

南京信息工程大学  
Nanjing University of Information Science & Technology

# 中国现场统计研究会高维数据统计分会 第八届学术研讨会

# 会议手册

主办单位：中国现场统计研究会高维数据统计分会  
承办单位：南京信息工程大学数学与统计学院

江苏·南京  
2024年4月

南京信息工程大学



# 目 录

---

会议委员会.....	1
会议须知.....	2
日程简表.....	3
会议日程.....	4
报告题目和摘要（按报告顺序） .....	13
通讯录（按工作单位首字母顺序） .....	48





# 会议委员会

## 程序委员会

主 席

王启华

委 员

艾明要

常晋源

林华珍

刘卫东

石 坚

唐年胜

王兆军

张立新

张日权

张新雨

赵 鹏

朱利平

邹国华

## 组织委员会

主 席

刘文军

委 员

曹春正

何丽丹

来 鹏

马倩婷

王 辉

卫雨婷

武业文

杨洲木

张 毅

朱连华



## 会议须知

欢迎您参加中国现场统计研究会高维数据统计分会第八届学术研讨会，本着圆满完成此次会议的宗旨，现将有关事宜告知如下：

### ◆ 一、酒店安排

南京华东饭店（江苏省南京市鼓楼区北京西路 67 号）

### ◆ 二、就餐安排

日期	就餐时间	就餐地点
4 月 12 日	晚餐 17: 30-20: 00	南京华东饭店 D 楼
4 月 13 日	早餐 07: 00-08: 00	南京华东饭店 B、D 楼
	午餐 12: 00-13: 00	南京华东饭店 D 楼
	晚宴 18: 00-19: 30	南京华东饭店 D 楼
4 月 14 日	早餐 07: 00-08: 00	南京华东饭店 B、D 楼
	午餐 12: 10-13: 00	南京华东饭店 D 楼

### ◆ 三、注意事项

1、请参会代表提前熟悉会议手册，按照会议议程准时参加各项学术活动，若会议议程时间、地点临时有变化，会务组将及时通知。

2、会议期间，请参会代表佩戴代表证参加各项活动，切记将您的手机调至静音或震动状态。期望您度过一个美好的会期！如遇任何困难，请及时联系会务组。

3、会务组联系人：

卫雨婷：18701109577

何丽丹：15922107730



## 日程简表

日期	时间	事项	地点
4月12日 (星期五)	09:00-20:00	报到注册	D楼一楼大厅
	17:30-20:00	晚餐	D楼一楼餐厅
4月13日 (星期六)	07:00-08:00	早餐	所住酒店
	08:20-09:05	开幕式、合影	D楼二楼大会堂
	09:05-09:55	大会特邀报告 1	
	09:55-10:45	大会特邀报告 2	
	10:45-11:05	茶歇	
	11:05-11:55	大会特邀报告 3	
	12:00-13:00	午餐	D楼一楼餐厅
	13:30-17:30	第一分会场报告	D楼二楼第一会议室
	13:30-17:30	第二分会场报告	D楼二楼第二会议室
	13:30-17:30	第三分会场报告	D楼二楼第三会议室
	13:30-17:30	第四分会场报告	D楼二楼嘉宾二厅
	15:20-15:40	茶歇	
	18:00-19:30	晚宴	D楼一楼餐厅
4月14日 (星期日)	07:00-08:00	早餐	所住酒店
	08:15-12:10	第一分会场报告	D楼二楼第一会议室
	08:15-12:10	第二分会场报告	D楼二楼第二会议室
	08:15-12:10	第三分会场报告	D楼二楼第三会议室
	08:15-12:10	第四分会场报告	D楼二楼嘉宾二厅
	10:05-10:20	茶歇	
	12:10-13:00	午餐	D楼一楼餐厅
	14:00-	自由交流、离会	



# 会议日程

## 4月12日

**会议报到:** 09:00-20:00, 南京华东饭店 D 楼一楼大厅

**晚 餐:** 17:30-20:00, 南京华东饭店 D 楼一楼餐厅

## 4月13日

**早 餐:** 07:00-8:00, 所住酒店

4月13日 08:20-09:05 开幕式及合影 南京华东饭店 D 楼二楼大会堂			
致 辞			主持人
南京信息工程大学校领导 致辞			刘文军
北京大学陈松蹊院士 致辞			
中国现场统计研究会高维数据统计分会理事长王启华教授 致辞			
全体人员合影			
4月13日 09:05-11:55 大会特邀报告 南京华东饭店 D 楼二楼大会堂			
时 间	报告题目	报告人	主持人
09:05-09:55	Ensemble Methods for Testing a Global Null with Applications to Whole Genome Sequencing Studies	林希虹 Harvard University	王启华
09:55-10:45	Weighted Residual Empirical Processes, Martingale Transformations, and Model Specification Tests with Diverging Number of Parameters	朱力行 北京师范大学	刘卫东
茶 歇			
11:05-11:55	Autoregressive Networks with Dependent Edges	常晋源 西南财经大学、中国科学院数学与系统科学研究院	艾明要

**午 餐:** 12:00-13:00, 南京华东饭店 D 楼一楼餐厅



4月13日 13:30-17:30 分会场报告		
时 间	报告题目和报告人	主持人
<b>第一分会场 南京华东饭店 D 楼二楼第一会议室</b>		
13:30-14:05	Max-linear Regression Models with Regularization (邀请报告) 张正军 中国科学院大学	邹国华
14:05-14:30	Two-stage Estimation and Bias-corrected Empirical Likelihood in a Partially Linear Single-index Varying-coefficient Model 薛留根 北京工业大学	
14:30-14:55	Probabilistic Exponential Family Inverse Regression and its Applications 庞道琳 上海交通大学	
14:55-15:20	Unconditional Quantile Regression for Streaming Data Sets 姜 荣 上海第二工业大学	
<b>第二分会场 南京华东饭店 D 楼二楼第二会议室</b>		
13:30-14:05	Manifold Principal Component Analysis and Matrix Elliptical Factor Model (邀请报告) 孔新兵 南京审计大学	张立新
14:05-14:30	High Dimension Portfolio Allocation Using Factor Models 黄 磊 西南交通大学	
14:30-14:55	The Optimality of Wide Neural Network in Large Dimensions 林 乾 清华大学	
14:55-15:20	Generalization and Risk Bounds for Recurrent Neural Networks 程学伟 湖南师范大学	





