

Fei Xue

Department of Statistics
Purdue University
MATH 506, 150 N. University St
West Lafayette, IN 47907, USA
✉ feixue@purdue.edu
📁 feixue-stat.github.io/

Employment

- 2021–current **Assistant Professor**, *Purdue University, USA*.
2019–2021 **Postdoc Researcher**, *The University of Pennsylvania, USA*.
Supervisor: Hongzhe Li

Education

- 2014–2019 **Ph.D. in Statistics**, *The University of Illinois at Urbana–Champaign, USA*.
Advisor: Annie Qu
2010–2014 **B.S. in Mathematics and Applied Mathematics**, *Fudan University, China*.

Research Interests

Data integration, missing data, deep learning, mediation analysis, mobile health, personalized modeling, precision medicine, high dimensional inference, and statistical genetics.

Papers

1. **Fei Xue**, Annie Qu (2021). Integrating multi-source block-wise missing data in model selection. *Journal of the American Statistical Association*, 116:536, 1914-1927. <https://doi.org/10.1080/01621459.2020.1751176>
2. **Fei Xue**, Yanqing Zhang, Wenzhuo Zhou, Haoda Fu, Annie Qu (2022). Multicategory angle-based learning for estimating optimal dynamic treatment regimes with censored data. *Journal of the American Statistical Association*, 117:539, 1438-1451. <https://doi.org/10.1080/01621459.2020.1862671>
3. **Fei Xue**, Xiwei Tang, Grace Kim, Allison Aiello, Karestan Koenen, Sandro Galea, Derek Wildman, Monica Uddin, Annie Qu (2022). Heterogeneous Mediation Analysis on Epigenomic PTSD and Traumatic Stress in a Predominantly African American Cohort. *Journal of the American Statistical Association*, 117:540, 1669-1683. <https://doi.org/10.1080/01621459.2022.2089572>
4. Xiwei Tang, **Fei Xue**, Annie Qu (2021). Individualized multi-directional variable selection. *Journal of the American Statistical Association*, 116:535, 1280-1296. <https://doi.org/10.1080/01621459.2019.1705308>

5. **Fei Xue**, Annie Qu (2022). Semi-standard partial covariance variable selection when irrerepresentable conditions fail. *Statistica Sinica*, 32: 1881-1909. <https://doi.org/10.5705/ss.202020.0495>
6. Simona Buetti, **Fei Xue**, Qiawen Liu, Juyoen Hur, Gavin Jun Peng Ng, Wendy Heller (2020). Perceived control in the lab and in daily life impact emotion-induced temporal distortions. *Timing & Time Perception*, 9(1): 88-122. <https://doi.org/10.1163/22134468-bja10018>
7. Grace S Kim, Alicia Smith, **Fei Xue**, Vasiliki Michopoulos, Adriana Lori, Don L. Armstrong, Allison E. Aiello, Karestan C. Koenen, Sandro Galea, Derek E. Wildman, Monica Uddin (2019). Methylomic profiles reveal sex-specific shifts in leukocyte composition associated with post-traumatic stress disorder. *Brain, Behavior, and Immunity*, 81: 280-291. <https://doi.org/10.1016/j.bbi.2019.06.025>
8. Agaz Wani, Allison Aiello, Grace S Kim, **Fei Xue**, Andrew Ratanatharathorn, Annie Qu, Karestan Koenen, Sandro Galea, Derek Wildman, Monica Uddin (2021). The impact of psychopathology, social adversity and stress-relevant DNAm on prospective risk for post-traumatic stress: A machine learning approach. *Journal of Affective Disorders*, 282: 894-905. <https://www.biorxiv.org/content/10.1101/2020.09.25.313635v1>
9. **Fei Xue**, Rong Ma, Hongzhe Li. Statistical inference for high-dimensional linear regression with blockwise missing data. *Statistica Sinica*, accepted. https://www3.stat.sinica.edu.tw/ss_newpaper/SS-2022-0104_na.pdf
10. **Fei Xue**, Hongzhe Li. An empirical Bayes regression for multi-tissue eQTL data analysis. *Statistica Sinica*, accepted. https://www3.stat.sinica.edu.tw/ss_newpaper/SS-2023-0201_na.pdf
11. Jiuchen Zhang, **Fei Xue**, Qi Xu, Jung-Ah Lee, Annie Qu. Individualized dynamic latent factor model with application in mobile health data. Under revision. <https://arxiv.org/abs/2311.12392>
12. Yujue Li, **Fei Xue**, Bingxuan Li, Yilin Yang, Zirui Fan, Juan Shu, Xiyao Wang, Jinjie Lin, Carlos Copana, Bingxin Zhao. Analyzing bivariate cross-trait genetic architecture in GWAS summary statistics with the BIGA cloud computing platform. <https://doi.org/10.1101/2023.04.28.538585>
13. **Fei Xue**, Bingxin Zhao. High-dimensional statistical inference for linkage disequilibrium score regression and its cross-ancestry extensions. Submitted. <https://arxiv.org/abs/2306.15779>
14. Huiming Xie, **Fei Xue**, Xiao Wang. Nonparametrically Identifiable Deep Latent Variable Models for MNAR Data. Submitted.
15. Kecheng Wei, **Fei Xue**, Qi Xu, Yubai Yuan, Yuxia Zhang, Guoyou Qin, Agaz Wani, Allison Aiello, Derek Wildman, Monica Uddin, Annie Qu. Time-varying mediation analysis for incomplete data with application to DNA methylation study for PTSD. Submitted. <https://www.biorxiv.org/content/10.1101/2024.02.06.579228v1>

