

# Kevin Z. Lin

University of Washington, Department of Biostatistics

Address: Hans Rosling Center for Population Health, 3980 15th Avenue NE, Seattle, WA 98195

Mobile: (206) 685-8334

Email : kzlin@uw.edu

Website: <https://linnykos.github.io/>

CV last updated: January 27, 2024

## EDUCATION

---

- **Princeton University** Princeton, NJ  
B.S.E. in Operations Research & Financial Engineering 2010 - 2014  
Advisors: Han Liu and Robert Vanderbei  
Certificates in “Statistics and Machine Learning” and “Applications of Computing”
- **Carnegie Mellon University** Pittsburgh, PA  
Ph.D. in Dietrich’s Statistics & Data Science, Masters in Machine Learning 2014 - 2020  
Thesis advisors: Kathryn Roeder and Jing Lei

## EMPLOYMENT

---

- **University of Pennsylvania** Philadelphia, PA  
Post-doctoral researcher in Wharton’s Statistics & Data Science 2020 - 2023  
Advisor: Nancy Zhang
- **University of Washington** Seattle, UW  
Assistant professor (tenure-track) in School of Public Health’s Biostatistics 2023 - Present

## HONORS, AWARDS, AND SCHOLARSHIPS

---

- **Teaching assistant excellence award recipient** Carnegie Mellon University  
For Course 36-350: “Statistical Computing” in Fall 2017 May 2017
- **Winner of Statistical excellence for early-career writing** Significance Magazine  
For article “We, the millennials: The statistical significance of political significance” June 2017  
*Competition held jointly with the Young Statisticians Section of Royal Statistical Society*
- **Honorable mention in student paper competition** American Statistical Association  
For paper “Dependency diagnostic: Visually understanding pairwise variable relationships” January 2018  
*For ASA section: Statistical Computing and Statistical Graphics*
- **PhD TAs of the year** Carnegie Mellon University  
For the Spring 2020 semester (1 of 2 total recipients) May 2020
- **Wikimedia Foundation Research Award of the Year** (Remote)  
For paper “Controlled analyses of social biases in Wikipedia bios” 2023
- **Biometrics Excellent Referee Awards** Atlanta, GA  
For the excellent reviews for the journal Biometrics (1 of 3 total recipients) 2024

## PROFESSIONAL ACTIVITIES

---

- Association member for: American Statistical Association; American Society of Human Genetics; International Chinese Statistical Association
- Peer-Reviewer for: Annals of Applied Statistics; Annals of Statistics; Bayesian Analysis; Biometrics; Biometrika; Electronic Journal of Statistics; IEEE Transactions of Network Science and Engineering; Information and Inference: A Journal of the IMA; Journal of American Statistical Association; Journal of Molecular Biology; Journal of the Royal Statistical Society: Series B; Nature Neuroscience; PLOS Genetics; Statistical Sinica; Statistics and Probability Letters; Statistics in Medicine; Technometrics
- Reviewer for Student Paper Award for the American Statistical Association's Section on Statistics in Genomics and Genetics (January 2023)

