

Danhyang Lee

| | | |
|---------------------|---|---|
| CONTACT INFORMATION | Baylor University Department of Statistical Science Marrs McLean Science Bldg Waco, TX 76706, USA | <i>E-mail:</i> Danhyang_Lee@baylor.edu <i>Phone:</i> +1 (515) 686-7112 |
| EMPLOYMENT | Assistant Professor <i>Department of Statistical Science, Baylor University, Waco, Texas, USA</i> | July 2024 - |
| | Assistant Professor <i>Department of Information Systems, Statistics and Management Science, Culverhouse College of Business, The University of Alabama, Tuscaloosa, Alabama, USA</i> | August 2019 - May 2024 |
| EDUCATION | Ph.D. in Statistics, Iowa State University , Ames, Iowa, USA • Dissertation Topic: “Topics on small area estimation, multilevel models, and semiparametric imputation” • Co-Advisors: Dr. Jae-Kwang Kim and Dr. Emily Berg | August 2019 |
| | M.S. in Statistics Kyungpook National University , Daegu, South Korea | February 2012 |
| | B.S. in Statistics Kyungpook National University , Daegu, South Korea | February 2010 |
| | B.S. in Mathematics Kyungpook National University , Daegu, South Korea | February 2010 |
| RESEARCH INTERESTS | Missing data analysis, Survey data integration, Bayesian modeling, Survey sampling methodology, Small area estimation, Causal inference | |
| PUBLICATIONS | (ABDC journal quality rating; * Ph.D. student; + corresponding author) | |
| | +Lee, D., and Chen, S. (2024+) Data integration with nonprobability sample: semiparametric model-assisted approach, <i>Scandinavian Journal of Statistics(A)</i> , Invited for 2nd Round Review. | |
| | Zhang, L.C., Sande, L. S., and Lee, D. (2024) Design-based predictive inference, <i>Journal of Official Statistics</i> , Accepted. | |
| | *Boumis, G., Geist, E. L., and Lee, D. (2023) Bayesian hierarchical modeling for probabilistic estimation of tsunami amplitude: uncertainty reduction and ungauged information. <i>Journal of Geophysical Research - Oceans</i> , Accepted. | |
| | Uehara, M., Lee, D., and Kim, J. K. (2023) Semiparametric response model with nonignorable nonresponse. <i>Scandinavian Journal of Statistics(A)</i> , Accepted. | |
| | +Lee, D., Zhang, L-C., and Chen, S. (2023) Robust quasi-randomisation-based estimation with ensemble learning for missing data. <i>Scandinavian Journal of Statistics(A)</i> , 50(3), 1263-1278. | |
| | Lee, D., and Kim, J. K. (2022) Semiparametric imputation using conditional Gaussian mixture models under item nonresponse. <i>Biometrics(A*)</i> , 78(1), 227-237. | |
| | Sang, H., Kim, J. K., and Lee, D. (2022) Semiparametric fractional imputation using Gaussian mixture models for handling multivariate missing data. <i>Journal of the American Statistical Association(A*)</i> , 117(538), 654-663. | |

Lee, D., Zhang, L-C., and Kim, J. K. (2021+) Maximum Entropy Classification for Record Linkage. *Survey Methodology*, Accepted.

Lee, D., Kim, J. K., and Skinner, C. J. (2019) Within-cluster Resampling for Multilevel Models under Informative Cluster Size. *Biometrika(A*)*, 106(4), 965-972.

Berg, E. and **Lee, D.** (2019) Prediction of Small Area Quantiles for the Conservation Effects Assessment Project Using a Mixed Effects Quantile Regression Model. *Annals of Applied Statistics(A*)*, 13(4), 2158-2188.

Berg, E. and **Lee, D.** (2019) Small Area Prediction of Quantiles for Zero-Inflated Data and an Informative Sample Design. *Statistics Theory and Related Fields*, 3(2), 114-128.

