Yoonkyung Lee

Department of Statistics
The Ohio State University
1958 Neil Avenue
Columbus, OH 43210

yklee@stat.osu.edu
http://www.stat.osu.edu/~yklee
614-292-9495 (office)
614-292-2096 (fax)

Education

2002 Ph.D. in Statistics, University of Wisconsin-Madison
Dissertation: *Multicategory Support Vector Machines, Theory, and Application to the Classification of Microarray Data and Satellite Radiance Data*
Advisor: Grace Wahba, Ph.D. (Co-advisor: Yi Lin, Ph.D.)

1996 M.Sc. in Statistics, Seoul National University, Korea
Thesis: *Image Data Analysis by Markov Random Field Models*
Advisor: Jong Woo Jeon, Ph.D.

1994 B.Sc. in Computer Science and Statistics, Seoul National University, Korea

Professional Experience

Positions

June 2016–present	Professor, Department of Statistics, The Ohio State University Professor, Department of Computer Science and Engineering, The Ohio State University (courtesy appointment)
July 2020–June 2022	Program Co-Director, Translational Data Analytics Institute
Sep–Oct 2019	Visiting Professor, Department of Mathematical Sciences, KAIST, Korea
Feb–June 2016	Associate Professor, Department of Computer Science and Engineering, The Ohio State University (courtesy appointment)
2008–2016	Associate Professor, Department of Statistics, The Ohio State University
Spring 2011	Visiting Associate Professor, Department of Statistics, Seoul National University, Korea
2002–2008	Assistant Professor, Department of Statistics, The Ohio State University
Jan–Mar 2007	Visiting Fellow, Statistical and Applied Mathematical Sciences Institute, Research Triangle Park, NC
1999–2002	Research Assistant, Department of Statistics, UW-Madison Supervisor: Grace Wahba, Ph.D.
Spring 1999	Teaching Assistant, Department of Statistics, UW-Madison
Fall 1998	Research Assistant, Department of Statistics, UW-Madison Supervisor: Michael Newton, Ph.D.
1997–1998	Teaching Assistant, Department of Statistics, UW-Madison

Teaching

STAT 3202	Introduction to Statistical Inference for Data Analytics
STAT 4201	Introduction to Mathematical Statistics I
STAT 428	Introduction to Probability and Statistics for Engineering and the Sciences
STAT 5301–5302	Intermediate Data Analysis I and II
STAT 621	Statistical Theory II
STAT 6500	Statistical Machine Learning
STAT 6560	Applied Multivariate Analysis
STAT 7410	Theory of the Linear Model
STAT 7620	Elements of Statistical Learning
STAT 7630	Nonparametric Function Estimation
STAT 881	Advanced Statistical Learning
STAT 882	Topics in Variable Selection and Model Selection

Research Supervision

- Doctoral Students
  - Xuerong Wang (current)
  - Haozhen Yu (current)
  - Zhenbang Jiao (2025), *Case Sensitivity in Lasso Regression: Diagnostics, Model Complexity, and Robust Model Evaluation.*
  - Lun Li (2024), *A Quasi-Likelihood Approach to Latent Space Modeling for Compositional Data: Computation, Model Diagnostics, and Applications.*
  - Ruochen Huang (2023), *Enhancing Exponential Family PCA: Statistical Issues and Remedies.*
  - Bo Luan (2022), co-advised with Yunzhang Zhu
    Model Complexity in Linear Regression: Extensions for Prediction and Heteroscedasticity.
  - Jianhao Zhang (2021), *Learning from Binary Matrix and Tensor Data with Sparsity.*
  - Jiae Kim (2020), co-advised with Steve MacEachern
    Nonlinear Generalizations of Linear Discriminant Analysis: the Geometry of the Common Variance Space and Kernel Discriminant Analysis.
  - Shanshan Tu (2019), co-advised with Yunzhang Zhu
    Case Influence and Model Complexity in Regression and Classification.
  - Tayler Blake (2018), *Nonparametric Covariance Estimation for Longitudinal Data.*
  - Liubo Li (2017), co-advised with Vince Vu
    Trend-Filtered Projections for Principal Component Analysis.
  - Jieyi Jiang (2017), co-advised with Steve MacEachern
    Realistic Predictive Risk: The Role of Penalty and Covariate Diffusion in Model Selection.
  - Andrew Landgraf (2015), *Generalized Principal Component Analysis: Dimensionality Reduction through the Projection of Natural Parameters.*

- John Lewis (2014), co-advised with Steve MacEachern
  Bayesian Restricted Likelihood Methods.
- Zhiyu Liang (2014), *Eigen-Analysis of Kernel Operators for Nonlinear Dimension Reduction and Discrimination.*
- Kazuki Uematsu (2012), *Statistical Consistency of Ranking: Bipartite and Multipartite Cases.*
- Rui Wang (2012), *Comparisons of Classification Methods in Efficiency and Robustness.*
- Cong Liu (2012), co-advised with Tao Shi
  Two Tales of Variable Selection for High Dimensional Data: Screening and Model Building.
- Yoonsuh Jung (2010), co-advised with Steve MacEachern
  Regularization of Case Specific Parameters: A New Approach for Improving Robustness and/or Efficiency of Statistical Methods.
- Youlan Rao (2009), co-advised with Jason Hsu
  Statistical Analysis of Microarray Experiments in Pharmacogenomics.
- Yonggang Yao (2008), *Statistical Applications of Linear Programming for Feature Selection via Regularization Methods.*
- Zhenhuan Cui (2007), *The Solution Paths of Multicategory Support Vector Machines: Algorithm and Applications.*

- Master Students
  - Steve Dignan (2024), *A Comparison of Logistic PCA and Selected Data Embedding Procedures for Binary Data with Application to Breast Cancer and Glioblastoma Data*

