Dongsun Yoon

Last updated: October 7, 2026

dsyoon@umich.edu dsyoonstat.github.io Google Scholar

Research Interests

Deep Learning Theory: Diffusion Models, Large Language Models

Statistics: Statistical Learning Theory, Minimax Theory, High-Dimensional Statistics

Education

Ph.D. in Statistics, University of Michigan — **Aug 2025 - Present**
Advisor: Saptarshi Chakraborty — *Ann Arbor, MI*

M.S. in Statistics, Seoul National University — **Mar 2023 - Aug 2025**
Advisor: Sungkyu Jung — *Seoul, South Korea*

B.S. in Statistics, Seoul National University — **Mar 2016 - Feb 2022**
Summa Cum Laude — *Seoul, South Korea*

Publications

Published / Accepted

[2] J. Hwang*, **D. Yoon***, S. Jung, M-J. Park, and I. Yeo. iLBA: An R package for confidentially disseminating aggregated frequency tables. *SoftwareX*, 35, 102880, 2026.

[1] **D. Yoon** and S. Jung. Adaptive Reference-Guided Estimation of Principal Component Subspace in High Dimensions. *Stat*, 14(3), e70081, 2025.

Submitted / In Preparation

[5] **D. Yoon**, J. Li, S. Chakraborty, A. Tewari. Minimax Optimal Prediction of Markov Chains with Effective Support Size. *In Preparation.*

[4] **D. Yoon**, S. Chakraborty. Masking Discrete Diffusion Models are Minimax Optimal for Estimating Distributions with Small Effective Support Size. *In Preparation.*

[3] **D. Yoon**, S. Chakraborty. Uniform Discrete Diffusion Models are Minimax Optimal for Estimating Distributions with Small Effective Support Size. *Submitted.*

* Equal contribution

Awards

Outstanding First-Year Ph.D. Student Award — **Sep 2026**
Department of Statistics, University of Michigan

Overseas PhD Fellowship ($13,000/yr) — **Fall 2025 – Spring 2030**
Korea Foundation for Advanced Studies

Best Poster Award (3rd Place) — **Dec 2024**
The Korean Statistical Society

Merit-Based Scholarship (60% Tuition) — **Spring 2024**
Department of Statistics, Seoul National University

Lecture & Research Scholarship (Full Tuition + $3,000/yr) — **Spring 2023 – Fall 2023**
Department of Statistics, Seoul National University

Dean's List **Aug 2020**
Department of Statistics, Seoul National University

National Science & Technology Scholarship (Full Tuition) **Spring 2020 – Fall 2021**
Korea Student Aid Foundation

Merit-Based Scholarship (Full Tuition) **Fall 2016 – Fall 2017**
Department of Statistics, Seoul National University

Work Experience

Shepherd23 **Mar 2023 – Feb 2025**
Machine Learning Engineer (Intern) *Seoul, South Korea*
- Implemented and maintained bandit algorithm for *Pickhound*, a product recommendation plug-in for eCommerce.

Bank of Korea **Jan 2022 – Mar 2023**
Junior Economist (Full-Time) *Seoul, South Korea*
- Produced GDP statistics of South Korea, specializing in financial and agricultural industries.
- Reconstructed *ECOS* (South Korean equivalent of *FRED*).

Teaching Experience

Department of Statistics, University of Michigan **Fall 2025 – Present**
Graduate Student Instructor *Ann Arbor, MI*
- Regression Analysis (STATS 500) [Office Hours/Grading], Fall 2026
- Introduction to Statistics and Data Analysis (STATS 250) [Lab Instruction/Office Hours/Grading], Winter 2026
- Introduction to Probability (MATH/STATS 425) [Office Hours/Grading], Fall 2025

Department of Statistics, Seoul National University **Spring 2023 – Spring 2024**
Teaching Assistant *Seoul, South Korea*
- First Year Graduate Student Seminar (M0000.008700) [Lab Instruction/Office Hours/Grading], Spring 2024
- Survival Analysis and Lab (326.412) [Lab Instruction/Office Hours/Grading], Fall 2023
- Regression Analysis and Lab (326.313) [Lab Instruction/Office Hours/Grading], Spring 2023

Presentations

Contributed

Statistical Guarantees of Uniform Discrete Diffusion Models with Effective Support Size (Talk), *2026 SIAM Student Mini-Symposium in Applied Mathematics*, Ann Arbor, MI, Aug 2026.

Augmented Estimation of Principal Component Subspace in High Dimensions (Talk), *International Symposium on Theories, Methodologies and Applications for Large Complex Data*, Tsukuba, Japan, Dec 2024.

Augmented Estimation of Principal Component Subspace in High Dimensions (Poster), *The Korean Statistical Society Conference*, Daejeon, South Korea, Nov 2024.

Invited

Software

iLBA [2]: R package for *Information Loss Bounded Aggregation*, a method for confidentially disseminating aggregated frequency tables. Adopted by *Statistics Korea* to publish census data.

argpca [1]: Python package for *Adaptive Reference-Guided PCA*, also known as *James-Stein PCA*.

Service

Reviewer: *SoftwareX*

Skills

Programming: Python, R, MATLAB, LaTeX

Language: Korean (Native), English (Fluent)