Lee, D., Arbuckle, J., Zhu, Z. and Nowatzke, L. (2018) Conditional Causal Mediation Analysis of Factors Influencing Cover Crop Adoption in Iowa, USA. *Water Resources Research(A)*, 54(11), 9566-9584

Lee, D., Nandram, B., and Kim, D. (2017). Bayesian Predictive Inference of a Proportion under a Two-Fold Small Area Model with Heterogeneous Correlations. *Survey Methodology*, 43(1), 69-92.

Lee, D., Kim, G., and Lee, K. E. (2013). Soil Moisture Prediction Using a Support Vector Regression. *Journal of the Korean Data and Information Science Society*, 24(2), 401-408.

OTHER RESEARCH
MANUSCRIPTS

Lee, D. (2016) Multivariate Regression Imputation Approach to the Analysis of Item Nonresponse in a Retail Trade Survey Data. *Proceedings of the Fifth International Conference on Establishment Surveys*.

WORKING PAPERS

(* Ph.D. student)

Lee, D., Cho, S. and Vakhutinsky, A. Demand estimation with missing data on covariates: Imputation-based approach.

Lee, D. and Chen, S. A novel composite weighting method for combining probability and non-probability samples.

*Boumis, G., Moftakhari, H. R., **Lee, D.** and Moradkhani, H. In search of non-stationary dependence between estuarine river discharge and storm surge using Bayesian dynamic copulas and large-scale climate indices

*Dong, A., and **Lee, D.** Multiple imputation using sparse conditional mixture modeling and its application to data integration.

GRANTS

Project title: New statistical tools for demand estimation applications

Sponsor: Culverhouse Summer Excellence in Research for Faculty Grant, Culverhouse College of Business, The University of Alabama

Role: PI

Amount: \$10,000 (May 2022 - Aug 2022)

HONORS AND
AWARDS

Culverhouse College of Business Early Achievement in Research Award by *Culverhouse College of Business, The University of Alabama, 2022*

1st Place of the Fifth International Conference Establishment Surveys Student Contest by *Fifth International Conference Establishment Surveys, 2016*

The Best Student Fellowship for graduate student by *Department of Statistics, Kyungpook National University, Daegu, South Korea, 2011*

The Alumni Award by *Department of Statistics, Kyungpook National University, Daegu, South Korea, 2010*

PRESENTATIONS

Data integration with nonprobability sample: semiparametric model-assisted approach. *Joint Statistical Meetings (JSM)*, Portland, Oregon, USA, August 2024.

Semiparametric imputation using conditional Gaussian mixture models. *UA Applied Statistics Seminar*, Tuscaloosa, AL, USA, 2021.

Semiparametric Imputation using Conditional Gaussian Mixture Model. *Joint Statistical Meetings (JSM)*, Denver, CO, July, 2019.

Parameter Estimation of Random Effect Models under Cluster-specific Nonignorable Nonresponse. *International Chinese Statistical Association (ICSA)*, Raleigh, NC, June, 2019.

Within-cluster Resampling for Generalized Linear Mixed Models under Informative Cluster Size. *Joint Statistical Meetings (JSM)*, Vancouver, Canada, 2018.

Small Area Quantile Prediction Using a Mixed Effects Quantile Regression Model Under Informative Sampling. *Statistical Society of Canada (SSC)*, Montreal, Canada, 2018.

Analysis of Clustered Data under Informative Cluster Sizes. *Canadian Statistical Sciences Institute (CANSSI) Workshop*, Montreal, Canada, 2018.

Causal Mediation Analysis to Iowa Nutrient Reduction Strategy Survey Data. *Center for Survey Statistics and Methodology (CSSM)*, Iowa, USA, 2017.

INVITED TALKS

Data integration with nonprobability sample: semiparametric model-assisted approach. *7th International Conference on Econometrics and Statistics (EcoSta 2024)*, Virtual, July 2024.