Honors and Awards

2023	Thomas and Jean Powers Teaching Award, Department of Statistics, OSU
2022	Lindley Prize honorable mention for the paper by Lewis, MacEachern, and Lee (2021)
2015	Elected Fellow of the American Statistical Association
2008	NSF travel award, The 7th World Congress in Probability and Statistics, Singapore, July 2008
2007	Statistical and Applied Mathematical Sciences Institute Fellowship
2003	NSF travel award, SRCOS/ASA Summer Research Conference, Jekyll Island, Georgia, June 2003
2003	Best student poster award, 12th Conference on Satellite Meteorology and Oceanography, American Meteorological Society Meeting 2003
2002	Student travel grant, IMS Mini-Meeting, University of Florida, January 2002
1996	Master's degree with honors, The Graduate School, Seoul National University
1995	Qualifying Examination passed with distinction, Department of Statistics, Seoul National University
1994	Dean's list upon graduation, College of Natural Sciences, Seoul National University
1991–1994	Seoul National University Alumni Fellowship

Grants

- *Uncovering a Conserved Role of Cell Polarity Signaling in Cellular Aging Using Budding Yeast and Hematopoietic Stem Cell Models*, President's Research Excellence (PRE) Accelerator Award, OSU, as a co-principal investigator (2023–2024). PIs: Hay-Oak Park (Molecular Genetics) and Brad Blaser (Medicine)
- *Model Evaluation in Modern Predictive Regimes: Case Influence and Model Complexity*, National Science Foundation, Division of Mathematical Sciences, Statistics Program as a principal investigator (2020–2024).
- *Nonlinear Dimension Reduction Methods*, National Science Foundation, Division of Mathematical Sciences, Statistics Program as a sole principal investigator (2015–2019).
- *Robust and Relevant Model Evaluation: Principles and techniques for handling weak prior information and contaminated data*, National Science Foundation, Division of Mathematical Sciences, Statistics Program as a co-principal investigator (2012–2016).
- *Robust and Relevant Model Evaluation*, Nationwide Center for Advanced Customer Insights as a co-principal investigator (2011–2012).
- *Blended Paradigm Inference with Application to Regression*, National Security Agency, Mathematical Sciences Program as a co-principal investigator (2010–2012).
- Conference on *Nonparametric Statistics and Statistical Learning*, National Science Foundation, Division of Mathematical Sciences, Statistics Program as a co-principal investigator (2009–2010).

Invited Talks

- Case Sensitivity in Regression and Beyond,
  Department of Statistics, Iowa State University, Ames, IA, March 10, 2025.
- Measuring Case Influence Locally,
  Joint Statistical Meetings, Portland, OR, August 3-8, 2024.
- Case Sensitivity in Regression and Beyond,
  Department of Statistics and Probability, Michigan State University, East Lansing, MI, March 14, 2024.
- Predictive Model Degrees of Freedom in Linear Regression,
  – The Korean Statistical Society Summer Conference, Busan, Korea, July 1, 2023
  – The Symposium on Statistics and Risk Management 2022, The Chinese University of Hong Kong, Hong Kong, December 9, 2022.
  – Department of Statistics, Florida State University, Tallahassee, FL, November 4, 2022.

- Assessment of Case Influence in Support Vector Machine,
  – CIMAT (Center of Research in Mathematics), Guanajuato, Mexico, November 25, 2020.
  – Department of Mathematics, University of Arizona, November 16, 2020.
  – The 12th International Conference of the ERCIM WG on Computational and Methodological Statistics, University of London, UK, December 14–16, 2019.
  – Department of Mathematical Sciences, KAIST, Daejeon, Korea, September 26, 2019.

- Statistical Learning with Kernels (Part I: Smoothing Splines, Part II: Support Vector Machines, Part III: Kernel Methods), Department of Mathematical Sciences, KAIST, Daejeon, Korea, September 30–October 8, 2019.

- Supervised Dimensionality Reduction for Exponential Family Data,
  – The 11th International Conference of the ERCIM WG on Computational and Methodological Statistics, University of Pisa, Italy, December 14–16, 2018.
  – Joint Statistical Meetings, Vancouver, Canada, July 28–August 2, 2018.
  – The Conference on Statistical Learning and Data Science, Columbia University, New York, June 4–6, 2018.

- Interview with Grace Wahba, The Pfizer Distinguished Statistician Colloquium Series, University of Connecticut, Storrs, CT, September 26–27, 2018.

- Dimensionality Reduction for Exponential Family Data,
  Computational Strategies for Large-Scale Statistical Data Analysis Workshop, International Centre for Mathematical Sciences, Edinburgh, UK, July 2–6, 2018.

- Cross Validation for Penalized Quantile Regression with a Case-Weight Adjusted Solution Path, The 2nd International Conference on Econometrics and Statistics, City University of Hong Kong, Hong Kong, June 19–21, 2018.

- The Geometry of Nonlinear Embeddings for Kernel Principal Component Analysis and Discriminant Analysis,
  TGDA@OSU TRIPODS Center Summer School, Mathematical Biosciences Institute, Columbus, OH, May 2018.

- A Characterization of Latent Factors for Binary Data with Application to a Study of the Comorbidity in ICU Patients,
  The Spring Conference of the Korean Statistical Society, Seoul, Korea, May 2017.

- The Remarkable Effectiveness of Principal Component Analysis,
  Korean Engineering Graduate Student Association Meeting, The Ohio State University, Columbus, OH, November 2016.

- Generalized Principal Component Analysis: Dimensionality Reduction through the Projection of Natural Parameters,
  – Data Science Research Group, School of Mathematical Sciences, Rochester Institute of Technology, NY, November 2017.
  – Department of Statistics, Purdue University, West Lafayette, IN, October 2017.
  – ICSA Applied Statistics Symposium, Chicago, IL, June 2017.
  – Department of Statistics, Ewha Womans University, Seoul, Korea, June 2017.

– The 9th International Conference of the ERCIM WG (European Research Consortium for Informatics and Mathematics Working Group) on Computational and Methodological Statistics, University of Seville, Spain, December 2016.
– International Indian Statistical Association Conference, Oregon State University, Corvallis, OR, August 2016.
– Department of Statistics and Probability, Michigan State University, East Lansing, MI, November 2015.
– Youngnam Mathematical Society Annual Meeting, Busan National University, Busan, Korea, June 2015.

- A Statistical View of Ranking: Midway between Classification and Regression,
  – Department of Statistics, The Ohio State University, Columbus, OH, September 2016.
  – Department of Computer Science and Engineering, The Ohio State University, Columbus, OH, December 2015.

