

SHANGHONG XIE

Department of Statistics, University of South Carolina
LeConte College, Columbia, SC 29208
SX2@mailbox.sc.edu
Personal website: <https://shanghongxie.github.io/>

EDUCATION

Columbia University in the City of New York Ph.D. in Biostatistics <i>Dissertation Title:</i> Statistical Methods for Constructing Heterogeneous Biomarker Networks <i>Advisor:</i> Yuanjia Wang	2019
University of Illinois at Urbana-Champaign M.S. in Statistics	2014
Sichuan University, Chengdu, China B.S. with highest honors in Statistics	2012

PROFESSIONAL EXPERIENCE

• Assistant Professor, <i>Department of Statistics</i> , University of South Carolina	2024 – Present
• Member, <i>Carolina Autism and Neurodevelopment (CAN) Research Center</i> , University of South Carolina	2024 – Present
• Member, <i>Institute for Mind and Brain (IMB)</i> , University of South Carolina	2024 – Present
• Associate Professor, <i>Department of Data Science</i> , School of Statistics, Southwestern University of Finance and Economics	2023 – 2024
• Assistant Professor, <i>Department of Data Science</i> , School of Statistics, Southwestern University of Finance and Economics	2021 – 2023
• Member, <i>Center of Statistical Research</i> , Southwestern University of Finance and Economics	2021 – 2024
• Postdoctoral Research Scientist, <i>Department of Biostatistics</i> , Mailman School of Public Health, Columbia University	2019 – 2021

RESEARCH INTERESTS

Statistical machine learning; network analysis; graphical model; precision medicine; functional data analysis; causal inference; mediation analysis; variable selection; high dimensional analysis; neuroimaging; biomarker; neurological and psychiatric diseases; mental health; COVID-19

AWARDS AND HONORS

• American Statistical Association (ASA) Mental Health Statistics Section Best Student Paper Award	2020
• International Conference on Health Policy Statistics (ICHPS) Student Travel Award	2018
• NYC Datathon (Data Science Competition) 1st Place Winning Team, among 1000+ participants	2017
• Columbia University Fellowship	2014 – 2017

PUBLICATIONS

An asterisk (*) is used to indicate corresponding author; An underline is used to indicate students under my supervision; A dagger (†) is used to indicate joint first author, equal contribution.

Peer-Reviewed Journal and Conference Articles

1. McDonnell, E., **Xie, S.**, Marder, K., Cui, F., and Wang, Y. (2024). Dynamic Undirected Graphical Models for Time-Varying Clinical Symptom and Neuroimaging Networks. *Statistics in Medicine* 43 (1), 4131-4147. (**An earlier version won ASA Statistics in Imaging Section First Prize Student Paper Award**)
2. Shi, B., Liu, Y., **Xie, S.**, Zhu, X., and Wang, Y. (2024). Network-Assisted Mediation Analysis with High-Dimensional Neuroimaging Mediators. Machine Learning for Healthcare Conference. *Proceedings of Machine Learning Research* 252, 1-22.
3. **Xie, S.***, Zeng, D., and Wang, Y. (2024). Identifying Temporal Pathways Using Biomarkers in the Presence of Latent Non-Gaussian Components. *Biometrics* 80 (2), ujae033.
4. **Xie, S.*** and Ogden, R. T. (2024). Functional Support Vector Machine. *Biostatistics* 25 (4), 1178-1194.
5. **Xie, S.***, Tarpey, T., Petkova, E., and Ogden, R. T. (2022). Multiple Domain and Multiple Kernel Outcome-weighted Learning for Estimating Individualized Treatment Regimes. *Journal of Computational and Graphical Statistics* 31 (4), 1375-1383.
6. **Xie, S.***, Wang, W., Wang, Q., Wang, Y., and Zeng, D. (2022). Evaluating Effectiveness of Public Health Intervention Strategies for Mitigating COVID-19 Pandemic. *Statistics in Medicine* 41 (9), 3820-3836.
7. COVID-19 Forecast Hub Consortium (2022). Evaluation of Individual and Ensemble Probabilistic Forecasts of COVID-19 Mortality in the US. *Proceedings of the National Academy of Sciences* 119 (15), e2113561119.
8. COVID-19 Forecast Hub Consortium (2022). The United States COVID-19 Forecast Hub Dataset. *Scientific Data* 9, 462.
9. **Xie, S.***, McDonnell, E., and Wang, Y. (2022). Conditional Gaussian Graphical Model for Estimating Personalized Disease Symptom Networks. *Statistics in Medicine* 41 (3), 543-553. (**An earlier version won ASA Mental Health Statistics Section Best Student Paper Award**)
10. **Xie, S.***, Zeng, D., and Wang, Y. (2021). Integrative Network Learning for Multi-modality Biomarker Data. *Annals of Applied Statistics* 15 (1), 64-87.
11. **Xie, S.***, Li, X., McColgan, P., Scahill, R. I., Zeng, D., and Wang, Y. (2020). Identifying Disease-associated Biomarker Network Features Through Conditional Graphical Model. *Biometrics* 76 (3), 995-1006. (**Cover story of *Biometrics* September 2020 issue; An earlier version won the International Conference on Health Policy Statistics (ICHPS) Student Travel Award**)
12. Goldman, J., **Xie, S.**, Green, D., Naini, A., Mansukhani, M. M., and Marder, K. (2021). Predictive Testing for Neurodegenerative Diseases in the Age of Next-generation Sequencing. *Journal of Genetic Counseling* 30, 553-562.
13. Wang, Q., **Xie, S.**, Wang, Y., and Zeng, D. (2020). Survival-Convolution Models for Predicting COVID-19 Cases and Assessing Effects of Mitigation Strategies. *Frontiers in Public Health* 8, 325. (**Our model was used by the Center of Disease Control and Prevention (CDC) for COVID-19 Ensemble Forecast; Our forecasts website: https://github.com/COVID19BIOSTAT/covid19_prediction; CDC ensemble forecast website: <https://www.cdc.gov/coronavirus/2019-ncov/covid-data/forecasting-us.html>**)