4月13日 13:30-17:30 分会场报告		
时 间	报告题目和报告人	主持人
<b>第三分会场 南京华东饭店 D 楼二楼第三会议室</b>		
13:30-14:05	Aggregated Representation Learning (邀请报告) 张新雨 中国科学院数学与系统科学研究院	程维虎
14:05-14:30	Strategic Online Learning via Sequentially Integrated Stochastic Gradient Descent Estimators 严晓东 山东大学	
14:30-14:55	Simultaneous Heterogeneous and Reduced-rank Learning for Multivariate Response Regression 吴 捷 安徽大学	
14:55-15:20	Group Penalized Multinomial Logit Models and Stock Return Direction Prediction 胡雪梅 重庆工商大学	
<b>第四分会场 南京华东饭店 D 楼二楼嘉宾二厅</b>		
13:30-14:05	Class-specific Joint Feature Screening in Ultrahigh-dimensional Mixture Regression (邀请报告) 徐 晨 鹏城国家实验室	张崇岐
14:05-14:30	AMDP: An Adaptive Detection Procedure for False Discovery Rate Control in High-dimensional Mediation Analysis 朱学虎 西安交通大学	
14:30-14:55	Test for High-dimensional Linear Hypothesis of Mean Vectors via Random Integration 李江豪 东北师范大学	
14:55-15:20	Testing the Effects of High-dimensional Covariates via Aggregating Cumulative Covariances 周叶青 同济大学	
<b>茶 歇</b>		



4月13日 13:30-17:30 分会场报告		
时 间	报告题目和报告人	主持人
<b>第一分会场 南京华东饭店 D 楼二楼第一会议室</b>		
15:40-16:15	High-dimensional Manifold-based Inference (邀请报告) 郑泽敏 中国科学技术大学	任好洁
16:15-16:40	Testing Conditional Quantile Independence with Functional Covariate 李 杰 中国人民大学	
16:40-17:05	Measures of Concordance and Testing of Independence in Multivariate Structure 邓文丽 江西师范大学	
17:05-17:30	Rank-based Indices for Testing Independence between Two High-dimensional Vectors 许 凯 安徽师范大学	
<b>第二分会场 南京华东饭店 D 楼二楼第二会议室</b>		
15:40-16:15	Covariate-shift Robust Adaptive Transfer Learning for High-dimensional Regression (邀请报告) 刘婧媛 厦门大学	陈 夏
16:15-16:40	Transfer Learning under Generalized Linear Models with Errors-in-variables 张正龙 安徽大学	
16:40-17:05	Integrative Learning of Linear Non-Gaussian Directed Acyclic Graphs 李轩宇 中国科学院大学	
17:05-17:30	Moment-assisted Subsampling Method for Maximum Likelihood Estimator with Large-scale Data 苏苗苗 北京邮电大学	



4月13日 13:30-17:30 分会场报告		
时 间	报告题目和报告人	主持人
<b>第三分会场 南京华东饭店 D 楼二楼第三会议室</b>		
15:40-16:15	Optimal Signal Detection in Covariance and Precision Matrices (邀请报告) 邱宇谋 北京大学	夏志明
16:15-16:40	Change-point Design-based Charting Schemes for Monitoring Process Variability 王 丹 西北大学	
16:40-17:05	Change Point Detection for Multivariate Nonparametric Regression with Deep Neural Networks 周厚林 安徽大学	
17:05-17:30	The Consistency of Estimators in Semiparametric EV Model with Asymptotically Almost Negative Associated Errors and Missing Responses 葛梅梅 安徽大学	
<b>第四分会场 南京华东饭店 D 楼二楼嘉宾二厅</b>		
15:40-16:15	Lasso-adjusted Treatment Effect Estimation under Covariate-adaptive Randomization (邀请报告) 马 维 中国人民大学	王学军
16:15-16:40	Kernel Cox Partially Linear Regression: Building Predictive Models for Cancer Patients' Survival 荣耀华 北京工业大学	
16:40-17:05	Accelerating the Elastic Net Penalized Cox Proportional Hazards Regression Using Safe Screening 王 洪 中南大学	
17:05-17:30	Stochastic Alternating Structure-adapted Proximal Gradient Descent Method with Variance Reduction for Nonconvex 贾泽慧 南京信息工程大学	



4月14日 08:15-12:10 分会场报告		
时 间	报告题目和报告人	主持人
<b>第一分会场 南京华东饭店 D 楼二楼第一会议室</b>		
08:15-08:50	Feature Screening for Cluster Analysis Using Quasi-likelihood (邀请报告) 席瑞斌 北京大学	李高荣
08:50-09:15	Group Feature Screening via Maximum Information Coefficient (MIC) for Ultra-high Dimensional Multi-Classification 陈婷婷 桂林理工大学	
09:15-09:40	Partitioning and Aggregating Cross-tissue and Tissue-specific Genetic Effects in Identifying Gene-trait Associations 宋 爽 清华大学	
09:40-10:05	Knockoff-based Statistics for the Identification of Putative Causal Genes in Genetic Studies 马诗洋 上海交通大学	
<b>第二分会场 南京华东饭店 D 楼二楼第二会议室</b>		
08:15-08:50	Graphical Models for Temporal Point Processes (邀请报告) 林 伟 北京大学	赵俊龙
08:50-09:15	Mathematical Imaging: From Variational Modelling to Deep Learning for Images 马倩婷 南京信息工程大学	
09:15-09:40	Grouped Heterogeneous Gaussian Graphical Models for High-dimensional Clustered Data 曾 鑫 厦门大学	
09:40-10:05	Asymptotic Properties of a Multi-colour Random Reinforced Urn Model with Multiple-drawing and Random Addition 杨 丽 东北师范大学	