- A Modern Look at Classical Multivariate Techniques,
  Part I: Regression, Part II: Classification, Part III: Dimensionality Reduction
  The 13th School of Probability and Statistics, CIMAT (Center of Research in Mathematics), Guanajuato, Mexico, March 16–20, 2015.

- Nonparametric Covariance Estimation with Shrinkage Toward Stationary Models,
  The Joint Statistical Meetings, Boston, MA, August 2014.

- A Statistical View of Ranking: Midway between Classification and Regression,
  Conference on Nonparametric Statistics for Big Data and Celebration to Honor Grace Wahba, Madison, WI, June 2014.

- Error Bars?: Discussion of "Error Bars in Experimental Biology" by Cumming et al.,
  Journal of Cell Biology (2007), Molecular Genetics 7802, Research Seminar: Cell Biology, Department of Molecular Genetics, The Ohio State University, December 2013.

- A Comparison of Classification Methods for Robustness,
  The Joint Statistical Meetings, San Diego, CA, August 2012.

- Two Tales of Variable Selection for High Dimensional Regression: Screening and Model Building, The second IMS-APRM (Institute of Mathematical Statistics Asia Pacific Rim Meeting), Tsukuba, Japan, July 2012.

- Statistical Analysis of Bipartite Ranking by Convex Risk Minimization,
  KSS International Conference on Statistics and Probability, Busan, Korea, July 2011.

- A Study of Relative Efficiency and Robustness of Classification Methods,
  – Department of Statistics, Seoul National University, Korea, April 2011.
  – Department of Applied Statistics, Yonsei University, Seoul, Korea, May 2011.
  – Department of Statistics, University of Seoul, Seoul, Korea, June 2011.

- Support Vector Machines for Classification: A Statistical Portrait,
  – The Spring Conference of Korean Statistical Society, KAIST, Daejeon, Korea, May 2011.
  – Ulsan National Institute of Science and Technology, Ulsan, Korea, June 2011.

– Department of Statistics, Duksung Women's University, Seoul, Korea, June 2011.
– Department of Informational Statistics, Hoseo University, Asan, Korea, June 2011.

- Statistical Analysis of Bipartite and Multipartite Ranking by Convex Risk Minimization,
  Summer School on Statistical Pattern Recognition, CIMAT (Center of Research in Mathematics), Guanajuato, Mexico, August 2010.

- Does Modeling Lead to More Accurate Classification?
  – Summer School on Statistical Pattern Recognition, CIMAT (Center of Research in Mathematics), Guanajuato, Mexico, August 2010.
  – ICSA Applied Statistics Symposium, Indianapolis, IN, June 2010.

- Comparison of the Efficiency of Classification Methods,
  – Department of Statistics, Indiana University, Bloomington, IN, March 2010.
  – Department of Food and Resource Economics, University of Delaware, Newark, DE, March 2010.

- Functional Component Pursuit,
  – ICSA Applied Statistics Symposium, San Francisco, CA, June 2009.
  – The first Institute of Mathematical Statistics - Asia Pacific Rim Meeting, Seoul, Korea, July 2009.
  – Department of Statistics, University of Seoul, Korea, July 2009.

- A Bahadur Type Representation of the Linear Support Vector Machine and its Relative Efficiency,
  Machine Learning Summer School (Theory and Practice of Computational Learning), University of Chicago, IL, June 2009.

- A Tutorial on Kernel Methods in a Regularization Framework,
  Fall Conference of Korean Statistical Society, Chung-Ang University, Seoul, Korea, October 2008.

- Linear Programming for Feature Selection via Methods of Regularization,
  – Department of Mathematical Information Science, Tokyo University of Science, Tokyo, Japan, July 2008.
  – Statistical Research Center for Complex Systems, Seoul National University, Seoul, Korea, July 2008.
  – International Conference on Machine Learning and Data Mining, Beijing, China, June 2008.

- A Regularization Approach to Screening and Selection of Biomarkers,
  – Department of Biostatistics and Bioinformatics, Emory University, Atlanta, GA, December 2008.
  – ENAR, Arlington, VA, March 2008.
  – Division of Biostatistics, Washington University in St. Louis, MO, November 2007.

- Another Look at Linear Programming for Feature Selection via Methods of Regularization,
  – Statistics-Econometrics Seminar, Columbia University, New York City, NY, November 2007.
  – Department of Statistics, Carnegie Mellon University, Pittsburgh, PA, September 2007.
  – Department of Statistics, Korea University, Seoul, Korea, August 2007.

- Kernel Methods in a Regularization Framework for Nonparametric Model Building,
  The Fall Workshop of ASA Cleveland Chapter, Cleveland, OH, October 2007.

- A Bahadur Representation of the Linear Support Vector Machine,
  – Department of Mathematics, Washington University in St. Louis, MO, November 2007.
  – Department of Statistics, University of Georgia, Athens, GA, March 2007.

- A Short Course on Kernel Methods in a Regularization Framework,
  Winter School, CIMAT, Guanajuato, Mexico, January 2007.

- A Tutorial on Support Vector Machines,
  – Quantitative Psychology Area Brownbag Series, Department of Psychology,
  The Ohio State University, October 2006.
  – North America Korean Statistical Association, The Joint Statistical Meetings, Seattle, WA,
  August 2006.

- The Solution Path of Multicategory Support Vector Machines, Joint Summer Research Conferences: Machine Learning, Statistics, and Discovery, Snowbird, UT, June 2006.

- Structured Statistical Learning with Support Vector Machine for Feature Selection and Prediction,
  – The Joint Statistical Meetings, Minneapolis, MN, August 2005.
  – WNAR, Fairbanks, AK, June 2005.
  – Graybill Conference, Fort Collins, CO, June 2005.

- A Sparse Solution Approach to Gene Selection for Cancer Diagnosis Using Microarray Data, CCF/OSU/CWRU Joint Biostatistics Symposium, Cleveland Clinic Foundation, May 2005.

- Multicategory Support Vector Machines: a Population View of Different Approaches, The Joint Statistical Meetings, San Francisco, CA, August 2003.

- Discussion of the Wald lecture III (Grace Wahba): Nonstandard Multicategory Support Vector Machine, The Joint Statistical Meetings, San Francisco, CA, August 2003.

