

Jooho Kim

✉ jooho991122 (at) gmail (dot) com 🏠 [Personal Website](#) 🐙 [GitHub](#)

Education

Seoul National University <i>MS in Statistics</i>	Seoul, South Korea Mar. 2024 – Feb. 2026
Korea University <i>BE in Food and Resource Economics, Double major in Statistics</i>	Seoul, South Korea Mar. 2018 – Feb. 2024
The University of Texas at Austin <i>Exchange Program, Economics</i>	Texas, United States Aug. 2022 – Dec. 2022

Research Interests

Missing Data, Survival Analysis, Causal Inference, Statistical Uncertainty Quantification in AI

Preprints

Kim, J. and Shin, Y. E. (2025). “Scalable and Efficient Multiple Imputation for Case-Cohort Studies via Influence Function-Based Supersampling.” arXiv:2511.14692. [\[paper\]](#) [\[software\]](#)

Research Experience

Prediction Model Lab, Seoul National University <i>Graduate Researcher (Advisor: Dr. Yei Eun Shin)</i>	Seoul, South Korea Jun. 2024 – Present
<ul style="list-style-type: none"> ◦ Led as the primary graduate researcher on a project funded by the National Research Foundation of Korea: “Multiple Imputation for Missing Covariates due to Epidemiological Cohort Sampling Designs”. ◦ Proposed an influence function-based supersampling approach to impute only a subset (e.g., 3%) of the missing covariates while preserving efficiency and unbiasedness. ◦ Devised weight calibration equations that reconcile heterogeneous sampling weights for a unified Cox proportional hazards regression analysis. ◦ Applied the proposed method to the NIH-AARP Diet and Health Study to assess expensive biomarkers associated with pancreatic cancer risk using the Cox proportional hazards model. 	
Urban Informatics Lab, The University of Texas at Austin <i>Undergraduate Research Assistant (Connected through Dr. Arya Farahi)</i>	Austin, United States Oct. 2022 – Dec. 2022
<ul style="list-style-type: none"> ◦ Aggregated and cleaned geotagged electric vehicle tweets using regular expressions to handle misspellings and variant notations of the U.S. states. ◦ Conducted hotspot analysis across the U.S. to identify regions with significant EV-related public sentiment. 	

Presentation

Scalable and Robust Multiple Imputation for Case-Cohort Studies via Influence Function-Based Supersampling Korean Statistical Society, Seoul, Korea (Oral Presentation).	Dec. 2025
Multiple Imputation for Incomplete Survival Data with Missing Covariates: Toward Valid Causal Inference The 2nd Symposium on Causal Inference, Seoul, Korea (Oral Presentation, English).	Jun. 2025

Honors and Awards

Fellowship for Fundamental Academic Fields <i>Awarded by Seoul National University for academic excellence and research potential</i>	2024, 2025
Graduate Research Fellowship in Science and Engineering <i>Awarded by the National Research Foundation of Korea through a competitive selection process</i>	2024 – 2025

Special Scholarship <i>Awarded by Korea University for academic excellence</i>	Fall 2022, Spring 2023
Semester High Honors <i>Recognized by Korea University for academic excellence</i>	Fall 2018, Spring 2022, Spring 2023
Agricultural Economics Alumni Scholarship <i>Awarded by Korea University, Department of Food and Resource Economics for academic excellence</i>	Spring 2022

Teaching Assistantship

Survival Data Analysis and Lab <i>Advanced Undergraduate Course</i>	Fall 2025
◦ Led hands-on lab sessions on survival analysis and graded assignments and exams.	
Selected Topics Seminar <i>Introductory Undergraduate Course</i>	Spring 2025
◦ Organized weekly discussion sessions on economics and statistics, and advised on data analysis projects.	
Mathematical Statistics 2 <i>Core Undergraduate Course</i>	Fall 2024
◦ Held office hours, graded assignments and exams, and prepared solution sets.	
Statistics Lab <i>Introductory Undergraduate Course</i>	Spring 2024
◦ Evaluated Python programming coursework, and held office hours.	

Projects

Modeling Risk Factors for Mortality and Hospitalization	Sep. 2025
◦ Analyzed risk factors for mortality and hospital stay using GLMM with multiple imputation, addressing repeated events and missing data.	
Weight Design Project for the Longitudinal Survey Panel	Oct. 2024
◦ Constructed stratified sampling weights using the R survey package and advised on missing data handling in the SNU student survey.	
Bitcoin Chart Pattern Image Recognition and Price Prediction Project	May 2022 – Jul. 2022
GitHub Repository	
◦ Implemented Monte Carlo Dropout in the N-BEATS time-series neural network to quantify predictive uncertainty and visualize prediction intervals.	
◦ Augmented chart image data using probability distributions, resulting in a 10% increase in accuracy.	
Optimizing Pricing Strategies for a Low-Demand Food Product	May 2022 – Jun. 2022
◦ Designed an online survey and conducted a conjoint analysis to identify consumer preferences.	
◦ Developed an algorithm in Python to estimate the profit-maximizing bundle price for the products.	
Data Visualization of Job Openings in Korea	Nov. 2021 – Jan. 2022
GitHub Repository (In Korean)	
◦ Extracted and preprocessed 36,000 job postings and 11,000 resumes by identifying HTML patterns.	

Skills & Languages

Software R, Python, LaTeX, SAS, ArcGIS, Stata, SPSS
Languages Fluent in both English and Korean