

# YOUNG-GEUN KIM

[younggeun91@unist.ac.kr](mailto:younggeun91@unist.ac.kr)

<https://kyg0910.github.io/>

<https://scholar.google.com/citations?user=HVqiptEAAAAJ>

## RESEARCH INTERESTS

---

My research interests revolve around developing innovative data science tools and promoting their dissemination on biomedical data. Research topics include, but are not limited to:

- Deep generative models for multi-modal biomedical data (e.g., neuroimaging and multi-omics)
- Deep learning for identifying biomarkers associated with mental illness
- Reinforcement learning-based health care

## PROFESSIONAL APPOINTMENTS

---

**Assistant Professotr** *Feb. 2026 - Present*

Department of Mathematical Sciences, Ulsan National Institute of Science and Technology (UNIST)

**Ast. Prof-Hannan Endow Visiting Scholar** *Aug. 2024 - Jan. 2026*

Department of Statistics and Probability, Michigan State University

## EDUCATION & TRAINING

---

**Adjunct Associate Research Scientist** *Jul. 2021 - Aug. 2024*

Department of Biostatistics, Columbia University

**Mentor:** Ying Liu, Ph.D.

**Postdoctoral Researcher** *Jul. 2021 - Aug. 2024*

Department of Psychiatry, Columbia University

Mental Health Data Science, New York State Psychiatric Institute

**Mentor:** Ying Liu, Ph.D.

**Postdoctoral Researcher** *Mar. 2021 - Jun. 2021*

Department of Statistics, Seoul National University

**Mentor:** Myunghee Cho Paik, Ph.D.

**Seoul National University** *Mar. 2015 - Feb. 2021*

Ph.D. in Statistics

Graduated with the Best Dissertation Award

**Advisor:** Myunghee Cho Paik, Ph.D.

**Dissertation:** Statistical distance of conditional distributions and its applications

**Seoul National University** *Mar. 2010 - Feb. 2015*

**Triple Major**

Graduated with Honors (Cum Laude)

B.S. in Industrial Engineering

B.S. in Statistics

B.S. in Mathematical Sciences

## HONORS & AWARDS

---

**Career Development Award**

*Dec. 2023*

Korean International Statistical Society

**Outstanding Reviewer Award**

Jul. 2022

Thirty-ninth International Conference on Machine Learning

**Best Dissertation Award**

Feb. 2021

College of Natural Sciences, Seoul National University

**Seoul National University Innovation Program Scholarship**

Mar. 2017 - Feb. 2018

Seoul National University

**1st Prize, Student Paper Competition**

June 2017

Korean Statistical Society

**PUBLICATIONS & PREPRINTS**

---

\*: First author; ‡: Corresponding author

**Journal**

1. Kim, S.\*, **Kim, Y.-G.**, and Wang, Y.‡ (2024). Temporal generative models for learning heterogeneous group dynamics of ecological momentary data. *Biometrics*. [\[Paper\]](#) [\[Code\]](#)
2. **Kim, Y.-G.**\*, Ravid, O.\*, Zheng, X., Kim, Y., Neria, Y., Lee, S., He, X.‡, and Zhu, X.‡ (2024). Explaining deep learning-based representations of resting state functional connectivity data: focusing on interpreting nonlinear patterns in autism spectrum disorder. *Frontiers in Psychiatry, section Computational Psychiatry*. [\[Paper\]](#) [\[Code\]](#)
3. **Kim, Y.-G.**\*, Lee, K., and Paik, M.C.‡ (2022). Conditional Wasserstein generator. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. [\[Paper\]](#) [\[Code\]](#)
4. **Kim, Y.-G.**\*, Kwon, Y., and Paik, M.C.‡ (2019). Valid oversampling schemes to handle imbalance. *Pattern Recognition Letters*, 125 (1): 661-667. [\[Paper\]](#) [\[Code\]](#)

