

# KYUSEONG CHOI



Cornell Tech



kyuseongchoi5.github.io

kc728@cornell.edu



EDUCATION	<b>Ph.D., Statistics and Data Science</b> , Cornell Tech Advisors: Prof. Raaz Dwivedi & Prof. Kengo Kato	2021–
	<b>M.S. Biostatistics</b> , University of Michigan, Ann Arbor Advisors: Prof. Jeremy Taylor & Prof. Peisong Han	2021
	<b>B.B.A. Business Administration</b> , Korea University, Seoul, Korea Military service (Korean Air Force) 2014 – 2016	2019
RESEARCH INTERESTS	Post-training large language models, Foundational models for matrix completion, Discrete diffusion, Distributional / personalized causal inference, High-dimensional CLTs, Kernel methods	
EXPERIENCE	<b>Applied Scientist intern</b> , Amazon P2 Science Research: Kernel uniform inference for conditional average treatment effect under endogeneity	2025.05-08
WORK IN PROGRESS	R1. <b>Kyuseong Choi</b> , Dwaipayan Saha, Woojeong Kim, Anish Agarwal, Raaz Dwivedi, “GRPO++: Improved policy optimization using ranked rewards”. R2. <b>Kyuseong Choi</b> , Kengo Kato, “Gaussian approximation in Reproducing Kernel Hilbert spaces, applications to Kernel Ridge Regression”.	2025
PREPRINTS	P1. <b>Kyuseong Choi</b> , Gabriela Ciolek, “Wild regenerative block bootstrap for Harris recurrent Markov chains”, <i>arxiv</i> . P2. Jacob Feitelberg, Dwaipayan Saha, <b>Kyuseong Choi</b> , Zaid Ahmad, Anish Agarwal, Raaz Dwivedi, “TabImpute: Accurate and Fast Zero-Shot Missing-Data Imputation with a Pre-Trained Transformer”, <i>arxiv</i> . P3. <b>Kyuseong Choi</b> , Jacob Feitelberg, Anish Agarwal, Raaz Dwivedi, “Learning counterfactual distributions via kernel methods”, <i>arxiv</i> . P4. Jacob Feitelberg, <b>Kyuseong Choi</b> , Anish Agarwal, Raaz Dwivedi, “Distributional matrix completion via nearest neighbors in the Wasserstein space”, <i>arxiv</i> .	2025
WORKSHOPS	W1. Caleb Chin, Aashish Khubchandani, Harshvardhan Maskara, <b>Kyuseong Choi</b> , Jacob Feitelberg, Albert Gong, Manit Paul, Tathagata Sadhukhan, Anish Agarwal, Raaz Dwivedi, “N <sup>2</sup> : A Unified Python Package and Test Bench for Nearest Neighbor-Based Matrix Completion”, <i>ICML CODEML</i> .	2025
PUBLICATIONS	J1. Albert Gong, <b>Kyuseong Choi</b> , Raaz Dwivedi, “Supervised kernel thinning”, <i>Neurips</i> . J2. <b>Kyuseong Choi</b> , Jeremy Taylor, Peisong Han, “Robust data integration from multiple external sources for generalized linear models with binary outcomes”, <i>Biometrics</i> . J3. Jeremy Taylor, <b>Kyuseong Choi</b> , Peisong Han, “Data integration: exploiting ratios of parameter estimates from a reduced external model”, <i>Biometrika</i> .	2024

CONTRIBUTED  
PRESENTATIONS

**Uniform inference for stationary distribution of diffusion process: wild regenerative block bootstrap for Harris recurrent Markov chains**

- 2025 JSM, Nashville Aug 2025
- 2025 INFORMS, Atlanta Oct 2025

**Test bench and unified python package for Nearest Neighbor-based matrix completion**

- 2025 ICML CODEML Workshop, Seattle July 2025

**Supervised kernel thinning**

- 2024 Neurips, Vancouver Dec 2024

**Learning counterfactual distributions via kernel methods**

- 2025 ICCOPT, LA July 2025
- 2024 INFORMS, Seattle Oct 2024
- 2025 SCI, ACIC, Detroit May 2025

**Gaussian approximation in Reproducing Kernel Hilbert spaces**

- 2024 Workshop on Translational Research on Data Heterogeneity, St. Louis Apr 2024

**Robust data integration in generalized linear models**

- 2023 JSM, Toronto Aug 2023

TEACHING  
EXPERIENCE

- T1. TA: Statistical principles for data science (ORIE 5160), *Cornell Tech.* 2024
- T2. TA: Linear algebra, Operations research and financial engineering, Theoretical statistics (MATH 1710, ORIE 5740, STSCI 4090), *Cornell University.* 2021 - 2024
- T3. TA: Intro to Regression, Intro to Biostatistics (BIOS 502, 521), *University of Michigan, Ann Arbor.* 2019–2020

SELECTED  
ACHIEVEMENTS

- Korea Foundation for Advanced Studies (KFAS) Program Scholarship in Business Administration Studies 2017 – 2018

ADDRESS, CONTACT  
INFO.

Mailing address : 1 E Loop Rd, 10044, NY  
Phone : 734-249-4777