## BIBLIOGRAPHY

---

**Bold** denotes PI and his mentored students. Items are ordered in chronological order.

### Refereed research articles

1. Vanderbei, R., **Lin, K. Z.**, Liu, H., and Wang, L. (2016). Revisiting compressed sensing: Exploiting the efficiency of simplex and sparsification methods. *Mathematical Programming Computation*, 8(3):253–269  
[DOI: 10.1007/s12532-016-0105-y](https://doi.org/10.1007/s12532-016-0105-y)
2. **Lin, K. Z.**, Sharpnack, J., Rinaldo, A., and Tibshirani, R. J. (2017). A sharp error analysis for the fused lasso, with application to approximate changepoint screening. In *Advances in Neural Information Processing Systems*, pages 6884–6893  
[DOI: 10.5555/3295222.3295432](https://doi.org/10.5555/3295222.3295432)
3. An, J.-Y., **Lin, K. Z.**, Zhu, L., Werling, D. M., Dong, S., Brand, H., Wang, H. Z., Zhao, X., Schwartz, G. B., Collins, R. L., Currall, B. B., Dastmalchi, C., Dea, J., Duhn, C., Gilson, M. C., Klei, L., Liang, L., Markenscoff-Papadimitriou, E., Pochareddy, S., Ahituv, N., Buxbaum, J. D., Coon, H., Daly, M. J., Shin Kim, Y., Marth, G. T., Neale, B. M., Quinlan, A. R., Rubenstein, J. L., Sestan, N., State, M. W., Willsey, A. J., Talkowski, M. E., Devlin, B., Roeder, K., and Sanders, S. J. (2018). Genome-wide de novo risk score implicates promoter variation in autism spectrum disorder. *Science*, 362(6420)  
[DOI: 10.1126/science.aat6576](https://doi.org/10.1126/science.aat6576)
4. Lei, J. and **Lin, K. Z.** (2020). Discussion of ‘Network cross-validation by edge sampling’. *Biometrika*, 107(2):285–287  
[DOI: 10.1093/biomet/asaa009](https://doi.org/10.1093/biomet/asaa009)
5. Wang, D., Zhao, Z., **Lin, K. Z.**, and Willett, R. (2021). Statistically and computationally efficient changepoint localization in regression settings. *Journal of Machine Learning Research*, 22:248–1  
[DOI: 10.5555/3546258.3546506](https://doi.org/10.5555/3546258.3546506)
6. Hyun, S., **Lin, K. Z.**, G'Sell, M., and Tibshirani, R. J. (2021). Post-selection inference for changepoint detection algorithms with application to copy number variation data. *Biometrics*, 77(3):1037–1049  
[DOI: 10.1111/biom.13422](https://doi.org/10.1111/biom.13422)

7. **Lin, K. Z.**, Liu, H., and Roeder, K. (2021b). Covariance-based sample selection for heterogeneous data: Applications to gene expression and autism risk gene detection. *Journal of the American Statistical Association*, 116(533):54–67  
DOI: 10.1080/01621459.2020.1738234
8. **Lin, K. Z.**, Lei, J., and Roeder, K. (2021a). Exponential-family embedding with application to cell developmental trajectories for single-cell RNA-seq data. *Journal of the American Statistical Association*, 116(534):457–470  
DOI: 10.1080/01621459.2021.1886106
9. Field, A., Park, C. Y., **Lin, K. Z.**, and Tsvetkov, Y. (2022). Controlled analyses of social biases in Wikipedia bios. In *Proceedings of the ACM Web Conference 2022*, pages 2624–2635  
DOI: 10.1145/3485447.3512134
10. Lei, J. and **Lin, K. Z.** (2022). Bias-adjusted spectral clustering in multi-layer stochastic block models. *Journal of the American Statistical Association*, pages 1–13  
DOI: 10.1080/01621459.2022.2054817
11. Guan, P. Y., Lee, J. S., Wang, L., **Lin, K. Z.**, Mei, W., Chen, L., and Jiang, Y. (2023). Destin2: Integrative and cross-modality analysis of single-cell chromatin accessibility data. *Frontiers in Genetics*, 14  
DOI: 10.3389/fgene.2023.1089936
12. **Lin, K. Z.** and Zhang, N. R. (2023). Quantifying common and distinct information in single-cell multimodal data with Tilted Canonical Correlation Analysis. *Proceedings of the National Academy of Sciences*, 120(32)  
DOI: 10.1073/pnas.2303647120
13. Chen, S., Zhu, B., Huang, S., Hickey, J. W., **Lin, K. Z.**, Snyder, M., Greenleaf, W. J., Nolan, G. P., Zhang, N. R., and Ma, Z. (2023). Integration of spatial and single-cell data across modalities with weakly linked features. *Nature Biotechnology*, pages 1–11  
DOI: 10.1038/s41587-023-01935-0

#### Other refereed scholarly publications

1. **Lin, K. Z.** (2017). We, the millennials: The statistical significance of political significance. *Significance*, 14(5):28–33  
DOI: 10.1111/j.1740-9713.2017.01073.x

#### Other non-refereed published scholarly publications

1. **Lin, K. Z.**, Qiu, Y., and Roeder, K. (2023). eSVD-DE: Cohort-wide differential expression in single-cell RNA-seq data using exponential-family embeddings. *bioRxiv*, pages 2023–11  
Biorxiv: 10.1101/2023.11.22.568369

---

#### PATENTS AND OTHER INTELLECTUAL PROPERTY

Below are publicly available software.

