Jooho Kim

✉ jooho991122 (at) gmail (dot) com ⌂ Personal Website ⧉ Jooho-Git

Education

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

Research Interest

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

Preprints

Kim, J., Saegusa, T., and Shin, Y. E. (2025). "Robust and Scalable Multiple Imputation for Case-Cohort Studies via Influence Function-based Supersampling." *Preprint.* [arXiv:xxxxx] ⧉	To be submitted

Research Experience

Prediction Modeling 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"*.
- Developed influence function-based supersampling approach to impute only a subset (e.g., 3%) of the missing covariate while preserving efficiency and unbiasedness.
- Devised weight calibration equations that reconcile heterogeneous samping weights for a unified Cox proportional hazards regression analysis.
- Applied the proposed method to NIH-AARP Diet and Health Study, reducing the imputation time by approximately 95% in SMC-FCS method.

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 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

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

Honors and Awards

Next Generation Scholarship for Fundamental Research	2024, 2025
Awarded by Seoul National University for academic and research potential	
Graduate Research Fellowship in Science and Engineering	Sep. 2024 – Aug. 2025
Awarded a research grant by the National Research Foundation of Korea (NRF) through a proposal review process	

Semester High Honors   2018 F, 2022 S, 2023 S
Awarded for achieving a semester GPA greater than 4.0/4.5

Agricultural Economics Alumni Scholarship   2021
Recognized for academic performance

Teaching Assistantship

Survival Data Analysis and Lab   Fall 2025
Advanced Undergraduate Course

- Led hands-on lab session 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 coursework, and held office hours.

Projects

Statistical Consulting: Risk Factors of 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

- Implemented 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 5%p increase in accuracy.

Analysis of Price and Marketing Strategies for a Low-Demand Product   May 2022 – Jun. 2022

- Designed an online survey and conducted a conjoint analysis to identify consumer preference.
- 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 English, Native in Korean