4月14日 08:15-12:10 分会场报告		
时 间	报告题目和报告人	主持人
<b>第三分会场 南京华东饭店 D 楼二楼第三会议室</b>		
08:15-08:50	Navigating Challenges in Classification and Outlier Detection: A Remedy Based on Semi-parametric Density Ratio Models (邀请报告) 刘玉坤 华东师范大学	何道江
08:50-09:15	Bidirectional Efficient Non-convex Adaptive Federated Learning 杨 光 南京审计大学	
09:15-09:40	The Optimality of Kernel Classifiers in Sobolev Space 赖建发 清华大学	
09:40-10:05	Solving Saddle Point Problems: A Landscape of Primal-dual Algorithm with Larger Stepsizes 姜 帆 南京信息工程大学	
<b>第四分会场 南京华东饭店 D 楼二楼嘉宾二厅</b>		
08:15-08:50	A Random Graph-based Autoregressive Model for Networked Time Series (邀请报告) 吴未迟 清华大学	陈占寿
08:50-09:15	Statistical Inference of High-dimensional Vector Autoregressive Time Series with Non-i.i.d. Innovations 张云翼 香港中文大学（深圳）	
09:15-09:40	Privacy-preserving Parametric Inference for Spatial Autoregressive Model 宋允全 中国石油大学（华东）	
09:40-10:05	空间分位面板自回归模型的贝叶斯 Elastic Net 变量选择及其应用 于卓熙 辽宁大学	
<b>茶 歇</b>		



4月14日 08:15-12:10 分会场报告		
时 间	报告题目和报告人	主持人
<b>第一分会场</b> 南京华东饭店 D 楼二楼第一会议室		
10:20-10:55	Model-free Variable Selection in High Dimension via Constrained Kernel Regression (邀请报告) 李长城 大连理工大学	曹春正
10:55-11:20	On Selection of Semiparametric Spatial Regression Models with Partial Differential Equation Regularization 刘 静 中国矿业大学	
11:20-11:45	A Two-stage Variable Selection Based on the Adjusted Adaptive Lasso in High-dimensional Cancer Classification 蒋青青 桂林理工大学	
11:45-12:10	fastCCLasso: Fast and Efficient Algorithm for Estimating Correlation Matrix from Compositional Data 张 琄 首都师范大学	
<b>第二分会场</b> 南京华东饭店 D 楼二楼第二会议室		
10:20-10:55	Causal Inference in Randomized Experiments for Dyadic Data with Interference (邀请报告) 苗 旺 北京大学	来 鹏
10:55-11:20	Multivariate Spatiotemporal Models with Low Rank Coefficient Matrix 蒲 丹 西南财经大学	
11:20-11:45	Correcting Spot Power Variation Estimator via Edgeworth Expansion 何丽丹 南京信息工程大学	
11:45-12:10	Generalization Error Curves for Analytic Spectral Algorithms under Power-law Decay 李易诚 清华大学	



4月14日 08:15-12:10 分会场报告		
时 间	报告题目和报告人	主持人
<b>第三分会场 南京华东饭店 D 楼二楼第三会议室</b>		
10:20-10:55	Construction of High-dimensional Designs for Computer Experiments (邀请报告) 何 煦 中国科学院数学与系统科学研究院	骆 威
10:55-11:20	Gaussian Variational Approximation with Composite Likelihood for Crossed Random Effect Models 徐礼柏 苏州大学	
11:20-11:45	Influence Exploration of Estimating Procedure on Adequacy Test of Semiparametric Censored Regression Models 孙志华 中国科学院大学	
11:45-12:10	Simultaneous Variable Selection and Estimation for Survival Data via the Gaussian Seamless- $L_0$ Penalty 刘自力 中南大学	
<b>第四分会场 南京华东饭店 D 楼二楼嘉宾二厅</b>		
10:20-10:55	度量空间回归的随机森林方法及理论 (邀请报告) 於 州 华东师范大学	张庆昭
10:55-11:20	儿科临床大数据分析面临的机遇与挑战 徐铄明 重庆医科大学附属儿童医院	
11:20-11:45	大模型的成功与高维数据算法存在的问题 夏慧异 池州学院	
11:45-12:10	Quantile Generalized Measures of Correlation 张新瑜 北京师范大学	

## 报告题目和摘要（按报告顺序）

4月13日 09:05-11:50

大会特邀报告 1

**报告题目：** Ensemble Methods for Testing a Global Null with Applications to Whole Genome Sequencing Studies

报告人：林希虹 Harvard University

**报告摘要：** Testing a global null is a canonical problem in statistics and has a wide range of applications. In view of the fact of no uniformly most powerful test, prior and/or domain knowledge are commonly used to focus on a certain class of alternatives to improve the testing power, e.g., the class of alternatives in the scenario of the same effect sign or signal sparsity. However, it is generally challenging to develop tests that are particularly powerful against a certain class of alternatives. In this paper, motivated by the success of ensemble learning methods for prediction or classification, we propose an ensemble framework for testing that mimics the spirit of random forests to deal with the challenges. Our ensemble testing framework aggregates a collection of weak base tests to form a final ensemble test that maintains strong and robust power. The key component of the framework is to introduce a certain random procedure in the construction of base tests. We then apply the framework to four problems about global testing in different classes of alternatives arising from Whole Genome Sequencing (WGS) association studies. Specific ensemble tests are proposed for each of these problems, and their theoretical optimality is established in terms of Bahadur efficiency. Extensive simulations are conducted to demonstrate type I error control and power gain of the proposed ensemble tests. In an analysis of the WGS data from the Atherosclerosis Risk in Communities (ARIC) study, the ensemble tests demonstrate substantial and consistent power improvement compared to other existing tests.