1. **covarianceSelection**: R package for selecting datasets with similar high-dimensional covariance matrices. (<https://github.com/linnykos/covarianceSelection>)
2. **selectiveModel**: R package for selective inference via binary segmentation for changepoint detection. (<https://github.com/linnykos/selectiveModel>, developed jointly with Sangwon Hyun)

3. **eSVD**: R package for performing dimension reduction for matrices with respect to an exponential-family distribution. (<https://github.com/linnykos/esvd>)
4. **networkSoSD**: R package for performing spectral clustering to estimate node communities based on a collection of networks. (<https://github.com/linnykos/networkSoSD>)
5. **tiltedCCA**: R package for estimating the shared and distinct geometry in paired multimodal data. (<https://github.com/linnykos/tiltedCCA>)
6. **eSVD2**: R package that extends the functionality of the previous eSVD package, and also implements specialized functions to perform subject-level differential expression for single-cell data. (<https://github.com/linnykos/eSVD2>)

## CONFERENCES AND SYMPOSIUMS

---

“id.” denotes identical title as the item above. Items are grouped by talks, and then ordered by chronological order.

### Invited departmental seminars

1. **University of California, Los Angeles: Department of Statistics** (12/2021).  
Exponential-family embedding with application to cell developmental trajectories for single-cell data
2. **University of California, Irvine: Department of Statistics** (12/2022). Tilted-CCA:  
Quantifying common and distinct information in multiomic single-cell data
3. **University of Washington: Department of Biostatistics** (01/2023). id.
4. **University of Chicago: Department of Statistics** (01/2023). id.
5. **University of California, Riverside: Department of Statistics** (01/2023). id.
6. **University of Waterloo: Department of Statistics** (01/2023). id.
7. **Pennsylvania State University: Department of Statistics** (01/2023). id.
8. **University of North Carolina: Department of Biostatistics** (01/2023). id.
9. **University of Michigan: Department of Statistics** (01/2023). id.
10. **University of Michigan: Department of Biostatistics** (02/2023). id.
11. **University of Illinois, Urbana-Champaign: Department of Statistics** (02/2023). id.
12. **Emory University: Department of Biostatistics** (02/2023). id.
13. **McGill University: Department of Epidemiology, Biostatistics, and Occupational Health** (09/2023). id.
14. **University of California, Santa Cruz: Department of Statistics** (10/2023). id.

## Conference/Seminar presentations

Note: (\*) denotes invited talks.

1. Modeling and Optimization: Theory and Applications (08/2014, Bethlehem, PA). Optimization for compressed sensing: New insights and alternatives
2. American Society of Human Genetics (10/2015, Baltimore, MD). Gaussian graphical model integrating microarray and sequencing data for autism risk gene detection
3. Joint Statistical Meeting (08/2016, Chicago, IL). id.
4. Conference on Neural Information Processing Systems (12/2017, Long Beach, CA). A sharp error analysis for the fused lasso, with application to approximate changepoint screening
5. Joint Statistical Meeting (08/2017, Baltimore, MD). Hypothesis testing for simultaneous variable clustering and correlation network estimation
6. Joint Statistical Meeting (08/2018, Vancouver, Canada). Dependency diagnostic: Visually understanding pairwise variable relationships
7. Joint Statistical Meeting (08/2019, Denver, CO). Exponential-family embedding with application to cell developmental trajectories for single-cell data
8. **Joint Statistical Meeting\*** (08/2020, Remote). id.
9. American Society of Human Genetics (10/2020, Remote). id.
10. Joint Statistical Meeting (08/2020, Remote). Time-varying stochastic block models via kernel smoothing, with application to RNA-seq data
11. **StatScale Seminar\*** (07/2021, Remote). id.
12. Symposium on Data Science and Statistics (06/2022, Pittsburgh, PA). Spectral clustering for multi-layer stochastic block models: Analysis of dynamic heterophilic networks
13. Joint Statistical Meetings (08/2022, Washington DC). Tilted-CCA: Quantifying common and distinct information in multiomic single-cell data
14. Institute of Mathematical Statistics New Researchers Conference 2023 (08/2023, Toronto, Canada). id.
15. **Scripps Research: Computational Biology and Bioinformatics Seminar\*** (08/2023, Remote). id.
16. Cell Symposia: The conceptual power of single-cell biology (08/2023, San Diego, CA). id.
17. American Society of Human Genetics (11/2023, Washington DC). id.
18. Keystone Symposia: Regulatory RNA - Emerging Mechanisms (12/2023, Banff, Canada). id.
19. Joint Statistical Meeting (08/2023, Toronto, Canada). eSVD: Cohort-level differential expression in single-cell data via matrix factorization
20. **UW Genome Science: Combi Seminar\*** (01/2024, Seattle, WA). Single-cell paired RNA & ATAC: Surveying broad multi-modal coordination in development and cancer resistance