14. Li, X.[†], **Xie, S.[†]**, McColgan, P., Tabrizi, S. J., Scahill, R. I., Zeng, D., and Wang, Y. (2018). Learning Subject-Specific Directed Acyclic Graphs with Mixed Effects Structural Equation Models from Observational Data. *Frontiers in Genetics* 9, 430.
15. Li, X., **Xie, S.**, Zeng, D., and Wang, Y. (2018). Efficient l_0 -norm Feature Selection Based on Augmented and Penalized Minimization. *Statistics in Medicine* 37 (3), 473-486.
16. Avissar, M.[†], **Xie, S.[†]**, Vail, B., Lopez-Calderon, J., Wang, Y., and Javitt, D. C. (2018). Meta-analysis of Mismatch Negativity to Simple versus Complex Deviants in Schizophrenia. *Schizophrenia Research* 19, 25-34.

GRANT SUPPORT

- **Principal Investigator** 01/2023 – 07/2024
 “Learning Temporal Causal Network from Biomarker Time Series Data”
 National Natural Science Foundation of China (NSFC), Department of Mathematical and Physical Sciences,
 Grant No 12201511. Funding rate 17%
- **Principal Investigator** 01/2022 – 12/2022
 “Dynamic Network Learning Using Neuroimaging Data”
 Southwestern University of Finance and Economics Startup Grant

COLLABORATIVE RESEARCH EXPERIENCE

New York State Psychiatric Institute (NYSPI) <i>Role: Statistician</i>	2016 – 2021
Columbia University Vagelos College of Physicians and Surgeons, Department of Neurology <i>Role: Statistician</i>	2019 – 2021
University College London, Institute of Neurology <i>Role: Collaborator</i>	2016 – 2021
Columbia University, School of Social Work <i>Role: Statistician</i>	2016

TEACHING EXPERIENCE

University of South Carolina, Department of Statistics <i>Instructor</i>	2024 – Present
Classroom Teaching	
• STAT 530, Applied Multivariate Statistics and Data Mining	Spring 2025
• STAT 509, Statistics for Engineers	Fall 2024
Computing Session Seminar	
• Computing Session 4: Using GitHub and Git	Fall 2024
Functional and Complex Data Analysis (FUNCODA) Working Group	2024 – Present
• Co-founder with Yuexuan Wu and William Consagra	
Southwestern University of Finance and Economics, School of Statistics <i>Instructor</i>	2022 – 2024

