

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> <i>Military Service, 2020 – 2021</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, Uncertainty Quantification

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. [\[paper\]](#) [\[software\]](#)
Recipient of the 2025 Korean Statistical Society Graduate Student Paper Award.

Research Experience

| | |
|---|--|
| Prediction Model Lab, Seoul National University <i>Graduate Researcher (Advisor: Dr. Yeì 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 Sampling Designs”. ◦ Proposed an influence function-based supersampling approach that imputes only a subset (e.g., 3%) of the missing covariates while preserving efficiency and unbiasedness. ◦ Devised weight calibration equations that integrate distinct sampling designs for a unified 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 (EV)-related tweets across the U.S. using regular expressions and bot probability scores. ◦ Conducted hotspot analysis to identify regions with significant EV-related public sentiment. | |

Contributed Presentations

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

Honors and Awards

| | |
|---|--|
| Graduate Student Paper Award <i>Awarded by the Korean Statistical Society (KSS) for a graduate student paper presentation.</i> | Dec. 2025 |
| 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> <ul style="list-style-type: none">◦ Led hands-on lab sessions on survival analysis and graded assignments and exams. | Fall 2025 |
| Selected Topics Seminar <i>Introductory Undergraduate Course</i> <ul style="list-style-type: none">◦ Organized weekly discussion sessions and advised on data analysis projects. | Spring 2025 |
| Mathematical Statistics 2 <i>Core Undergraduate Course</i> <ul style="list-style-type: none">◦ Held office hours, graded assignments and exams, and prepared solution sets. | Fall 2024 |
| Statistics Lab <i>Introductory Undergraduate Course</i> <ul style="list-style-type: none">◦ Evaluated Python programming assignments and exams and held office hours. | Spring 2024 |

Employment

| | |
|--|-----------------------|
| Hankuk University of Foreign Studies – Insight Camp <i>Academic Mentor</i> <ul style="list-style-type: none">◦ Provided residential academic mentoring and taught math and English classes in a 4-week program.◦ Created problem-solving exercises and offered individualized tutoring. | Jan. 2019 – Feb. 2019 |
|--|-----------------------|

Projects

| | |
|---|----------------------|
| Survey Design and Analysis for the SNU Panel of Youth (SNUPY) <ul style="list-style-type: none">◦ Constructed stratified cross-sectional and longitudinal sampling weights using the R <code>survey</code> package.◦ Advised on missing data imputation and weight calibration. | Oct. 2024; Dec. 2025 |
| Modeling Risk Factors for Mortality and Hospitalization <ul style="list-style-type: none">◦ Analyzed clinical risk factors for mortality and hospital stay using GLMM and multiple imputation, addressing repeated events and missing data. | Sep. 2025 |
| Bitcoin Chart Pattern Image Recognition and Price Prediction GitHub Repository <ul style="list-style-type: none">◦ Implemented Monte Carlo Dropout in the N-BEATS time-series neural network to quantify and visualize predictive uncertainty. | May 2022 – Jul. 2022 |

- 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