Grants

- 2022-2025 **Principal Investigator**, *Integrative approaches with applications in eQTL analysis and randomized trials*, NSF DMS 2210860.
- 2023-2026 **Co-Principal Investigator**, *Machine Learning for Predictive Breeding Technology*, Beck's Superior Hybrids, INC.

Selected Awards

- 2022 **IMS New Researchers Travel Award**, *Institute of Mathematical Statistics, USA*.
- 2019 **ICSA New Researcher Award**, *International Chinese Statistical Association, USA*.
- 2019 **IMS Hannan Graduate Student Travel Award**, *Institute of Mathematical Statistics, USA*.
- 2018 **ASA Student Paper Award**, *Statistical Learning and Data Science Section of the American Statistical Association, USA*.
- 2018 **Horace W. Norton Prize for Outstanding Thesis Research Finalist**, *Department of Statistics, UIUC, USA*.
- 2018 **The Third Place in Student Poster Competition**, *The Fourth International Workshop on the Statistical Analyses of Multi-Outcome Data, USA*.
- 2013 **Meritorious Winner**, *Mathematical Contest in Modeling, USA*.
- 2010 **Bronze medal**, *Chinese Mathematical Olympiad, China*.

Students

Huiming Xie, Xiaochen Yang, Xiaojing Sun, Hyunwoo Chung

Teaching Experience

Instructor at Purdue

- Spring 2024 Stat 417: **Statistical Theory** (two sections)
- Spring 2023 Stat 528: **Introduction to Mathematical Statistics**
- Fall 2022 Stat 517: **Statistical Inference**
- Spring 2022 Stat 517: **Statistical Inference**
- Fall 2021 Stat 517: **Statistical Inference**

Instructor at UIUC

- Fall 2016 Stat 200: **Statistical Analysis**
Independently led a class of 40 undergraduate students, gave lectures, held office hours and designed exams and homework assignments

Talks

Invited

- June 2023 **ICSA Applied Statistics Symposium, Ann Arbor, MI**.
High-dimensional statistical inference for linkage disequilibrium score regression and its cross-ancestry extensions