- Multicategory Support Vector Machines: Theory and its Applications.
  – Statistical Research Center for Complex Systems, Seoul National University, Korea, August 2003.
  – Department of Statistics, University of Tennessee, Knoxville TN, April 2003.

- Classification of Multiple Cancer Types by Multicategory Support Vector Machines Using Gene Expression Profiles.
  – Summer Research Conference in Statistics, *Statistics in Genetics, Molecular Biology and Bioinformatics*, Jekyll Island, GA, June 2003.
  – Division of Pediatric Informatics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, December 2002.

- Multicategory Support Vector Machines.
  – Department of Biostatistics, University of Michigan, January 2002.
  – Department of Statistics, The Ohio State University, January 2002.

– Department of Statistics, University of Illinois, Urbana-Champaign, February 2002.
– Department of Statistics, Texas A & M, February 2002.
– Department of Statistics, University of California, Davis, February 2002.
– Department of Statistics, University of Florida, Gainesville, February 2002.

Contributed Talks

- Predictive Model Degrees of Freedom in Linear Regression, IMS Annual Meeting, London, UK, June, 2022.
- Supervised Dimensionality Reduction for Exponential Family Data, *Women in Statistics and Data Science Conference*, Cincinnati, OH, October 2018.
- Generalized Principal Component Analysis: Dimensionality Reduction through the Projection of Natural Parameters, *The 4th IMS Asia Pacific Rim Meeting*, The Chinese University of Hong Kong, Hong Kong, June 2016.
- Statistical Consistency of Multipartite Ranking, *The Joint Statistical Meetings*, Montreal, Canada, August 2013.
- Statistical Analysis of Bipartite and Multipartite Ranking by Convex Risk Minimization (poster), *Statistical Science - Making A Difference*, 50th Anniversary Conference of the Department of Statistics, Madison, WI, June 2010.
- Comparison of the Efficiency of Classification Methods, Data Mining and Statistical Learning study group, Department of Statistics, The Ohio State University, April 2010.
- A Bahadur Representation of the Linear Support Vector Machine,
  – *The 7th World Congress in Probability and Statistics*, Singapore, July 2008.
  – Data Mining and Statistical Learning study group, Department of Statistics, The Ohio State University, October 2008.
- Regularization Approach to Screening and Selection of Biomarkers, *The First International Symposium on Biopharmaceutical Statistics*, Shanghai, China, July 2008.
- Quarterly presentations at Data Mining and Statistical Learning study group, Department of Statistics, The Ohio State University, 2007.
- Structured Multicategory Support Vector Machine with ANOVA decomposition,
  – *New Researchers Conference*, York University, Toronto, Canada, August 2004.
  – *The Joint Statistical Meetings*, Toronto Canada, August 2004.
- Quarterly presentations at Statistical Genetics Journal Club, Department of Statistics, The Ohio State University, 2003–2004.
- Classification of Satellite Radiance Data by Multicategory Support Vector Machines (poster), *Third Conference on Artificial Intelligence Applications to the Environmental Science*, Long Beach, CA, February 2003. Won the best student poster award.

- Multicategory Support Vector Machines, Department of Statistics, UW-Madison, April 2002.
- Classification of Multiple Cancer Types by Multicategory Support Vector Machines Using Gene Expression Data, *ENAR 2002*, Crystal City VA, March 2002.
- Multicategory Support Vector Machines with Application to Cancer Classification Using Gene Expression Data (poster), *IMS Mini-Meeting on Imaging, Classification and Clustering*, University of Florida, Gainesville FL, January 2002.
- Multicategory Support Vector Machines, *the 33rd Symposium on the Interface*, Costa Mesa CA, June 2001.

Publications

The following papers are available via my website `http://www.stat.osu.edu/~yklee`.

- Jiao, Z. and Lee, Y., Assessment of Case Influence in the Lasso with a Case-Weight Adjusted Solution Path. June 2024. arXiv:2406.00493. To appear in *Technometrics*.
- Kim, J., Lee, Y., and Liang, Z., The Geometry of Nonlinear Embeddings in Kernel Discriminant Analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 45, No. 4, 5203–5217, 2023.
- Huang, R. and Lee, Y., Debiasing Principal Component Score Estimation in Exponential Family PCA for Sparse Count Data. In the Proceedings of the Joint Statistical Meetings, 1124–1143, 2022.
- Lewis, J. R., MacEachern, S. N., and Lee, Y., Bayesian Restricted Likelihood Methods: Conditioning on Insufficient Statistics in Bayesian Regression (with Discussion). *Bayesian Analysis*, **16**, 1393–1462, 2021.
- Landgraf, A. J. and Lee, Y., Dimensionality Reduction for Binary Data through the Projection of Natural Parameters. *Journal of Multivariate Analysis*, **180**, 104668, 2020.
- Blake, T. A. and Lee, Y., Nonparametric Covariance Estimation with Shrinkage toward Stationary Models. *WIREs Computational Statistics*, **12**, No. 6, e1507, 2020.
- Landgraf, A. J. and Lee, Y., Generalized Principal Component Analysis: Projection of Saturated Model Parameters. *Technometrics*, **62**, Issue 4, 459–472, 2020. `https://doi.org/10.1080/00401706.2019.1668854`
- Uematsu, K., and Lee, Y., On Theoretically Optimal Ranking Functions in Bipartite Ranking. *Journal of the American Statistical Association*, **112**, No. 519, 1311–1322, 2017.
- Uematsu, K., and Lee, Y., Statistical Optimality in Multipartite Ranking and Ordinal Regression. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **37**, No. 5, 1080–1094, 2015.
- Jung, Y., Lee, Y., and MacEachern, S. N., Efficient Quantile Regression for Heteroscedastic Models. *Journal of Statistical Computation and Simulation*, **85**, Issue 13, 2548–2568, 2015.

- Lee, Y. and Wang, R., Does Modeling Lead to More Accurate Classification?: A Study of Relative Efficiency in Linear Classification. *Journal of Multivariate Analysis*, **133**, 232–250, 2015.

- Lee, Y., Comments on Support Vector Machines Maximizing Geometric Margins for Multi-classification. *TOP* (an official journal of the Spanish Society of Statistics and Operations Research), **22**, Issue 3, 852–855, 2014.