大会特邀报告 2

**报告题目：** Weighted Residual Empirical Processes, Martingale Transformations, and Model Specification Tests with Diverging Number of Parameters

报告人：朱力行 北京师范大学

**报告摘要：** This paper proposes a new methodology for testing the parametric forms of the mean and variance functions based on weighted residual empirical processes and their martingale transformations in regression models. The dimensions of the parameter vectors can be divergent as the sample size goes to infinity. We study the convergence of weighted residual empirical processes and their martingale transformation under the null and alternative hypotheses in diverging dimension settings. The proposed tests based on weighted residual empirical processes can detect local alternatives distinct from the null at the fastest possible rate of order  $n^{-1/2}$  but are not asymptotically distribution-free. While tests based on martingale transformed weighted residual





empirical processes can be asymptotically distribution-free, yet, unexpectedly, can only detect the local alternatives converging to the null at a much slower rate of order  $n^{-1/4}$ , which is somewhat different from existing asymptotically distribution-free tests based on martingale transformations. As the tests based on the residual empirical process are not distribution-free, we propose a smooth residual bootstrap and verify the validity of its approximation in diverging dimension settings. Simulation studies and a real data example are conducted to illustrate the effectiveness of our tests.

---

大会特邀报告 3

**报告题目:** Autoregressive Networks with Dependent Edges

**报告人:** 常晋源 西南财经大学、中国科学院数学与系统科学研究院

**报告摘要:** We propose an autoregressive framework for modelling dynamic networks with dependent edges. It encompasses the models which accommodate, for example, transitivity, density-dependent and other stylized features often observed in real network data. By assuming the edges of network at each time are independent conditionally on their lagged values, the models, which exhibit a close connection with temporal ERGMs, facilitate both simulation and the maximum likelihood estimation (MLE) in the straightforward manner. Due to the possible large number of parameters in the models, the initial MLEs may suffer from slow convergence rates. An improved estimator for each component parameter is proposed based on an iteration based on the projection which mitigates the impact of the other parameters (Chang et al., 2021; Chang, Shi and Zhang, 2023). Based on a martingale difference structure, the asymptotic distribution of the improved estimator is derived without the stationarity assumption. The limiting distribution of the estimator is not normal in general, and it reduces to normal when the underlying process satisfies some mixing conditions. Illustration with a transitivity model was carried out in both simulation and with two real network data sets.



4月13日 13:30-15:20

第一分会场

分会场报告1 (邀请报告)

**报告题目:** Max-linear Regression Models with Regularization

报告人: 张正军 中国科学院大学

**报告摘要:** In this talk, I will introduce a new class of max-linear regression models to take advantage of easy interpretable features embedded in linear regression models. It can be seen that linear relation is a special case of max-linear relation. We develop an EM algorithm based maximum likelihood estimation procedure. The consistency and asymptotics of the estimators for parameters are proved. To advance max-linear models to deal with high dimensional predictors, we adopt the common strategy of regularization in the high dimensional regression literature. We demonstrate the broad applicability of max-linear models using simulation examples and real applications in econometric and business modeling. The results, in terms of predictability, show a significant improvement compared with solely using regular regression models and other existing machine learning methods. The results enhance our understanding of the relationship between the response variable and the predictors, and among the predictors as well. Joint work (JoE) with Qirong Cui, Yuqing Xu, Vincent Chan.

-----

分会场报告2

**报告题目:** Two-stage Estimation and Bias-corrected Empirical Likelihood in a Partially Linear Single-index Varying-coefficient Model

报告人: 薛留根 北京工业大学

**报告摘要:** In this talk, we study the estimation and empirical likelihood (EL) of the parameters of interest in a partially linear single-index varying-coefficient model. A two-stage method is presented to estimate the regression parameters and the coefficient functions. The asymptotic distributions of the proposed estimators are obtained. Meanwhile, a bias-corrected EL ratio for the regression parameters is proposed. It is shown that the ratio is asymptotically standard chi-squared. The result can be directly used to construct the EL confidence regions of the regression parameters. Simulation studies are carried out to evaluate the finite sample behavior of the proposed method. An application example of a real data set is given.

-----

分会场报告3

**报告题目:** Probabilistic Exponential Family Inverse Regression and Its Applications

报告人: 庞道琳 上海交通大学

**报告摘要:** Rapid advances in high-throughput sequencing technologies have led to the fast accumulation of high-dimensional data, which is harnessed for understanding the implications of various factors on human disease and health. While dimension reduction plays an essential role in high-dimensional regression and classification, existing methods often require the predictors to be continuous, making them unsuitable



for discrete data, such as presence-absence records of species in community ecology and sequencing reads in single-cell studies. To identify and estimate sufficient reductions in regressions with discrete predictors, we introduce probabilistic exponential family inverse regression (PrEFIR), assuming that, given the response and a set of latent factors, the predictors follow one-parameter exponential families. We show that the low-dimensional reductions result not only from the response variable but also from the latent factors. We further extend the latent factor modeling framework to the double exponential family by including an additional parameter to account for the dispersion. This versatile framework encompasses regressions with all categorical or a mixture of categorical and continuous predictors. We propose the method of maximum hierarchical likelihood for estimation, and develop a highly parallelizable algorithm for its computation. The effectiveness of PrEFIR is demonstrated through simulation studies and real data examples.