- June 2023 **International Purdue Statistics Symposium (IPSS), West Lafayette, IN.**
Individualized dynamic latent factor model with application in mobile health data
- May 2023 **Midwest Machine Learning Symposium (MMLS), Chicago, IL.**
Deep generative modeling for nonignorable missing data with nonparametric identifiability
- May 2023 **BFF8, Cincinnati, OH.**
Statistical inference for high-dimensional block-wise missing data
- Dec. 2022 **CMStatistics (Hybrid), King's College London, UK.**
An Empirical Bayes Regression for Multi-tissue eQTL Data Analysis
- Nov. 2022 **Lecture Series-Trends in Statistics, National University of Singapore, Singapore.**
Statistical inference for high-dimensional block-wise missing data
- Apr. 2022 **CCBB Seminar, Indiana University School of Medicine, Indianapolis, IN.**
Statistical inference for high-dimensional block-wise missing data
- Apr. 2022 **New Challenges and Novel Solutions in Statistics and Data Science, Irvine, CA.**
An Empirical Bayes Regression for Multi-tissue eQTL Data Analysis
- March 2022 **ENAR Spring Meeting, Houston, TX.**
Heterogeneous Mediation Analysis on Epigenomic PTSD and Traumatic Stress in a Predominantly African American Cohort
- March 2022 **Seminar, Rutgers University, New Brunswick, NJ.**
Statistical inference for high-dimensional block-wise missing data
- Feb. 2022 **Biostatistics Seminar, University of Louisville, Louisville, KY.**
Statistical inference for high-dimensional block-wise missing data
- Feb. 2022 **Statistics Seminar, Indiana University-Purdue University Indianapolis, Indianapolis, IN.**
Heterogeneous Mediation Analysis on Epigenomic PTSD and Traumatic Stress in a Predominantly African American Cohort
- Jan. 2022 **Bioinformatics Seminar, Purdue University, West Lafayette, IN.**
Heterogeneous Mediation Analysis on Epigenomic PTSD and Traumatic Stress in a Predominantly African American Cohort
- Nov. 2021 **EAC-ISBA (Hybrid), Atlantic City, New Jersey, USA and Kunming, Yunnan, China.**
An empirical Bayes regression for multi-tissue eQTL data analysis
- Oct. 2021 **NESS (Hybrid), Kingston, RI, USA.**
Statistical inference for high-dimensional linear regression with blockwise missing data
- Sep. 2021 **The ICSA Applied Statistics Symposium, Online.**
Statistical inference for high-dimensional linear regression with blockwise missing data
- Aug. 2021 **JSM, Online.**
Integrating multi-source block-wise missing data in model selection
- March 2021 **ENAR Spring Meeting, Online.**
Heterogeneous mediation analysis for causal inference

- Dec. 2020 **The ICSA Applied Statistics Symposium, Online.**
Multicategory angle-based learning for estimating optimal dynamic treatment regimes with censored data
- Nov. 2020 **Queen's University, Ontario, Canada.**
Multicategory angle-based learning for estimating optimal dynamic treatment regimes with censored data
- Dec. 2019 **Fudan University, Shanghai, China.**
Integrating multi-source block-wise missing data in model selection
- Dec. 2019 **East China Normal University, Shanghai, China.**
Integrating multi-source block-wise missing data in model selection
- Dec. 2019 **The 11th ICSA International Conference, Hangzhou, China.**
Integrating multi-source block-wise missing data in model selection
- Nov. 2019 **Workshop on LEADERS, New Haven, CT, USA.**
Heterogeneous mediator selection for high-dimensional data
- June 2019 **The ICSA Applied Statistics Symposium, Raleigh, NC, USA.**
Integrating multi-source block-wise missing data in model selection
- Oct. 2017 **Yunnan University, Kunming, China.**
Variable selection for highly correlated predictors
- Contributed**
- Aug. 2022 **JSM, Washington, DC, USA.**
An empirical Bayes regression for multi-tissue eQTL analysis
- Aug. 2020 **JSM, Online.**
Multicategory angle-based learning for estimating optimal dynamic treatment regimes with censored data
- Aug. 2018 **JSM, Vancouver, Canada.**
Variable selection for highly correlated predictors
- Nov. 2018 **Bohrer Workshop, Champaign, IL, USA.**
Integrating multi-source block-wise missing data in model selection
- Nov. 2017 **Bohrer Workshop, Champaign, IL, USA.**
Variable selection for highly correlated predictors