- Liu, C., Shi, T., and Lee, Y., Two Tales of Variable Selection for High Dimensional Regression: Screening and Model Building. *Statistical Analysis and Data Mining*, **7**, 140–159, 2014.

- Yao, Y., and Lee, Y., Another Look at Linear Programming for Feature Selection via Methods of Regularization. *Statistics and Computing*, **24**, Issue 5, 885–905, 2014.

- Liang, Z. and Lee, Y., Eigen-Analysis of Nonlinear PCA with Polynomial Kernels. *Statistical Analysis and Data Mining*, **6**, 529–544, 2013.

- Lewis, J. R., MacEachern, S. N. and Lee, Y., Robust Inference via the Blended Paradigm. In the *Proceedings of the Joint Statistical Meetings*, 1773–1786, 2012.

- Lee, Y., MacEachern, S. N., and Jung, Y., Regularization of Case-Specific Parameters for Robustness and Efficiency. *Statistical Science*, **27**, 350–372, 2012.

- Lee, Y., Support Vector Machines for Classification: A Statistical Portrait. In *Statistical Methods in Molecular Biology*, Heejung Bang, Xi Kathy Zhou, Heather L. Van Epps, and Madhu Mazumdar, eds, in the series of Methods in Molecular Biology, Humana Press, 347–368, 2010.

- Hsu, J., Rao, Y., Lee, Y., Chang, J., Bergsteinsdottir, K., Magnusson, M.K., Wang, T., Steingrimsson, E., Design and Analysis of Microarray Experiments for Pharmacogenomics. In *Multiple Testing Problems in Pharmaceutical Statistics*, Dmitrienko, A., Tamhane, A.C., Bretz, F., eds., Chapman & Hall/CRC Biostatistics Series, 239–264, 2009.

- Koo, J.-Y., Lee, Y., Kim, Y., and Park, C., A Bahadur Representation of the Linear Support Vector Machine. *Journal of Machine Learning Research*, **9**, 1343–1368, 2008.

- Rao, Y., Lee, Y., Jarjoura, D., Ruppert, A. S., Liu, C., Hsu, J. C., and Hagan, J. P., A Comparison of Normalization Techniques for MicroRNA Microarray Data. *Statistical Applications in Genetics and Molecular Biology*, **7**, Issue 1, Article 22, 2008.

- Lee, Y., Kim, Y., Lee, S., and Koo, J.-Y., Structured Multicategory Support Vector Machine with ANOVA decomposition. *Biometrika*, **93**, 555–571, 2006.

- Lee, Y., (Book Review) Semiparametric Regression by David Ruppert, M. P. Wand, and R. J. Carroll. *Journal of the American Statistical Association*, **101**, 1722–1723, 2006.

- Lee, Y. and Cui, Z., Characterizing the Solution Path of Multicategory Support Vector Machines. *Statistica Sinica*, **16**, 391–409, 2006.

- Lee, Y., Lin, Y., and Wahba, G., Multicategory Support Vector Machines, Theory, and Application to the Classification of Microarray Data and Satellite Radiance Data. *Journal of the American Statistical Association*, **99**, 67–81, 2004.

- Lee, Y., Wahba, G., and Ackerman, S., Classification of Satellite Radiance Data by Multicategory Support Vector Machines. *Journal of Atmospheric and Oceanic Technology*, **21**(2), 159–169, 2004.

- Lee, Y. and Lee, C.-K., Classification of Multiple Cancer Types by Multicategory Support Vector Machines Using Gene Expression Data. *Bioinformatics*, **19**, 1132–1139, 2003.

- Lee, Y., Multicategory Support Vector Machines, Theory, and Application to the Classification of Microarray Data and Satellite Radiance Data. Ph.D. thesis, Department of Statistics, University of Wisconsin-Madison, 2002.

- Wahba, G., Lin, Y., Lee, Y., and Zhang, H., Optimal Properties and Adaptive Tuning of Standard and Nonstandard Support Vector Machines. In *Nonlinear Estimation and Classification*, Denison, D. D., Hansen, M. H., Holmes, C. C., Mallick, B., and Yu, B., eds, Springer, New York, 129–147, 2003.

- Lin, Y., Wahba, G., Zhang, H., and Lee, Y., Statistical Properties and Adaptive Tuning of Support Vector Machines. *Machine Learning*, **48**, 115–136, 2002.

- Lin, Y., Lee, Y., and Wahba, G., Support Vector Machines for classification in nonstandard situations. *Machine Learning*, **46**, 191–202, 2002.

- Lee, Y., Lin, Y. and Wahba, G., Multicategory Support Vector Machines. *Computing Science and Statistics 33*, 498–512, 2001.

Unpublished Manuscripts

- Huang, R. and Lee, Y., Debiasing Sample Loadings and Scores in Exponential Family PCA for Sparse Count Data. December 2023. arXiv:2312.13430.

- Luan, B., Lee, Y., and Zhu, Y., On Measuring Model Complexity in Heteroscedastic Linear Regression. April 2022. arXiv:2204.07021.

- Luan, B., Lee, Y., and Zhu, Y., Predictive Model Degrees of Freedom in Linear Regression. June 2021. arXiv:2106.15682.

- Zhang, J. and Lee, Y., Sparse Logistic Tensor Decomposition for Binary Data. June 2021. arXiv:2106.14258.

- Tu, S., Zhu, Y., Lee, Y., Gu, Q., and Yu H., Cross Validation for Penalized Quantile Regression with a Case-Weight Adjusted Solution Path. Technical Report No. 893, Department of Statistics, The Ohio State University, February 2019. arXiv:1902.07770.

- Jung, Y., MacEachern, S. N., and Lee, Y., Window Width Selection for L_2 Adjusted Quantile Regression. Technical Report No. 835, Department of Statistics, The Ohio State University, April 2010.
- Rao, Y., Lee, Y., and Hsu, J. C., Determination of Sample Size for Validation Study in Pharmacogenomics. Technical Report No. 834, Department of Statistics, The Ohio State University, January 2010.