- Reading & Writing Scientific Articles: Advanced Topics in Machine Learning (*Graduate*) 11 students enrolled Spring 2024
- Machine Learning and Data Mining (*Undergraduate*) 72 students enrolled Fall 2023
- An Introduction to Machine Learning (*Graduate*) 63 students enrolled Fall 2023
- Reading & Writing Scientific Articles: Advanced Topics in Machine Learning (*Graduate*) 12 students enrolled Spring 2023
- Machine Learning and Data Mining (*Undergraduate*) 40 students enrolled Spring 2023
- Machine Learning and Data Mining (*Undergraduate*) 2 classes, 121 students enrolled in total Fall 2022
- Reading & Writing Scientific Articles: Advanced Topics in Machine Learning (*Graduate*) 8 students enrolled Spring 2022
- Machine Learning and Data Mining (*Undergraduate*) 2 classes, 132 students enrolled in total Spring 2022

Massive Open Online Course (MOOC: XuetangX) 2023
Co-Instructor (Graduate Course)

- Machine Learning

Columbia University, Department of Biostatistics 2015 - 2018
Co-Instructor (Graduate Course)

- Statistical Collaboration for Interdisciplinary Research Spring 2018

Teaching Assistant (Graduate Course)

- Randomized Clinical Trial II Fall 2016
- Generalized Linear Models Fall 2016
- Design of Medical Experiments Spring 2016
- Analysis of Longitudinal Data Fall 2015

University of Illinois at Urbana-Champaign, Department of Statistics 2013 – 2014
Teaching Assistant (Graduate Course)

- Sampling and Categorical Data Spring 2014
- Applied Multivariate Analysis Fall 2013

MENTORING ACTIVITIES

University of South Carolina, Department of Statistics
PhD Dissertation Advisor

- Sehun An: PhD student (co-advisor with Ting Fung Ma) 2025 – Present

PhD Dissertation Committee Member

- Jiasong Duan: PhD student 2025 – Present
- Sijian Fan: PhD student 2024 – Present

Southwestern University of Finance and Economics, School of Statistics

Thesis Advisor

• Hexuan Song (Master thesis)	2022 – Present
• Guishan Xiang (Master thesis)	2022 – 2024
• Anqi Hua (Bachelor thesis)	2023 – 2024
• Yi Jiang (Bachelor thesis)	2023 – 2024
• Lu Li (Bachelor thesis)	2023 – 2024
• Lijuan Guo (Bachelor thesis)	2022 – 2023
• Silu Liu (Bachelor thesis)	2022 – 2023
• Qi Yang (Bachelor thesis)	2022 – 2023
• Yangjie Yin (Bachelor thesis)	2022 – 2023
• Ruiying Li (Bachelor thesis)	2021 – 2022
• Mengjie Li (Bachelor thesis)	2021 – 2022
• Xing Wang (Bachelor thesis)	2021 – 2022
• Qiyu Wang (Bachelor thesis)	2021 – 2022

Academic Advisor

• Supervised over 40 undergraduate students, received excellent evaluations	2022 – 2024
---	-------------

Columbia University, Department of Biostatistics

Mentor

• Zexi Cai (PhD student) PhD dissertation projects	2022 – Present
• Bin Yang (Master student) Research project	2021 – 2021
• Erin Mcdonnell (PhD student, Advisor: Professor Yuanjia Wang) Dissertation projects	2019 – 2021
• Bihui Sun (Master student, Advisor: Professor Yuanjia Wang) Practicum project	2019 – 2020

ACADEMIC SERVICES

Editorial Board

• <i>Frontiers in Public Health</i>	2023 – Present
-------------------------------------	----------------

Journal and Conference Review

- *Annals of Applied Statistics*
- *Biometrics*
- *Briefings in Bioinformatics*
- *BMC Bioinformatics*
- *Computational Statistics and Data Analysis*
- *Journal of the American Statistical Association*
- *Scientific Reports*
- *Statistics in Medicine*

- *Statistics in Biosciences*
- *Statistical Methods in Medical Research*
- *Stat*
- *IEEE International Conference on Bioinformatics and Biomedicine*

Grant Review

- Full-time Member, 2024 – Present
NIH/NINDS Huntington’s Disease Biospecimen Resource Access Committee (HD-BRAC)