Consulting Experience

- 2016–2019 **Statistical Consultant, Illinois Statistics Office, UIUC.**
Provided statistical consultation to 40+ clients
- Clients from Eli Lilly and Company, Pace Suburban Bus, Accel Entertainment, BioFortis Inc, Carle Foundation Hospital, and University of Illinois
 - In the fields of biology, business, ecology, engineering, finance, marketing, medicine, psychology, sociology, etc.
- Covered a wide range of services
- Grant proposal and medical protocol review
 - Experimental design (survey design, clinical trial design)
 - Data analysis (ANOVA, regression, factor analysis, mixed models, etc.),
 - Prediction (market prediction, yield prediction)
 - Software tutorial and technical report writing

Referee Experience

Journal of the American Statistical Association (10)
Journal of the Royal Statistical Society: Series B (4)
Annals of Statistics (1)
Journal of Machine Learning Research (2)
Electronic Journal of Statistics (2)
Journal of Nonparametric Statistics (2)
Journal of Computational and Graphical Statistics (2)
Technometrics (1)
PLOS One (1)
PLOS Genetics (1)
Biometrics (3)
Bioinformatics (1)
The Canadian Journal of Statistics (2)
Statistics in Medicine (2)
Statistica Sinica (3)
Lifetime Data Analysis (1)
Science China Mathematics (1)
Journal of Multivariate Analysis (1)
TEST (1)
Stat (1)
Computational Statistics and Data Analysis (2)
Fertility and Sterility (1)

Service

- 2023–2024 **Member of Talent-Based Hires Committee.**
Purdue University
- 2021–current **Chief Executive Officer of Purdue Statistics Alumni Association (PSAA).**
Purdue University
- 2021–2024 **Member of SLDS Student Paper Award Committee.**
American Statistical Association
- 2022–2023 **Member of Academic Program Committee for International Purdue Statistics Symposium (IPSS).**
Purdue University
- 2022–2023 **Member of Local Organizing Committee for International Purdue Statistics Symposium (IPSS).**
Purdue University
- June 2023 **Session Organizer of Purdue Statistics Alumni.**
IPSS, West Lafayette, IN
- June 2023 **Session Organizer of Recent Developments in Data Integration.**
IPSS, West Lafayette, IN

- March 2023 **Panel Organizer of Networking and career development.**
PSAA March Event, online
- March 2023 **Organizer of Purdue Statistics Alumni Association (PSAA) March Event.**
- December 2022 **Session Organizer of Recent Developments in Statistical Machine Learning.**
CMStatistics, London, UK
- August 2022 **Session Organizer of Recent Developments in Causal Inference.**
JSM, Washington, DC, USA
- March 2022 **Panel Organizer of Collaboration and Networking.**
PSAA March Event, online
- March 2022 **Organizer of Purdue Statistics Alumni Association (PSAA) March Event.**
- 2022 **Member of External Review Committee.**
Purdue University
- 2021–2022 **Member of Data Science Search Committee.**
Purdue University
- 2017–2019 **Committee Member of Graduate Student Invited Seminar.**
The University of Illinois at Urbana–Champaign
- July 2017 **Session Chair of Machine Learning Applications.**
JSM, Baltimore, MD, USA
- June 2017 **Session Chair of Prediction and inference for high-dimensional data and non-parametric models.**
The ICSA Applied Statistics Symposium, Chicago, IL, USA

Professional Memberships

American Statistical Association
Institute of Mathematical Statistics
International Chinese Statistical Association