Professional Membership and Service

- Member of the American Statistical Association
- Member of the Institute of Mathematical Statistics
- Member of the Korean International Statistical Society
- Editorial Board Reviewer of *Journal of Machine Learning Research* since July 2020
- Associate Editor of *Econometrics and Statistics* since May 2019
- Associate Editor of *Chemometrics and Intelligent Laboratory Systems* since October 2018
- Associate Editor of *Journal of the Korean Statistical Society* since January 2017
- Associate Editor of *Statistical Analysis and Data Mining* since June 2011
- Board of Directors, the Korean International Statistical Society, 2023–2028
- External reviewer of a tenure and promotion case, September 2024
- Organizer of Invited Session on *Model Assessment, Explainability and Validation*, ASA Conference on Statistical Learning and Data Science, Newport Beach, CA, November 2024
- Judge of student presentations, ASA DataFest at The Ohio State University, April 2024
- Organizer of Topic-Contributed Paper Session on *New and Unexpected Risk Landscapes for Overparameterized Models in Machine Learning* for the Joint Statistical Meetings, Toronto, Canada, August 2023
- Panelist of Computational and Data-Enabled Science and Engineering in Mathematical and Statistical Sciences Program, National Science Foundation, 2023
- Reviewer of a proposal submitted to US-Israel Binational Science Foundation, 2023
- External reviewer of a tenure and promotion case, March 2023
- Reviewer of TDAI Interdisciplinary Research Pilot Awards, February 2023
- Organizer of Invited Session on *Emerging Statistical Issues in Overparameterized Modeling* for Computational and Methodological Statistics 2022, London, UK, December 2022 (hybrid)

- Program Chair of the inaugural TDAI Foundations CoP Deep Learning Summer School, Columbus, Ohio, June 2022 (hybrid)
- Panelist of Statistics Program, Division of Mathematical Sciences, National Science Foundation, 2020
- External reviewer of a tenure and promotion case, September 2020
- Reviewer for Neural Information Processing Systems (NeurIPS) Conference 2020
- Organizer of Invited Session on *Multiple Facets of Dimension Reduction* for The 4th International Conference on Econometrics and Statistics, Hong Kong, June 2021 (virtual)
- Organizer of Invited Session on *Recent Advances in Statistical Learning* for The 3rd International Conference on Econometrics and Statistics, Taiwan, June 2019
- Panelist of Statistics Program, Division of Mathematical Sciences, National Science Foundation, 2018
- Organizer of Invited Session on *Statistical Computing and Optimization* for The 2nd International Conference on Econometrics and Statistics, Hong Kong, June 2018
- Program Chair of the Korean International Statistical Society, 2015–2016
- Organizer of KISS Invited Session on *Dimensionality Reduction* for International Indian Statistical Association Conference, Corvallis, OR, 2016
- Member of the Student Paper Award Committee of ASA Nonparametric Statistics Section for the Joint Statistical Meetings 2014
- Member of Program Committee for The Joint Applied Statistics Symposium, Portland, Oregon, June 15–18, 2014
- Panelist of Statistics Program, Division of Mathematical Sciences, National Science Foundation, 2013
- General Methodology Co-Chair, Program Committee for the Joint Statistical Meetings, San Diego, CA, 2012
- Member of Senior Program Committee for the Fourth Asian Conference on Machine Learning (ACML), Singapore, November 2012
- Member of Program Committee for Conference on Statistical Learning and Data Mining, Ann Arbor, MI, June 5–7, 2012
- Member of Senior Program Committee for the Third Asian Conference on Machine Learning (ACML), Taoyuan, Taiwan, November 2011
- Member of Program Committee for the First International Conference on Pattern Recognition Applications and Methods, Vilamoura, Portugal, February 2012

- Member of Organizing Committee for the Korean Statistical Society International Conference on Statistics and Probability, Busan, Korea, July 2011
- Member of Screening Committee for the Journal of Korean Statistical Society, April–June 2011
- Guest Editor for Special Issue on *Nonparametrics and Statistical Learning* of the ASA Journal *Statistical Analysis and Data Mining*, 2010–2011
- Reviewer of Papers Submitted to Neural Information Processing Systems Conference 2010
- Member of Local Program Committee for Conference on Nonparametric Statistics and Statistical Learning, Columbus, Ohio State University, May 2010
- Member of Program Committee for Artificial Intelligence and Statistics 2010, Sardinia, Italy
- Reviewer of a paper submitted to Conference on Learning Theory (COLT) 2009
- Member of Program Committee for the 12th International Conference on Applied Stochastic Models and Data Analysis 2007
- Member of Program Committee for Artificial Intelligence and Statistics 2007
- Organizer of IMS Invited Session on *Semi-Supervised Learning* for the Joint Statistical Meetings 2006
- Reviewer of a proposal to The Ohio Supercomputer Center, November 2005
- Reviewer of papers submitted to Neural Information Processing Systems Conference 2003
- Reviewer of a proposal submitted to US-Israel Binational Science Foundation, 2003
- Reviewer of papers for the following journals:
  Artificial Intelligence
  The American Statistician
  The Annals of Applied Statistics
  The Annals of Statistics
  Annals of the Institute of Statistical Mathematics
  Bernoulli
  BMC Bioinformatics
  Bioinformatics
  Biometrics
  Biometrika
  Chemometrics and Intelligent Laboratory Systems
  Computational Statistics and Data Analysis
  Contemporary Mathematics (AMS)
  Data Mining and Knowledge Discovery
  Econometrics and Statistics
  Electronic Journal of Statistics
  European Journal of Operational Research

IEEE Transactions on Information Theory
IEEE Transactions on Neural Networks
IEEE Transactions on Neural Networks and Learning Systems
Information and Inference: A Journal of the IMA
International Statistical Review
Journal of the American Statistical Association
Journal of Artificial Intelligence Research
Journal of Computational and Graphical Statistics
Journal of Machine Learning Research
Journal of Nonparametric Statistics
Journal of the Royal Statistical Society: Series B
Journal of Statistical Planning and Inference
Probability Theory and Related Fields
SIAM Journal on Optimization
Soft Computing
Statistica Sinica
Statistical Analysis and Data Mining
Statistics and Computing
Statistics in Medicine
Technometrics