---

分会场报告 4

**报告题目:** Unconditional Quantile Regression for Streaming Data Sets

**报告人:** 姜荣 上海第二工业大学

**报告摘要:** In this talk, we introduce a renewable parameter estimation method for unconditional quantile regression (UQR) applied to streaming data sets. To address this, we present a novel approach involving smoothing logistic regression estimation. Subsequently, we propose a renewable estimator tailored for UQR with streaming data, which relies on current data and summary statistics derived from historical data. Theoretically, our proposed estimators exhibit equivalent asymptotic properties to the standard version computed directly on the entire dataset, without any additional constraints. Both simulations and real data analysis are conducted to illustrate the finite sample performance of the proposed methods.



## 第二分会场

分会场报告 1 (邀请报告)

**报告题目:** Manifold Principal Component Analysis and Matrix Elliptical Factor Model

**报告人:** 孔新兵 南京审计大学

**报告摘要:** In this talk, we, for the first time, propose the matrix elliptical factor model, by taking the heavy tails of observations (e.g., financial returns) into consideration. Manifold principal component analysis (MPCA) is, for the first time, introduced to estimate the row/column loading spaces. MPCA first performs singular value decomposition (SVD) for each “local” matrix observation, and then finds the “center” of the locally estimated spaces across all observations, while all existing PCA methods first integrate data across observations and then do eigenvalue decomposition of the sample covariance matrices. We propose two versions of MPCA algorithms to estimate the factor loading matrices robustly, without any moment constraints on the factors and the idiosyncratic errors. Theoretical convergence rates of the corresponding estimators of the factor loading matrices, factor score matrices and common components are derived under mild conditions. Asymptotic distribution of the standard error for the empirical subspace “center” deviating from the population one is provided, which is of independent interest. In the end, we provide robust and consistent estimators of the row/column factor numbers based on the eigenvalue-ratio approach. Numerical studies and real example on financial returns check the flexibility of our model and the validity of the MPCA methods.

---

分会场报告 2

**报告题目:** High Dimension Portfolio Allocation Using Factor Models

**报告人:** 黄磊 西南交通大学

**报告摘要:** In finance, mean–variance portfolio (MVP) is often applied to solve portfolio allocation problems. However, when investing in high dimension stocks, the classical MVP is no longer applicable. For large portfolio allocation, this study constructs a comprehensive MVP model, called an unconstrained regression model with latent factors. It applies a factor structure with latent factors to modify the covariance matrix estimation in the classical MVP, using an unconstrained penalized regression to estimate the allocation. The empirical results show that this model can control the risk with Sharpe ratio reaching 6 times as many as MVP model, and even improve about 26% compared with the sub-optimal portfolio model.

---

分会场报告 3

**报告题目:** The Optimality of Wide Neural Network in Large Dimensions

**报告人:** 林乾 清华大学

**报告摘要:** We perform a study on the generalization ability of a two-layer wide neural network for large dimensional data (where the sample size  $n$  is polynomially depending on the dimension  $d$  of the samples, i.e.,  $n \asymp d^\gamma$  for some  $\gamma > 0$ ). We first build a general tool to characterize the lower bound and upper bound of the kernel



regression for large dimensional data through the metric entropy  $\bar{\varepsilon}_n^2$  and the Mendelson complexity  $\varepsilon_n^2$  respectively. When the target function falls into the RKHS associated with the neural tangent kernel (NTK) defined on  $\mathbb{S}^d$ , we utilize the new tool to show that the minimax rate of the excess risk of kernel regression with NTK is  $n^{-1/2}$  when  $n \asymp d^\gamma$  where  $\gamma = 2, 4, 8, 12, \dots$ . We then further determine the optimal rate of the excess risk of kernel regression with NTK for all the  $\gamma > 0$  and find that the curve of optimal rate varying along  $\gamma$  exhibits several new phenomena including the multiple descent behavior and the periodic plateau behavior. As a direct corollary, we know the above claims hold for wide neural network as well.

---

分会报告 4

**报告题目:** Generalization and Risk Bounds for Recurrent Neural Networks

报告人: 程学伟 湖南师范大学

**报告摘要:** Recurrent Neural Networks (RNNs) have achieved great success in the prediction of sequential data. However, their theoretical studies are still lagging behind because of their complex interconnected structures. In this paper, we establish a new generalization error bound for vanilla RNNs, and provide a unified framework to calculate the Rademacher complexity that can be applied to a variety of loss functions. When the ramp loss is used, we show that our bound is tighter than the existing bounds based on the same assumption on the Frobenius and spectral norms of the weight matrices and a few mild conditions. Moreover, we derive a sharp estimation error bound for RNN-based estimators obtained through empirical risk minimization (ERM) in multi-class classification problems when the loss function satisfies a Bernstein condition.



### 第三分会场

分会场报告 1 (邀请报告)

**报告题目:** Aggregated Representation Learning

报告人: 张新雨 中国科学院数学与系统科学研究院

**报告摘要:** This paper studies the general framework of representation learning and develops a frequentist model averaging framework to combine different representation learning methods. The weight choice criterion is K-fold cross-validation criterion. We prove that the model averaging estimator is asymptotically optimal under certain assumptions.

-----

分会场报告 2

**报告题目:** Strategic Online Learning via Sequentially Integrated Stochastic Gradient Descent Estimators

报告人: 严晓东 山东大学

**报告摘要:** The stochastic gradient descent algorithm, often depicted as SGD, has been widely employed in various fields of artificial intelligence and is a prototype of online learning algorithms. In the article, we propose a novel and general framework of one-sided testing for streaming data based on SGD. The proposed method constructs an online-updated test statistic sequentially by integrating the selected batch-specific estimators or its opposite, which is referred to as opposite online learning. Notably, the batch-specific online estimators are chosen strategically according to the proposed sequential tactics designed by the two-armed bandit process. Theoretical results prove the strategy's advantage, ensuring that the test statistic distribution is optimal under the null hypothesis. We also supply the theoretical evidence of power enhancement compared with classical test statistics. In application, the proposed method is appealing for statistical inference of two-sided testing and it is scalable and adaptable for any model. Finally, the superior finite-sample performance is evaluated by simulation studies.