**Peer-reviewed Conference**

1. **Kim, Y.-G.**\*, Hu, M.-C., Nunes, E. V., Luo, S. X.‡, and Liu, Y.‡ (2025). Optimizing contingency management with reinforcement learning. *IEEE International Conference on Healthcare Informatics*. Selected as a long presentation [\[Paper\]](#) [\[Code\]](#)
2. Yu, W.\*, Qu, G., **Kim, Y.-G.**, Xu, L., and Zhang, A.‡ (2025). A Novel GNN Framework Integrating Neuroimaging and Behavioral Information to Understand Adolescent Psychiatric Disorders. *Medical Imaging with Deep Learning*. [\[Paper\]](#) [\[Code\]](#)
3. **Kim, Y.-G.**\*, Liu, Y.‡, and Wei, X. (2023). Covariate-informed representation learning to prevent posterior collapse of iVAE. *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics (AISTATS 2023)*. [\[Paper\]](#) [\[Code\]](#)
4. Kim, M.\*, **Kim, Y.-G.**, Kim, D., Kim, Y., and Paik, M.C.‡ (2021). Kernel-convoluted deep neural networks with data augmentation. *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI 2021)*. [\[Paper\]](#) [\[Code\]](#)
5. **Kim, Y.-G.**\*, Kwon, Y., Chang, H., and Paik, M.C.‡ (2020). Lipschitz continuous autoencoders in application to anomaly detection. *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics (AISTATS 2020)*. [\[Paper\]](#) [\[Code\]](#)

**Patents**

1. Paik, M.C.‡, **Kim, Y.-G.**, and Lee, K. (2024). Method and apparatus for conditional data generation using conditional Wasserstein generator. *KR Patent 102734936B1*. [\[Info\]](#)
2. Paik, M.C.‡, **Kim, Y.-G.**, and Chang, H. (2021). Learning method and learning device for high-dimension unsupervised anomaly detection using kernalized Wasserstein autoencoder to lessen too

many computations of Christophel function, and testing method and testing device using the same.  
*KR Patent 102202842B1*. [\[Info\]](#)

## Preprints

1. **Kim, Y.-G.\***, Lee, K., Choi, Y., Won, J.-H., and Paik, M.C.<sup>‡</sup> (2023). Wasserstein geodesic generator for conditional distributions (under *Major Revision* at Journal of Machine Learning Research). [\[ArXiv\]](#)[\[Code\]](#)
2. Yang, B.\*, **Kim, Y.-G.**, and Wang, Y.<sup>‡</sup> Deep Representation Learning for Optimizing Treatment Decisions with Electroencephalogram Biomarkers. Submitted to Biometrics.  
- This work was selected as the Runner-up in the student paper competition for the Statistics in Imaging Section of the ASA in 2025.
3. Zheng, X.\*, Ravid, O., Barry, R. A.J., Kim, Y., Wang, Q., **Kim, Y.-G.**, Zhu, X.<sup>‡</sup> and He, X.<sup>‡</sup> (2024). Denoising Variational Autoencoder as a Feature Reduction Pipeline for the diagnosis of Autism based on Resting-state fMRI. [\[ArXiv\]](#)

## Work in Progress

1. **Kim, Y.-G.\*<sup>‡</sup>** and Liu, Y. Mid-VAE: Multi-modal, Identifiable, and Disentangled Representation Learning for Children's Structural Brain Imaging. Work in progress  
- Preliminary results were presented at ABCD AIIM conference.

## GRANTS & FUNDING

---

### Active

#### USA-Funded

- **Statistical understanding of adversarial training in neural networks** *Aug. 2025 – Jul. 2028*  
\$180,000  
National Science Foundation DMS PD 18-1269  
Role: **Co-PI** (Aug. 2025 – Jan. 2026); **International Consultant** (Jan. 2026 – Jul. 2028).  
Role transitioned due to relocation to UNIST, South Korea (**PI**: Dr. Yue Xing, Michigan State University)

### Past

#### USA-Funded

- **A data science framework for empirically evaluating and deriving reproducible and transferrable RDoC constructs in youth (R01)** *Jul. 2021 - Aug. 2024*  
National Institutes of Health NIMH  
Role: **Postdoctoral Researcher**

#### Republic of Korea-Funded

- **Deep learning with incomplete and sequential data: Application to biomedical data** *Mar. 2020 - Jun. 2021*  
National Research Foundation  
Role: **Graduate Student**
- **Development of low-yield trackers via causal inference** *May 2019 - Nov. 2019*  
SK Telecom  
Role: **Graduate Student**