Committees

- Ph.D. Candidacy Exam Committee

April 2025	Xuerong Wang (Statistics)
April 2025	Meghna Karla (Electrical and Computer Engineering, Advisor: Kiryung Lee)
April 2025	Wenyan Luo (Mathematics, Advisor: Dustin Mixon)
April 2025	Fangyi Wang (Statistics, Advisors: Sebastian Kurtek and Yuan Zhang)
October 2024	Jae Ho Chang (Statistics, Advisor: Subhadeep Paul)
May 2024	Xinyu Zhang (Statistics, Advisor: Steve MacEachern)
December 2023	Haozhen Yu (Statistics)
December 2023	Zhenbang Jiao (Statistics)
April 2023	Rui Zhang (Statistics, Advisors: Oksana Chkrebtii and Dongbin Xiu)
March 2023	Seonho Kim (Electrical and Computer Engineering, Advisor: Kiryung Lee)
December 2022	Meijia Shao (Statistics, Advisor: Yuan Zhang)
December 2022	Lun Li (Statistics)
November 2022	Pashmeen Kaur (Statistics, Advisor: Peter Craigmile)
May 2022	Chenxi Zhou (Statistics, Advisor: Vince Vu)
May 2022	Ruochen Huang (Statistics)
February 2022	Xinyun Jiang (City and Regional Planning, Advisor: Yasuyuki Motoyama)
December 2021	Qiuyu Gu (Statistics, Advisor: Yunzhang Zhu)
November 2021	Daryl Swartzentruber (Statistics, Advisor: Elly Kaizar)

June 2021	Renxiong Liu (Statistics, Advisor: Yunzhang Zhu)
December 2020	Prateek Sasan (Statistics, Advisor: Vince Vu)
August 2020	Bo Luan (Statistics, Co-Advisor: Yunzhang Zhu)
April 2020	Zhaojun Li (Psychology, Advisor: Paul De Boeck)
August 2019	Jianhao Zhang (Statistics)
July 2019	Dewei Zhang (Integrated Systems Engineering, Advisor: Sam Davanloo)
June 2019	Jiae Kim (Statistics, Co-Advisor: Steve MacEachern)
May 2019	Shuyi Wang (Statistics, Advisors: Lo-Bin Chang and Dongbin Xiu)
February 2019	Min Ho Cho (Statistics, Advisors: Sebastian Kurtek and Steve MacEachern)
August 2018	Mahan Mansouri (Integrated Systems Engineering)
May 2018	Shanshan Tu (Statistics)
May 2018	Shuhan Tang (Statistics)
March 2018	Xiaoyi Chen (Statistics)
March 2018	Joonsuk Park (Psychology)
February 2017	Justin Strait (Statistics)
November 2016	Jangho Park (Integrated Systems Engineering)
August 2016	Qian Qian (Statistics)
August 2016	Liubo Li (Statistics)
May 2016	Huseyin Bahtiyar (Industrial and Systems Engineering)
April 2016	Jieyi Jiang (Statistics)
April 2016	Andrew Bean (Statistics)
July 2015	Zhifei Yan (Statistics)
April 2015	Rong Lu (Biostatistics)
December 2014	Jiaqi Zaetz (Statistics)
September 2014	Fabian Benitez-Quiroz (Electrical and Computer Engineering, Advisor: Aleix Martinez)
April 2014	Ran Wei (Statistics)
April 2014	Andrew Landgraf (Statistics)
November 2013	John Lewis (Statistics)
July 2013	Sungmin Kim (Statistics)
January 2013	Tayler Blake (Statistics)
December 2012	Zhiyu Liang (Statistics)
December 2012	Siyoen Kil (Statistics)
November 2012	Alice Hinton (Biostatistics)
July 2012	Jiangyong (Matt) Yin (Statistics)
July 2012	Jonathan Bradley (Statistics)
May 2012	Harsha Gangammanavar (Integrated Systems Engineering)
September 2011	Rui Wang (Statistics)
August 2011	Dunke Zhou (Statistics)
February 2011	Cong Liu (Statistics)
December 2010	Kazuki Uematsu (Statistics)
November 2010	Di You (Electrical and Computer Engineering, Advisor: Aleix Martinez)
June 2010	Seoungbum Kim (Civil Engineering)*
March 2010	Tao Xiao (Biostatistics)
December 2009	John Draper (Statistics)
December 2009	Prasenjit Kapat (Statistics)

August 2009	Yoonsuh Jung (Statistics)
March 2009	Jared Schuetter (Statistics)
August 2008	Yi Liu (Statistics)
February 2008	Onur C. Hamsici (Electrical and Computer Engineering, Advisor: Aleix Martinez)
June 2007	Youlan Rao (Statistics)
August 2006	Kaushik Sinha (Computer Science and Engineering)
August 2006	Dianne Bautista (Statistics)
May 2006	Lili Yu (Statistics)
March 2006	Zhenhuan Cui (Statistics)
March 2006	Yanxing Zhao (Statistics)
December 2005	Brandon Moore (Electrical and Computer Engineering)*
September 2005	Qingzhao Yu (Statistics)
May 2005	Bin Li (Statistics)
August 2004	Haiyan Xu (Statistics)
December 2003	Charalampos Papachristou (Statistics)
June 2003	Changyi Park (Statistics)
May 2003	Yufeng Liu (Statistics)
March 2003	Sijin Liu (Statistics)

