Advances in Interdisciplinary Statistics and Combinatorics (AISC)**, October 2018, Greensboro North Carolina

Contributed Talks

1. *Distilling causal effects: stable subgroup estimation via distillation trees in causal inference* (poster), **American Causal Inference Conference (ACIC)**, May 2024, Seattle WA
2. *Properties and Applications of Feature Whitening* (talk), **Joint Statistical Meetings (JSM)**, August 2022, Washington D.C.
3. *Unsupervised Feature Decorrelation for Variable Selection* (talk), **Joint Statistical Meetings (JSM)**, August 2021, Virtual
4. *Constructing a Polygenic Risk Score for Childhood Obesity Using Functional Data Analysis* (poster), **CSHL Biology of Genomes**, May 2021, Virtual
5. *Simultaneous Outlier Detection and Feature Selection Using Mixed-Integer Programming* (talk), **Society for Advancement of Chicanos/Hispanics & Native Americans in Science (SACNAS) Conference**, October 2020, Virtual
6. *Simultaneous Outlier Detection and Feature Selection Using Mixed-Integer Programming* (talk), **Women in Statistics and Data Science (WSDS)**, October 2020, Virtual
7. *Simultaneous Outlier Detection and Feature Selection Using Mixed-Integer Programming* (talk), **Joint Statistical Meetings (JSM)**, August 2020, Virtual
8. *Feature De-Correlation for Variable Selection Using a Constrained Whitening Approach* (talk), **Women in Statistics and Data Science (WSDS)**, October 2019, Bellevue Washington
9. *Feature Selection in L_0 Norm: A Viable Approach* (talk), **Joint Statistical Meetings (JSM)**, August 2018, Vancouver Canada
10. *Feature Selection in L_0 Norm: A Viable Approach* (talk), **4th Conference of the International Society for Nonparametric Statistics (ISNPS)**, June 2018, Salerno Italy
11. *Feature Selection in L_0 Norm: A Viable Approach* (talk), **Symposium on Data Science and Statistics (SDSS)**, May 2018, Reston Virginia
12. *Efficient Bounds for Best Subset Selection using Mixed Integer Linear Programming* (poster session and speed talk), **Women in Statistics and Data Science Conference (WSDS)**, October 2017, La Jolla California
13. *Efficient Bounds for Best Subset Selection using Mixed Integer Linear Programming* (talk), **Joint Statistical Meetings (JSM)**, August 2017, Baltimore Maryland.
14. *Predicting Random and Rather Erratic Road Conditions* (talk), **SIAM Front Range Student Conference**, February 2015, Denver Colorado.