- **Statistical approaches to deep learning: New methods for convolutional neural networks in application to medical imaging data** *Mar. 2017 - Feb. 2020*  
National Research Foundation  
Role: **Graduate Student**
- **Deep Learning for the CT based Acute Cerebral Infarction Classification and Lesion Segmentation** *July 2016 - May 2019*  
National Research Foundation  
- Collaborated with Seoul National University Bundang Hospital  
Role: **Graduate Student**
- **New Robust Methods for Missing or Censored Covariates** *Mar. 2016 - Nov. 2016*  
National Research Foundation  
Role: **Graduate Student**

Under Review

## SELECTED TALKS

---

### Invited Talks

- (Scheduled) **Kim, Y.-G.** (2026). Multi-Modal Representation Learning with Partially Pairwise Observations. *The 2026 International Chinese Statistical Association (ICSA) China Conference, Shenzhen, China.*
- **Kim, Y.-G.**, Kwon, Y., Chang, H., and Paik, M.C. (2019). Lipschitz continuous autoencoders in application to anomaly detection. *IMS-China International Conference on Statistics and Probability, Dalian, China.*
- **Kim, Y.-G.**, Lee, K., Choi, Y., Won, J.-H., and Paik, M.C. (2025). Wasserstein geodesic generator for conditional distributions. *The 2025 International India Statistical Association (IISA), Lincoln, NE.*
- **Kim, Y.-G.**, Luo, S. X., Brandt, L., Cheung, K., Nunes, E. V., Roll, J., and Liu, Y. (2024). Optimizing contingency management interventions in substance use disorder treatment with reinforcement learning. *The Joint Statistical Meetings (JSM), Portland, OR.*
- **Kim, Y.-G.** and Liu, Y. (2024). Deep Identifiable Generative Models for Multi-Modal Data Analysis. *The 2024 International Chinese Statistical Association (ICSA) Applied Statistics Symposium, Nashville, TN.*
- **Kim, Y.-G.**, Kwon, Y., Chang, H., and Paik, M.C. (2019). Lipschitz continuous autoencoders in application to anomaly detection. *IMS-China International Conference on Statistics and Probability, Dalian, China.*

### Contributed Talks

- **Kim, Y.-G.**, Liu, Y., and Wei, X. (2023). Covariate-informed representation learning to prevent posterior collapse of iVAE. *The Twenty Third International Conference on Artificial Intelligence and Statistics (AISTATS 2023), Palau de Congressos, Valencia, Spain.*<sup>†</sup>
- **Kim, Y.-G.**, Kwon, Y., Chang, H., and Paik, M.C. (2020). Lipschitz continuous autoencoders in application to anomaly detection. *The 23rd International Conference on Artificial Intelligence and Statistics (AISTATS 2020), Virtual conference due to COVID-19.*
- **Kim, Y.-G.** and Liu, Y. (2024). Explaining Nonlinear Patterns in Children's Structural MRI with Multi-modal Identifiable VAE. *The ABCD Insights & Innovations Meeting, MD.*<sup>†</sup>

- **Kim, Y.-G.**, Liu, Y., Brandt, L., Cheung, K., Nunes, E. V., Roll, J., and Luo, S. X. (2023). Optimizing contingency management in substance use disorder treatment using off-policy policy evaluation. *Eastern North American Region (ENAR) 2023 Spring meeting*.

† indicates a poster presentation.

## TEACHING EXPERIENCE

---

### Instructor

- **MTH344: Mathematical Statistics**
- **STT441 Sections 1 and 2: Probability and Statistics I: Probability** *Fall 2025*  
Michigan State University  
- An undergraduate-level course covering topics such as normal approximation, sampling distributions, parameter estimation, and elementary tests of hypotheses. [\[Syllabus\]](#)
- **STT351: Probability and Statistics for Engineering** *Summer 2025*  
Michigan State University  
- An undergraduate-level course covering topics such as probability models and random variables, estimation, confidence intervals, hypothesis testing, simple linear regression, and applications to engineering. [\[Syllabus\]](#)
- **STT890: Statistical Problems** *Summer 2025*  
Michigan State University  
- A graduate-level course for individualized study on selected problems
- **STT997: Advanced Topics in Statistics** *Spring 2025*  
Michigan State University  
- **Developed a new course** [\[CourseMaterial\]](#)  
- A graduate-level course covering recent topics in deep generative models and their applications