Enhanced Statistical Inference with Semiparametric Models for Nonignorable Nonresponse and Data Integration Applications. Sungkyunkwan University, Seoul, South Korea, July 2024.

Enhanced Statistical Inference with Semiparametric Models for Nonignorable Nonresponse and Data Integration Applications. Kyungpook National University, Daegu, South Korea, June 2024.

Enhancing Demand Estimation in the Presence of Missing Covariates. *Decision Sciences Institute 54rd Annual Conference*, Phoenix, Arizona, USA, October 2023.

Demand estimation with missing data on covariates: Imputation-based approach. *Decision Sciences Institute 53rd Annual Conference*, Houston, TX, USA, November 2022.

Demand estimation with missing data on covariates: Imputation-based approach. *23rd ACM Conference on Economics and Computation (EC'22)*, Boulder, CO, USA, July 2022.

Imputation approach to data integration: Combining information from survey and administrative data. *Virtual International Statistical Institute (ISI) World Statistics Congress (WSC) 2021*.

Semiparametric imputation using conditional Gaussian mixture models. *2020 SDSU Data Science Symposium*, Brookings, South Dakota, USA, 2020.

Within-cluster Resampling for Multilevel Models under Informative Cluster Size. *Korean Statistical Society (KSS) Fall Conference*, Seoul, South Korea, 2018.

Multivariate Regression Imputation Approach to the Analysis of Item Nonresponse in a Retail Trade Survey Data. *Fifth International Conference on Establishment Surveys*, Geneva, Switzerland, 2016.

ACADEMIC
EXPERIENCE

Instructor

Spring 2020 - Spring 2024

Department of Information Systems, Statistics and Management Science, The University of Alabama, Tuscaloosa, Alabama, USA

- STAT 260 Statistical Data Analysis
- STAT 440 Statistical Programming & Computing with R
- STAT 540 Statistical Programming & Computing with R
- STAT 547 Data Visualization and Analytics in R
- STAT 697 Bayesian Statistics

Spring 2014; Spring 2012 - Spring 2013

Department of Information Statistics, Andong National University, Andong, South Korea

- Elementary Statistics
- Mathematical Statistics II
- Computational Statistics
- Multivariate Analysis

Lab Instructor

Spring 2019

Department of Statistics, Iowa State University, Ames, Iowa, USA

- STAT326 Introductory Business Statistics II

Teaching Assistant

Fall 2014 - Spring 2016

Department of Statistics, Iowa State University, Ames, Iowa, USA

- STAT500 Statistical Methods I
- STAT543 Theory of Probability and Statistics II
- STAT407 Multivariate Analysis
- STAT305 Engineering Statistics

PROFESSIONAL
EXPERIENCE

Journal Referee

- Computational Statistics and Data Analysis
- Journal of Computational and Graphical Statistics
- Journal of Survey Statistics and Methodology
- Journal of the Korean Statistical Society
- Scandinavian Journal of Statistics
- Statistics in Medicine
- Statistica Sinica
- Statistics and Its Interface
- Survey Methodology
- The Annals of Applied Statistics

Statistics Korea, Daejeon, South Korea

- Mass imputation under imperfect matching, September 2018 - November 2018.

Center for Survey Statistics and Methodology, Ames, Iowa, USA

- Causal mediation analysis for Nutrient Reduction Strategy Project, August 2016 - August 2018.
- Small area quantile prediction for Conservation Effects Assessment Project, August 2016 - May

2018.

Worcester Polytechnic Institute, Massachusetts, USA

- A two-fold Bayesian model for a binomial response variable, January 2013 - February 2013.

National Evidence-based Healthcare Collaborating Agency, Seoul, Korea

- Bayesian Meta Analysis using WinBUGS, July 2012 - October 2012.

COMPUTER SKILLS

- Extensive experience with R/Rcpp, SAS, WinBUGS, Matlab, and SPSS.
- **Certification:** SAS certified advanced programmer for SAS 9, 2011.