

Jooho Kim

✉ jooho991122 (at) gmail (dot) com 🏠 [Personal Website](#) 🔗 [Jooho-Git](#)

Education

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

Research Interest

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

Preprints

Kim, J. , Saegusa, T., and Shin, Y. E. (2025). “Robust and Scalable Multiple Imputation for Case-Cohort Studies via Influence Function-based Supersampling.” <i>Preprint.</i> [arXiv:2501.xxxxx] 🔗	To be submitted
--	-----------------

Research Experience

Prediction Modeling 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 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 sampling weights for a unified Cox proportional hazards regression 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 efficiency. 	
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 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

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

Honors and Awards



Next Generation Scholarship for Fundamental Research <i>Awarded by Seoul National University for outstanding academic performance and research potential</i>	2024, 2025
--	------------

Graduate Research Fellowship in Science and Engineering <i>Awarded a research grant by the National Research Foundation of Korea (NRF) through a competitive proposal review process</i>	Sep. 2024 – Aug. 2025
Semester High Honors <i>Awarded for achieving a semester GPA greater than 4.0/4.5</i>	2018 F, 2022 S, 2023 S
Agricultural Economics Alumni Scholarship <i>Recognized for outstanding academic performance</i>	2021

Teaching Assistantship

Survival Data Analysis and Lab <i>Advanced Undergraduate Course</i>	Fall 2025
<ul style="list-style-type: none"> ◦ Led a five hour hands-on lab session on survival analysis and graded assignments. 	
Selected Topics Seminar <i>Introductory Undergraduate Course</i>	Spring 2025
<ul style="list-style-type: none"> ◦ Organized weekly discussion sessions on economics and statistics, and advised students on data analysis for poster projects. 	
Mathematical Statistics 2 <i>Core Undergraduate Course</i>	Fall 2024
<ul style="list-style-type: none"> ◦ Held office hours, graded assignments and exams, and prepared solution sets. 	
Statistics Lab <i>Introductory Undergraduate Course</i>	Spring 2024
<ul style="list-style-type: none"> ◦ Evaluated Python programming coursework, and held office hours. 	

Projects

Statistical Consulting: Risk Factors of Mortality and Hospitalization	Sep. 2025
<ul style="list-style-type: none"> ◦ 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
<ul style="list-style-type: none"> ◦ 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 	
<ul style="list-style-type: none"> ◦ 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. 	
Analysis of Price and Marketing Strategies for a Low-Demand Product	May 2022 – Jun. 2022
<ul style="list-style-type: none"> ◦ 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)	
<ul style="list-style-type: none"> ◦ 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, L^AT_EX, SAS, ArcGIS, STATA, SPSS
Languages Fluent in English, Native in Korean