## UNIVERSITY SERVICE

---

- UW Department of Biostatistics: Equity, Diversity, and Inclusion; Committee member (2023)
  - Lead coordinator for the Wellness working group subdivision (Winter 2023)
- UW School of Public Health: Equity, Diversity, Inclusion; Committee member, representing the Department of Biostatistics (2023)
- UW Department of Biostatistics: Teaching track assistant professor search; Committee member (2023)
- UW School of Public Health: Resume Clinic; Reviewer (Autumn 2023)
- UW School of Public Health: Mock Interview; Interviewer (Winter 2023)

## TEACHING HISTORY

---

### **At Princeton University**

1. Course designer for ORF 350: Analysis of Big Data (Spring 2012, Spring 2013, Spring 2014)

### **At Carnegie Mellon University**

1. Teaching assistant for 46-921 & 46-923: Financial Data Analysis I and II (Spring 2014)
2. Teaching assistant for 36-217: Probability Theory and Random Processes (Spring 2015)
3. Teaching assistant for 36-350: Statistical Computing (Fall 2015, Fall 2016, Fall 2017)
4. Guest lecturer for 36-750: Statistical Computing (Fall 2016, Fall 2017, Fall 2018, Fall 2019, Fall 2020)
5. Assistant instructor for 36-350: Statistical Computing (Spring 2018)
6. Instructor for 36-350: Statistical Computing (Summer 2018)
7. Data science initiative project fellow for 36-490: Undergraduate Research (Spring 2019)
8. Co-instructor for 36-469: Statistical Genomics and High-Dimensional Inference (Spring 2020).  
With Kathryn Roeder

### **At University of Washington**

1. Instructor for BIOST 582: Student Seminar (Autumn 2023)

### **Independent study supervision**

1. Yuhong Li (2023), University of Washington, Department of Biostatistics, Master of Science - Thesis
2. Zhaoheng Li (2023), University of Washington, Department of Biostatistics, PhD candidate
3. Yimin Zhao (2023), University of Washington, Department of Biostatistics, PhD candidate
4. Wenjing (Tati) Zhang (2023), University of Washington, Department of Biostatistics, Master of Science - Thesis

## Teaching and other professional development

- Certified by CMU's Eberly Center's Future Faculty Program, which included two observed lectures in two different semesters (Fall 2019 to Summer 2020)
- Certified with Gatekeeper certificate by the QPR's (Question, Persuade, Refer) suicide prevention program (February 2020)
- Certified by Mental Health First Aid USA (Fall 2020)

## ADVISING AND FORMAL MENTORING

---

### MS and PhD committees in non-chair role

1. Doctoral Committee Member
  - (a) Connor Finkbeiner (2023), University of Washington, Department of Genome Sciences, PhD candidate. (Chair: Manu Setty)
2. Qualifying Exam Member
  - (a) Anna Spiro (2023), University of Washington, Department of Computer Science, PhD candidate. (Chair: Sara Mostafavi)

### Other mentoring

Project Supervision (outside of independent study supervision)

1. Zeyu (Jerry) Wei (2023), University of Washington, Department of Statistics, PhD (on project jointly with Yen-Chi Chen)
2. Wenbin (Derek) Wu (2023), University of Washington, Department of Statistics, Master of Science - Advanced Methods and Data Analysis
3. Haoye Yang (2023), University of Washington, Department of Statistics, Master of Science - Advanced Methods and Data Analysis
4. Shizhao (Joshua) Yang (2023), University of Washington, Department of Biostatistics, Master of Science - Capstone

### Department-assigned Academic Advisor

1. Alejandro Hernandez (2023), University of Washington, Department of Biostatistics, Master of Science - Capstone
2. Yutong Jin (2023), University of Washington, Department of Biostatistics, Master of Science - Capstone