* as a Graduate School Representative

- Ph.D. Dissertation Committee

April 2025	Zhenbang Jiao (Statistics)
Nov 2024	Seonho Kim (Electrical and Computer Engineering, Advisor: Kiryung Lee)
July 2024	Patrick Millican (Physics, Advisor: Richard Furnstahl)*
July 2024	Pashmeen Kaur (Statistics, Advisor: Peter Craigmile)
May 2024	Lun Li (Statistics)
November 2023	Meijia Shao (Statistics, Advisor: Yuan Zhang)
July 2023	Renxiong Liu (Statistics, Advisor: Yunzhang Zhu)
July 2023	Ruochen Huang (Statistics)
April 2023	Eunseop Kim (Statistics, Advisors: Steve MacEachern and Mario Peruggia)
December 2022	Qiuyu Gu (Statistics, Advisor: Yunzhang Zhu)
August 2022	Chenxi Zhou (Statistics, Advisor: Vince Vu)
August 2022	Elan Weiss (Materials Science and Engineering, Advisor: Wolfgang Windl)*
August 2022	Prateek Sasan (Statistics, Advisors: Vince Vu and Subhadeep Paul)
December 2021	Bo Luan (Statistics, Co-Advisor: Yunzhang Zhu)
May 2021	Jianhao Zhang (Statistics)
April 2021	Shuyi Wang (Statistics, Advisors: Lo-Bin Chang and Dongbin Xiu)
November 2020	Jiae Kim (Statistics, Co-Advisor: Steve MacEachern)
July 2020	Min Ho Cho (Statistics, Advisors: Sebastian Kurtek and Steve MacEachern)
April 2020	Shuhan Tang (Statistics, Advisors: Peter Craigmile and Yunzhang Zhu)
August 2019	Qian Qian (Statistics, Advisors: Vince Vu and Yunzhang Zhu)
August 2019	Meraj Ahmed Khan (Computer Science and Engineering)*
July 2019	Shanshan Tu (Statistics, Co-Advisor: Yunzhang Zhu)
April 2018	Tayler Blake (Statistics)

August 2017	Liubo Li (Statistics, Co-Advisor: Vince Vu)
August 2017	Jieyi Jiang (Statistics, Co-Advisor: Steve MacEachern)
August 2017	Zhifei Yan (Statistics, Advisor: Vince Vu)
July 2017	Andrew Bean (Statistics, Advisors: Steve MacEachern and Xinyi Xu)
July 2017	Tuo Liu (Economics)*
April 2016	Qichao Que (Computer Science and Engineering, Advisor: Mikhail Belkin)
August 2015	Jiaqi Zaetz (Statistics, Advisor: Sebastian Kurtek)
July 2015	Yang Zhang (Computer Science and Engineering)*
May 2015	Andrew Landgraf (Statistics)
July 2014	John Lewis (Statistics, Co-Advisor: Steve MacEachern)
April 2014	Sungmin Kim (Statistics, Advisor: Tao Shi)
January 2014	Alice Hinton (Biostatistics, Advisor: Raj Nagaraja)
December 2013	Zhiyu Liang (Statistics)
August 2013	Yanfei Zhu (History of Art)*
August 2012	Kazuki Uematsu (Statistics)
August 2012	Rui Wang (Statistics)
May 2012	Dunke Zhou (Statistics, Advisor: Tao Shi)
May 2012	Harrison Smith (Electrical and Computer Engineering)*
April 2012	Cong Liu (Statistics, Co-Advisor: Tao Shi)
February 2012	John Draper (Statistics, Advisor: Prem Goel)
November 2011	Di You (Electrical and Computer Engineering, Advisor: Aleix Martinez)
February 2011	Kevin Driver (Physics)*
May 2010	Yoonsuh Jung (Statistics, Co-Advisor: Steve MacEachern)
March 2010	Jared Schuetter (Statistics, Advisor: Tao Shi)
June 2009	Youlan Rao (Statistics, Co-Advisor: Jason Hsu)
August 2008	Yonggang Yao (Statistics)
August 2008	Sivaramakrishnan Narayanan (Biomedical Informatics)*
August 2008	Onur C. Hamsici (Electrical and Computer Engineering, Advisor: Aleix Martinez)
August 2007	Zhenhuan Cui (Statistics)
May 2007	Bae-Geun Kim (Economics)*
August 2006	Bin Li (Statistics, Advisor: Prem Goel)
May 2006	Sijin Liu (Statistics, Advisor: Xiaotong Shen)
September 2005	Haiyan Xu (Statistics, Advisor: Jason Hsu)
August 2005	Changyi Park (Statistics, Advisor: Xiaotong Shen)
May 2005	Tao Wang (Statistics, Advisor: Jason Hsu)
April 2004	Yufeng Liu (Statistics, Advisor: Xiaotong Shen)

- Master Thesis Committee

July 2024	Steve Dignan (Statistics)
April 2020	Xiaoyu Liang (Geography, Advisor: Desheng Liu)

- Undergraduate Thesis Committee

May 2018	Mingfeng Li (Civil, Environmental, and Geodetic Engineering, Advisor: Andre Carrel)

- Exam Committee

Ph.D. Qualifying Exam II	September 2006, March 2007, September 2007 March 2008, September 2008 (Chair), March 2009 September 2009, March 2010, September 2010 September 2011 (Chair), March 2012 August 2013 (Chair), and January 2014
Ph.D. Qualifying Exam I	October 2005, January 2006, August 2012, and January 2013
Master of Applied Statistics Exam	May 2003, May 2004, and November 2004 January 2015, May 2015, January 2016 May 2016 (Chair), January 2017, January 2018 May 2018, January 2019, May 2019 (Chair) January 2021, May 2021, January 2022 (Chair) May 2022, January 2023, May 2023, January 2024 May 2024, January 2025 (Chair)

- Departmental Committee

2023–2025	Curriculum Committee (Co-Chair)
2023–2024	Anniversary Event Committee
2022–2023	PhD Program Revision (Co-Chair)
2020–2022	Executive Advisory Committee
2021–2022	Search Committee (Chair)
2021–2022	External Review Committee (Chair)
2020-2021	Curriculum Committee (Undergraduate Majors)
2015–2019	Curriculum Committee (Chair)
Fall 2018	Chair Reappointment Review Committee
2017–2019	Executive Advisory Committee
2015–2017	Search Committee
2013–2014	Executive Advisory Committee, Curriculum Committee, Search Committee
2012	Colloquium Committee, Curriculum Committee, Search Committee
2011	Search Committee, Admissions Committee
2010	Semester Conversion of Undergraduate Minor in Statistics Program
Spring 2009	Seminar Committee
2008–2010	Admissions Committee
2006–2007	Curriculum Committee
2002–2006	Computer Advisory Committee
Fall 2003	Search Committee (Statistical Genetics)

- College or University Service

2023 – present	Arts and Sciences Curriculum Committee Natural and Mathematical Sciences Subcommittee
July 2020–June 2022	Co-Lead, Foundations and AI Community of Practice, TDAI
Fall 2014–2017	Statistics Liaison with College of Engineering Core Curriculum and College Services Committee
April–August 2014	Data Analytics Collaborative Advisory Committee

Last updated April 13, 2025