Conference Service

- Member, Program Committee, 2024 – Present
2025 International Chinese Statistical Association (ICSA) China Conference
- Chair, Invited session “New developments in the frontiers of precision medicine and data science”, 2024
7th International Conference on Econometrics and Statistics (EcoSta)
- Member, Student Paper Competition Committee for ASA Mental Health Statistics Section 2023 – 2024
- Organizer, Invited session “Topics in healthcare and biostatistics”, R conference in China 2023
- Organizer, Invited session “Novel machine learning methods to advance precision medicine using
big biomarker data”, ICSA China 2023
- Reviewer, Student Paper Competition, 2022
International Conference on Health Policy Statistics (ICHPS)
- Chair, Invited session “Statistical research in rapid response to COVID-19 pandemic: forecasts, risk
factors, therapeutics, and vaccine trials”, Joint Statistical Meetings (JSM) 2021
- Chair, Topic-contributed session “Topics in clustering”, JSM 2018

Departmental Committees

University of South Carolina

- Member, Data Science and Analytics Programs Committee 2024 – Present
- Member, Professional-Track Faculty Hiring Committee 2024 – Present

Southwestern University of Finance and Economics

- Organizer, Departmental Seminar, School of Statistics 2023
- Member, Postdoctoral Evaluation Committee, School of Statistics 2022 – 2024
- Member, Master Thesis Committee, School of Statistics 2022 – 2024
- Member, Bachelor Thesis Committee, School of Statistics 2022 – 2024
- Member, Graduate Student Admissions Committee, School of Statistics 2022 – 2024
- Member, Curriculum Committee, School of Statistics 2021 – 2024

Memberships

- American Statistical Association (ASA)
- International Biometric Society, Eastern North American Region (ENAR)
- International Chinese Statistical Association (ICSA)
- Institute of Mathematical Statistics (IMS)
- New England Statistical Society (NESS)

SOFTWARE

R package ‘APML0’

Co-maintainer

- Augmented and penalized minimization method for regularized linear, logistic, and Cox models with ℓ_0 penalty, flexible for ℓ_1 , ℓ_2 , and network type regularized regression
- Most intensive computation codes written in C++
- Available on CRAN: <https://cran.r-project.org/web/packages/APML0/index.html>
- Downloaded 41,593 times as of 9/2023

R package ‘Covariate-dependent-network’

Maintainer

- Estimate covariate-dependent networks through conditional Gaussian graphical model, in which both the mean and precision matrix depend on covariates
- Most intensive computation codes written in C++
- Available on GitHub: <https://github.com/shanghongxie/Covariate-dependent-network>

R package ‘INL’

Maintainer

- Integrative network learning for multi-modality data
- Most intensive computation codes written in C++
- Available on GitHub: <https://github.com/shanghongxie/INL>

Matlab toolbox ‘OWMKL’

Maintainer

- Outcome weighted multiple kernel learning (OWMKL) for estimating individualized treatment rules
- Available on GitHub: <https://github.com/shanghongxie/OWMKL>

R package ‘FSVM’

Maintainer

- Functional support vector machine for classification and regression problems
- Available on GitHub: <https://github.com/shanghongxie/FSVM>

R package ‘ICATemporalNetwork’

Maintainer

- Temporal causal network learning, adjusting for latent non-Gaussian components and separating the temporal network from the contemporaneous network
- Available on GitHub: <https://github.com/shanghongxie/ICATemporalNetwork>

PRESENTATIONS AND POSTERS

“Identifying Temporal Pathways Using Biomarkers in the Presence of Latent Non-Gaussian Components.”, ICSA Applied Statistics Symposium, Storrs, CT, *Invited session* 5/2025

“Identifying Temporal Pathways Using Biomarkers in the Presence of Latent Non-Gaussian Components.”, Lifetime Data Science Conference (LiDS), New York, NY, *Invited session* 5/2025

“Identifying Temporal Pathways Using Biomarkers in the Presence of Latent Non-Gaussian Components.”, Eastern North American Region (ENAR), New Orleans, LA, *Contributed session* 3/2025

“Network-Assisted Mediation Analysis with High-Dimensional Neuroimaging Mediators.”, ABCD Insights & Innovation Meeting (AIIM) Meeting, Bethesda, MD, Poster session 3/2025

