Jooho Kim

jooho991122 (at) gmail (dot) com Personal Website GitHub

Education

Seoul National University — Seoul, South Korea
MS in Statistics — Mar. 2024 – Feb. 2026

Korea University — Seoul, South Korea
BE in Food and Resource Economics, Double major in Statistics — Mar. 2018 – Feb. 2024

The University of Texas at Austin — Texas, United States
Exchange Program, Economics — 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 — Seoul, South Korea
Graduate Researcher (Advisor: Dr. Yei Eun Shin) — Jun. 2024 – Present

- 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 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 — Austin, United States
Undergraduate Research Assistant (Connected through Dr. Arya Farahi) — Oct. 2022 – Dec. 2022

- Aggregated and cleaned geotagged electric vehicle (EV)-related tweets using regular expressions and bot probability scores.
- Conducted hotspot analysis across the U.S. to identify regions with significant EV-related public sentiment.

Presentation

Scalable and Efficient Multiple Imputation for Case-Cohort Studies via Influence Function-Based Supersampling — Dec. 2025
Korean Statistical Society, Seoul, Korea (Oral Presentation).

Multiple Imputation for Incomplete Survival Data with Missing Covariates: Toward Valid Causal Inference — Jun. 2025
The 2nd Symposium on Causal Inference, Seoul, Korea (Oral Presentation, English).

Honors and Awards

Fellowship for Fundamental Academic Fields — 2024, 2025
Awarded by Seoul National University for academic excellence and research potential

Graduate Research Fellowship in Science and Engineering — 2024 – 2025
Awarded by the National Research Foundation of Korea through a competitive selection process

Special Scholarship Fall 2022, Spring 2023
Awarded by Korea University for academic excellence

Semester High Honors Fall 2018, Spring 2022, Spring 2023
Recognized by Korea University for academic excellence

Agricultural Economics Alumni Scholarship Spring 2022
Awarded by Korea University, Department of Food and Resource Economics for academic excellence

Teaching Assistantship

Survival Data Analysis and Lab Fall 2025
Advanced Undergraduate Course

- Led hands-on lab sessions on survival analysis and graded assignments and exams.

Selected Topics Seminar Spring 2025
Introductory Undergraduate Course

- Organized weekly discussion sessions on economics and statistics, and advised on data analysis projects.

Mathematical Statistics 2 Fall 2024
Core Undergraduate Course

- Held office hours, graded assignments and exams, and prepared solution sets.

Statistics Lab Spring 2024
Introductory Undergraduate Course

- Evaluated Python programming assignments and exams and held office hours.

Employment

Hankuk University of Foreign Studies – Insight Camp Jan. 2019 – Feb. 2019
Academic Mentor

- Provided residential academic mentoring and taught math and English classes in a 4-week program.
- Created problem-solving exercises and offered individualized tutoring.

Projects

Modeling Risk Factors for Mortality and Hospitalization Sep. 2025

- Analyzed clinical risk factors for mortality and hospital stay using GLMM and 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 and visualize predictive uncertainty.
- 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 a Python algorithm to estimate the profit-maximizing bundle price for the food product.

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