-----

分会场报告 3

**报告题目:** Simultaneous Heterogeneous and Reduced-rank Learning for Multivariate Response Regression

报告人: 吴捷 安徽大学

**报告摘要:** Heterogeneous data are now ubiquitous in many applications in which correctly identifying the subgroups from a heterogeneous population plays an important role. Although there is an increasing body of literature on subgroup detection, existing methods mainly focus on the univariate response setting. In this paper, we propose a heterogeneous and reduced-rank learning framework to simultaneously identify the subgroup structure and estimate the covariate effects for heterogeneous multivariate response regression. In particular, our approach uses rank-constrained pairwise fusion penalization and conducts the subgroup analysis without requiring prior knowledge regarding the individual subgroup memberships. We implement the proposed approach



by an alternating direction method of multipliers (ADMM) algorithm and show its convergence. We also establish the asymptotic properties for the resulting estimators under mild and interpretable conditions. A predictive information criterion is proposed to select the rank of the coefficient matrix with theoretical support. The effectiveness of the proposed approach is demonstrated through simulation studies and a real data application.

---

分会场报告 4

**报告题目:** Group Penalized Multinomial Logit Models and Stock Return Direction Prediction

报告人: 胡雪梅 重庆工商大学

**报告摘要:** Multinomial logit model (MLM) has been proposed as the most frequently regression model for multi-category response and the widely used functional form for discrete choice probabilities. To deal with correlated data, in this paper we propose G-LASSO/G-SCAD/G-MCP penalized MLM model to exert class discovery and class prediction for multi-category classification problems. Firstly, we develop a group coordinate descent algorithm to simultaneously complete group selection and group estimation, and prove its convergence under mild conditions. Secondly, we apply the training set and group estimations to obtain class probability estimators, choose the Bayes classifier to identify class index information, and introduce the testing set and a few measures to assess multi-category prediction performance. Simulations show that the proposed methods outperform LASSO/SCAD/MCP penalized MLM, 3 deep learning methods and 3 machine learning methods in terms of Kappa, PDI, Optimal or Average Accuracy. Finally, we combine group penalized MLM with 58 technical indicators to predict up trends, sideways trends and down trends for stock returns, and show that the proposed methods outperform the other 9 methods in terms of Accuracy, PDI, Kappa and HUM. Therefore, the proposed method can not only accommodate the correlation information, but also improve multi-category prediction performance by shrinking group coefficients.

## 第四分会场

分会场报告 1 (邀请报告)

**报告题目:** Class-specific Joint Feature Screening in Ultrahigh-dimensional Mixture Regression

**报告人:** 徐晨 鹏城国家实验室

**报告摘要:** Finite mixture of regression models are ubiquitous for analyzing complex data. They aim to detect heterogeneity in the effects of a set of features on a response over a finite number of latent classes. When the number of features is large, a direct fitting of mixture regressions can be computationally infeasible and often leads to a poor interpretative value. One practical strategy is to screen out most irrelevant features before an in-depth analysis. In this paper, we propose a novel method for feature screening in ultrahigh-dimensional Gaussian finite mixture of regressions. The new method is built upon a sparsity-restricted expectation-approximation-maximization algorithm, which simultaneously removes varying sets of irrelevant features from multiple latent classes. In the screening process, joint effects between features are naturally accounted and class-specific screening results are produced without ad hoc steps. These merits give the new method an edge to outperform the existing screening methods. The promising performance of the method is supported by both theory and numerical examples including a real data analysis.

---

分会场报告 2

**报告题目:** AMDP: An Adaptive Detection Procedure for False Discovery Rate Control in High-dimensional Mediation Analysis

**报告人:** 朱学虎 西安交通大学

**报告摘要:** High-dimensional mediation analysis is often associated with a multiple testing problem for detecting significant mediators. Assessing the uncertainty of this detecting process via false discovery rate (FDR) has garnered great interest. To control the FDR in multiple testing, two essential steps are involved: ranking and selection. Existing approaches either construct p-values without calibration or disregard the joint information across tests, leading to conservatism in FDR control or non-optimal ranking rules for multiple hypotheses. In this paper, we develop an Adaptive Mediation Detection Procedure (referred to as "AMDP") to identify relevant mediators while asymptotically controlling the FDR in high-dimensional mediation analysis. AMDP produces the optimal rule for ranking hypotheses and proposes a data-driven strategy to determine the threshold for mediator selection. This novel method captures information from the proportions of composite null hypotheses and the distribution of p-values, which turns the high dimensionality into an advantage instead of a limitation. The numerical studies on synthetic and real data sets illustrate the performances of AMDP compared with existing approaches.





## 分会场报告 3

**报告题目:** Test for High-dimensional Linear Hypothesis of Mean Vectors via Random Integration

报告人: 李江豪 东北师范大学

**报告摘要:** In this paper, we investigate hypothesis testing for the linear combination of mean vectors across multiple populations through the method of random integration. We have established the asymptotic distributions of the test statistics under both null and alternative hypotheses. Additionally, we provide a theoretical explanation for the special use of our test statistics in situations when the nonzero signal in the linear combination of the true mean vectors is weakly dense. Moreover, Monte-Carlo simulations are presented to evaluate the suggested test against existing high-dimensional tests. The findings from these simulations reveal that our test not only aligns with the performance of other tests in terms of size but also exhibits superior power.

