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, Robust Deep Learning

Working Paper

J. Kim, Y-E. Shin. Influence-Based Super-Sampling for Efficient Multiple Imputation in Case–Cohort Studies — In preparation

Research Experience

Prediction Modeling Lab, Seoul National University Seoul, South Korea
Graduate Student Researcher (Advisor: Yei Eun Shin) Jun. 2024 – Present

- Served as a primary graduate researcher on the NRF-funded project: *"Assessing Bias and Efficiency of Imputation Methods for Missing Data due to Epidemiological Cohort Sampling Designs"*.
- Developed influence function-based sampling to impute only a subset (e.g., 10%) of the missing covariate while preserving efficiency and unbiasedness.
- Devised novel weight calibration equations that reconcile heterogeneous samping weights for unified Cox proportional hazards model analysis.
- Applied the proposed methods to NIH-AARP cohort data with over 300,000 records, reducing the imputation time by approximately 95% without loss of statistical validity.

Urban Informatics Lab, The University of Texas at Austin Austin, United States
Undergraduate Research Assistant (Connected through Arya Farahi) Oct. 2022 – Dec. 2022

- Aggregated and processed geotagged electric vehicle tweets using extensive 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.
- Filtered out automated and bot-generated accounts to construct a reliable large-scale dataset.

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 outstanding academic performance and research potential

- Received 23M KRW in total (≈ 17K USD).

NRF Graduate Research Fellowship in Science and Engineering Sep. 2024 - Aug. 2025
Received national fellowship through competitive review process of research proposals

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 outstanding academic performance

Teaching Assistantship

Survival Data Analysis and Lab Fall 2025
Advanced Undergraduate Course

- Led a five hour hands-on lab session on survival analysis and graded assignments.

Selected Topics Seminar Spring 2025
Introductory Undergraduate Course

- Organized weekly discussion sessions on economics and statistics, and advised students on data analysis for poster 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: Prediction of Mortality and Hospitalization Sep. 2025

- Analyzed the impact of clinical covariates on mortality and hospital stay using GLMM with multiple imputation, considering repeated hospitalizations and missingness.

Weight Design Project for the Longitudinal Survey Panel Oct. 2024

- Provided statistical consultation on analyzing SNU students' survey data, with a focus on calibrated design weights and handling missing data.

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.
- Revised the optimization function to address uncertainty quantification issues.
- Augmented chart image data using probability distributions resulting in 5%p increase in accuracy.

Analyzing Price and Marketing Strategies of a Ramen Company May 2022 - Jun. 2022

- Conducted a conjoint analysis to identify the product features most demanded by consumers.
- 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 36K job postings and 11K resumes by identifying html patterns.
- Filtered out 400 stop words using the "term frequency-inverse document frequency" method.

Skills & Languages

Software R, Python, LaTeX, SAS, ArcGIS, STATA, SPSS
Languages Fluent in English, Native in Korean