“Identifying Temporal Pathways Using Biomarkers in the Presence of Latent Non-Gaussian Components.”, Statistics Seminar, Department of Mathematical Sciences, Indiana University-Indianapolis, *Invited talk* 2/2025

“Identifying Temporal Pathways Using Biomarkers in the Presence of Latent Non-Gaussian Components.” Functional and Complex Data Analysis (FUNCODA) Working Group, University of South Carolina 12/2024

“Identifying Temporal Pathways Using Biomarkers in the Presence of Latent Non-Gaussian Components.” Conference on Statistical Learning and Data Science (SLDS), California, Contributed session 11/2024

“Identifying Temporal Pathways Using Biomarkers in the Presence of Latent Non-Gaussian Components.” International Conference on Econometrics and Statistics (EcoSta), Beijing, China, *Invited session* 07/2024

“Identifying Disease-associated Biomarker Network Features by Integrating Multi-modality Data.” Department of Statistics, Virginia Tech, *Invited talk* 08/2023

“Identifying Temporal Pathways Using Biomarkers in the Presence of Latent Non-Gaussian Components.” Hangzhou International Conference on Frontiers of Data Science, Hangzhou, China, *Invited session* 08/2023

“Identifying Temporal Pathways Using Biomarkers in the Presence of Latent Non-Gaussian Components.” ICSA China, Chengdu, China, *Invited session* 07/2023

“Identifying Disease-associated Biomarker Network Features through Graphical Models.” Department of Epidemiology and Biostatistics, University of Georgia, *Invited talk* 04/2023

“Evaluating Effectiveness of Public Health Intervention Strategies for Mitigating COVID-19 Pandemic.” New England Statistics Symposium (NESS), Hybrid, *Invited session* 05/2022

“Integrative Network Learning for Multi-modality Biomarker Data.” Center for Statistical Science, Tsinghua University, *Invited talk* 11/2021

“Identifying Temporal Pathways Using High-Dimensional Biomarkers.” Joint Statistical Meetings (JSM), Virtual, Topic-contributed session 08/2021

“Evaluating Effectiveness of Public Health Intervention Strategies for Mitigating COVID-19 Pandemic.” Columbia University Data Science Day, Oral poster session 04/2021

“Integrative Network Learning for Multi-modality Biomarker Data.” Department of Biostatistics and Computational Biology & Del Monte Neuroscience Institute, University of Rochester, *Invited talk* 01/2021

“Integrative Network Learning for Multi-modality Biomarker Data.” Division of Biostatistics, Department of Public Health Sciences, University of Virginia, *Invited talk* 10/2020

“Survival-Convolution Models for Predicting COVID-19 Cases and Assessing Effects of Mitigation Strategies.” Data Science Conference on COVID-19, Presentation session 08/2020

“Conditional Gaussian Graphical Model for Estimating Personalized Disease Symptom Networks.” JSM, Virtual, Topic-contributed session 08/2020

“Integrative Network Learning for Multi-modality Biomarker Data.” Eastern North American Region (ENAR), Virtual, Topic-contributed session 03/2020

“Statistical Methods for Constructing Heterogeneous Biomarker Networks.” Division of Biostatistics, Department of Population Health, New York University School of Medicine, *Invited talk* 11/2019

“Integrative Network Learning for Multi-modality Biomarker Data.” ICSA Applied Statistics Symposium, Raleigh, NC, *Invited session* 06/2019

“Estimating Heterogeneous Biomarker Networks and Their Effects on Disease Outcome.” JSM, Vancouver, Canada, Topic-contributed session 07/2018

“Learning Subject-Specific Directed Acyclic Graphs (DAGs) from High-Dimensional Biomarker Data.” Conference on Statistical Learning and Data Science (SLDS), New York, NY, Poster session 06/2018

“Learning Subject-Specific Directed Acyclic Graphs (DAGs) from High-Dimensional Biomarker Data.” ENAR, Atlanta, GA, Poster session 03/2018

“High-dimensional Subject-Specific Network Analysis for Disentangling Genetic Mutation-Phenotype Pathways.” ICHPS, Charleston, SC, Poster session 01/2018

CAREER DEVELOPMENT

Neuroimaging Short Courses	Harvard University, Martinos Center for Biomedical Imaging
• FreeSurfer Course	04/2017
• Structural and Functional Connectivity via MRI	10/2016