-----

## 分会场报告 4

**报告题目:** Testing the Effects of High-dimensional Covariates via Aggregating Cumulative Covariances

报告人: 周叶青 同济大学

**报告摘要:** In this article, we test for the effects of high-dimensional covariates on the response. In many applications, different components of covariates usually exhibit various levels of variation, which is ubiquitous in high dimensional data. To simultaneously accommodate such heteroscedasticity and high dimensionality, we propose a novel test based on an aggregation of the marginal cumulative covariances, requiring no prior information on the specific form of regression models. Our proposed test statistic is scale-invariance, tuning free and convenient to implement. The asymptotic normality of the proposed statistic is established under the null hypothesis. We further study the asymptotic relative efficiency of our proposed test with respect to the state-of-art universal tests in two different settings: one is designed for high-dimensional linear model and the other is introduced in a completely model-free setting. A remarkable finding reveals that, thanks to the scale-invariance property, even under the high-dimensional linear models, our proposed test is asymptotically much more powerful than existing competitors for the covariates with heterogeneous variances while maintaining high efficiency for the homoscedastic ones. Supplementary materials for this article are available online.



4月13日 15:40-17:30

### 第一分会场

分会场报告 5 (邀请报告)

**报告题目:** High-dimensional Manifold-based Inference

报告人: 郑泽敏 中国科学技术大学

**报告摘要:** Multi-task learning is a widely used technique for harnessing information from various tasks. Recently, the sparse orthogonal factor regression (SOFAR) framework, based on the sparse singular value decomposition (SVD) within the coefficient matrix, was introduced for interpretable multi-task learning, enabling the discovery of meaningful latent feature-response association networks across different layers. However, conducting precise inference on the latent factor matrices has remained challenging due to orthogonality constraints inherited from the sparse SVD constraint. In this paper, we suggest a novel approach called high-dimensional manifold-based SOFAR inference (SOFARI), drawing on the Neyman near-orthogonality inference while incorporating the Stiefel manifold structure imposed by the SVD constraints. By leveraging the underlying Stiefel manifold structure, SOFARI provides bias-corrected estimators for both latent left factor vectors and singular values, for which we show to enjoy the asymptotic mean-zero normal distributions with estimable variances. We introduce two SOFARI variants to handle strongly and weakly orthogonal latent factors, where the latter covers a broader range of applications. We illustrate the effectiveness of SOFARI and justify our theoretical results through simulation examples and a real data application in economic forecasting.

---

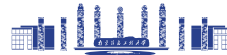
分会场报告 6

**报告题目:** Testing Conditional Quantile Independence with Functional Covariate

报告人: 李杰 中国人民大学

**报告摘要:** We propose a new nonparametric conditional independence test for a scalar response and a functional covariate over a continuum of quantile levels. We build a Cramér-von Mises type test statistic based on an empirical process indexed by random projections of the functional covariate, effectively avoiding the “curse of dimensionality” under the projected hypothesis, which is almost surely equivalent to the null hypothesis. The asymptotic null distribution of the proposed test statistic is obtained under some mild assumptions. The asymptotic global and local power properties of our test statistic are then investigated. We specifically demonstrate that the statistic is able to detect a broad class of local alternatives converging to the null at the parametric rate. Additionally, we recommend a simple multiplier bootstrap approach for estimating the critical values. The finite-sample performance of our statistic is examined through several Monte Carlo simulation experiments. Finally, an analysis of an EEG data set is used to show the utility and versatility of our proposed test statistic.

分会场报告 7



**报告题目:** Measures of Concordance and Testing of Independence in Multivariate Structure

**报告人:** 邓文丽 江西师范大学

**报告摘要:** Two random variables are concordant if one variable is large and then the other one tends to be large. Spearman's rank correlation and Kendall's tau can be used to measure the trend of both variables rising and falling simultaneously. For a multivariate case, most studies are based on average Spearman's rank correlation or average Kendall's tau, which compute bivariate measures of concordance for all pairs of variables and then average the results. A new measure of concordance which considers all the random variables simultaneously is proposed in this paper. The distribution and other relevant properties of this statistic are deduced. Since it is a U-statistic, this statistic follows an asymptotically normal distribution. Furthermore, a nonparametric test method for the independence of multivariate random variables is proposed.

-----  
分会场报告 8

**报告题目:** Rank-based Indices for Testing Independence between Two High-dimensional Vectors

**报告人:** 许凯 安徽师范大学

**报告摘要:** To test independence between two high-dimensional random vectors, we propose three tests based on the rank-based indices derived from Hoeffding's  $D$ , Blum-Kiefer-Rosenblatt's  $R$  and Bergsma-Dassios-Yanagimoto's  $\tau^*$ . Under the null hypothesis of independence, we show that the distributions of the proposed test statistics converge to normal ones if the dimensions diverge arbitrarily with the sample size. We further derive an explicit rate of convergence. Thanks to the monotone transformation-invariant property, these distribution-free tests can be readily used to generally distributed random vectors including heavily tailed ones. We further study the local power of the proposed tests and compare their relative efficiencies with two classic distance covariance/correlation based tests in high dimensional settings. We establish explicit relationships between  $D$ ,  $R$ ,  $\tau^*$  and Pearson's correlation for bivariate normal random variables. The relationships serve as a basis for power comparison. Our theoretical results show that under a Gaussian equicorrelation alternative, (i) the proposed tests are superior to the two classic distance covariance/correlation based tests if the components of random vectors have very different scales; (ii) the asymptotic efficiency of the proposed tests based on  $D$ ,  $\tau^*$  and  $R$  are sorted in a descending order.



## 第二分会场

分会场报告 5 (邀请报告)

**报告题目:** Covariate-shift Robust Adaptive Transfer Learning for High-dimensional Regression

报告人: 刘婧媛 厦门大学

**报告摘要:** The main challenge that sets transfer learning apart from traditional supervised learning is the distribution shift, reflected as the shift between the source and target models and that between the marginal covariate distributions. High-dimensional data introduces unique challenges, such as covariate shifts in the covariate correlation structure and model shifts across individual features in the model. In this work, we tackle model shifts in the presence of covariate shifts in the high-dimensional regression setting. Furthermore, to learn transferable information which may vary across features, we propose an adaptive transfer learning method that can detect and aggregate the feature-wise transferable structures. Non-asymptotic bound is provided for the estimation error of the target model, showing the robustness of the proposed method to high-dimensional covariate shifts.