### Guest Lecturer

- **Statistical Practice and Research for Interdisciplinary Sciences (SPRIS)** *Spring 2025*  
Columbia University  
- Graduate-level course on interdisciplinary research topics in Biostatistics.  
- Gave the lecture “Variational Autoencoders and Their Applications to Multi-modal Data Analysis.”
- **STT990: Statistics & Probability** *Fall 2024*  
Michigan State University  
- Graduate-level seminar course.  
- Gave the lecture “Deep Generative Model: A Statistical Perspective.”
- **Statistical Practice and Research for Interdisciplinary Sciences (SPRIS)** *Spring 2024*  
Columbia University  
- Graduate-level course on interdisciplinary research topics in Biostatistics.  
- Gave the lecture “Recent Topics on Conditional Generative Models.”
- **Deep Learning: A Statistical Perspective** *Fall 2021*  
Seoul National University  
- Graduate-level course on deep learning.  
- Gave the lecture “Conditional Image Synthesis and Its Applications” in English.

## MENTORSHIP EXPERIENCE

---

## Co-mentoring Graduate Students at Columbia University

- Bin Yang, Ph.D. Candidate, Department of Biostatistics April. 2024 - Present
  - Conducted regular weekly meetings with Dr. Yuanjia Wang.
  - Provided mentorship on the following work:  
Yang, B., **Kim, Y.-G.**, and Wang, Y. Deep Representation Learning for Optimizing Treatment Decisions with Electroencephalogram Biomarkers. Submitted to Biometrics.
  - This work was selected as the Runner-up in the student paper competition for the Statistics in Imaging Section of the ASA in 2025.
- Soohyun Kim, Ph.D., Department of Biostatistics Mar. 2022 - Sep. 2024
  - Conducted regular weekly meetings with Dr. Yuanjia Wang.
  - Provided mentorship on the doctoral dissertation and the following paper:  
Kim, S., **Kim, Y.-G.**, and Wang, Y. (2024). Temporal generative models for learning heterogeneous group dynamics of ecological momentary data. *Biometrics*. [\[Paper\]](#) [\[Code\]](#)
- Zekai Jin, Master Student, Department of Biostatistics Dec. 2022 - Oct. 2023
  - Conducted regular bi-weekly meetings with Dr. Seonjoo Lee.
  - Provided mentorship on deep learning-based EEG denoising methods.

## Doctoral Dissertation Committee

- Bin Yang, Department of Biostatistics, Columbia University
  - Role: Committee member

## OTHER PROFESSIONAL ACTIVITIES

---

### Event Chairing

- (Scheduled) Invited Session at 2026 International Chinese Statistical Association-China Conference (role: **Speaker**); Title: Multi-Modal Representation Learning with Partially Pairwise Observations.
- Invited Session at East Asia Chapter of International Society for Bayesian Analysis 2025 (role: **Chair**); Title: Recent Developments in a Bayesian Framework
- Invited Session at 2025 International Indian Statistical Association (role: **Speaker**); Title: Wasserstein Geodesic Generator for Conditional Distributions
- Invited Session at Joint Statistical Meetings 2024 (role: **Organizer & Speaker**); Title: Reliable and Cost-effective Mental Health Care with Reinforcement Learning [\[Info\]](#)
- Invited Session at 2024 International Chinese Statistical Association (role: **Speaker**); Title: Recent Advances in Precision Medicine and Adaptive Experiments [\[Info\]](#)
- Invited Session at Eastern North American Region 2023 (role: **Chair**); Title: Advanced Methods for Analyzing Large-Scale Neuroimaging Data from Nationwide Consortia for Mental Health Research [\[Info\]](#)
- Oral Presentation Session at International Conference on Machine Learning 2022 (role: **Chair**); Title: Theory [\[Info\]](#)
- Invited Session at IMS-China International Conference on Statistics and Probability 2019 (role: **Speaker**); Title: Lipschitz continuous autoencoders in application to anomaly detection.

### Reviewer (Journal)

- Expert Systems with Applications (42 submissions)

- JAMA Psychiatry (2 submissions)
- Biostatistics (1 submission)
- Physica A: Statistical Mechanics and its Applications (1 submission)
- Statistics and Data Science in Imaging (2 submissions)
- International Journal of Computer Assisted Radiology and Surgery (3 submissions)
- Journal of the Korean Statistical Society (2 submissions)
- Journal of Computing Science and Engineering (2)

**Reviewer (Conference)**

- International Conference on Machine Learning 2022
- International Conference on Artificial Intelligence and Statistics 2022, 2023, and 2026