---

分会场报告 6

**报告题目:** Transfer Learning under Generalized Linear Models with Errors-in-variables

报告人: 张正龙 安徽大学

**报告摘要:** This paper investigates transfer learning within the framework of generalized linear models with errors-in-variables. In practical scenarios, the presence of measurement errors in covariates, originating from both machine measurements and human observations, can introduce inconsistency into maximum likelihood estimation methods. We propose a transfer learning approach for handling covariate measurement errors by establishing an errors-in-variables transfer learning likelihood using linear correlation-ratio. Theoretical analysis demonstrates the robust performance of our method even when dealing with covariate measurement errors. Extensive simulation experiments validate the effectiveness of the proposed approach. Finally, we apply our method to analyze real data from the Maryland River Biological Resources Survey, showcasing its superior analytical performance compared to methods that only use the target data.

---

分会场报告 7

**报告题目:** Integrative Learning of Linear Non-Gaussian Directed Acyclic Graphs

报告人: 李轩宇 中国科学院大学

**报告摘要:** The Directed Acyclic Graph model (DAG) is a fundamental model to represent the causal relationship between random variables. We consider the problem of learning multiple independent linear non-Gaussian directed acyclic graphs (LiNGAMs) in high dimensional cases, where the graphs share the same sparse topological structure. We propose a novel integrative learning method consisting of the



Group Lasso to do integrative neighborhood regression and permutation independence test to identify the topological layers of the directed acyclic graphs in a bottom-up fashion. We also establish the consistency result of our approach under mild conditions, and theoretically demonstrate the quantitative advantage of our integrative learning method over separate learning methods. The effectiveness of our proposed method is further supported by the numerical comparison against popular competitors in various simulated examples as well as a real dataset application.

---

分会场报告 8

**报告题目:** Moment-assisted Subsampling Method for Maximum Likelihood Estimator with Large-scale Data

报告人: 苏苗苗 北京邮电大学

**报告摘要:** The maximum likelihood estimator for a parametric conditional density model can be computationally cumbersome for large-scale datasets, especially when the likelihood function requires integral computations. Subsampling offers an efficient solution for this problem. This paper proposes a moment-assisted subsampling method which can improve the estimation efficiency of any estimating equation-based subsampling estimators. Some sample moments can be efficiently computed even if the sample size of the whole dataset is huge. The proposed method incorporates informative sample moments of the whole data through the generalized method of moments. The resulting subsampling estimator can be computed rapidly and is asymptotically normal with a smaller asymptotic variance than the corresponding estimator without incorporating sample moments of the whole data. The proposed method is further extended to accommodate model misspecification. The asymptotic variance of the moment-assisted estimator relies on the moment function of the sample moment. We derive the optimal moment function that minimizes the resulting asymptotic variance in terms of Loewner order. The optimal moment function depends on the data generating process and might be challenging to calculate. We provide its data-adaptive approximation that is easy to compute. When the approximation is sufficiently accurate, the proposed estimator can achieve the same efficiency as the whole data-based maximum likelihood estimator. Numerical results demonstrate the promising performance of the moment-assisted method across various scenarios.



### 第三分会场

分会场报告 5 (邀请报告)

**报告题目:** Optimal Signal Detection in Covariance and Precision Matrices

**报告人:** 邱宇谋 北京大学

**报告摘要:** This talk considers testing for high dimensional covariance and precision matrices by deriving the detection boundaries as a function of the signal sparsity and signal strength. It first shows that the optimal detection boundary for testing sparse means is the minimax detection lower boundary for testing covariance and precision matrices. Multi-level thresholding tests are proposed and are shown to be able to attain the detection lower boundaries over a substantial range of the sparsity parameter, implying that the multi-level thresholding tests are sharp optimal in the minimax sense over the range. The asymptotic distributions of the multi-level thresholding statistic for covariance and precision matrices are derived by developing a novel U-statistic decomposition to handle the complex dependence among the elements of the estimated covariance and precision matrices. The superiority in the detection boundary of the multi-level thresholding test over the existing tests are also demonstrated.

分会场报告 6

**报告题目:** Change-point Design-based Charting Schemes for Monitoring Process Variability

**报告人:** 王丹 西北大学

**报告摘要:** Most of the recommended nonparametric control charts apply to monitor location parameter. In this paper, we propose a nonparametric exponentially weighted moving average (EWMA) control chart for monitoring scale parameter, which combines the Ansari-Bradley test and the framework of change point detection. As with most nonparametric control charts, a large number of historical observations are used to motivate the control chart, and we propose a nonparametric EWMA control chart that achieves the same purpose with a very small number of historical observations. Simulation results demonstrate that our control chart is effective and robust for monitoring scale parameter shifts. For illustration of our main results, an applied example demonstrates the practicality of this control chart for monitoring scale parameter.

分会场报告 7

**报告题目:** Change Point Detection for Multivariate Nonparametric Regression with Deep Neural Networks

**报告人:** 周厚林 安徽大学

**报告摘要:** In this paper, we propose a CUSUM-type test statistic to detect whether there is a change point in multivariate nonparametric regression. The test statistic is based on the estimators obtained by using fully connected deep neural networks (DNNs). DNNs can be adaptive to the intrinsic low-dimension structure of the multivariate



nonparametric regression functions, which empower them to circumvent the curse of dimensionality. We derive the theoretical properties of the proposed test statistic under the null and alternative hypotheses. Moreover, we establish the consistency for the change point location estimator. The finite sample performance of the proposed method is illustrated by Monte Carlo simulations and an empirical data analysis.

---

分会场报告 8

**报告题目：** The Consistency of Estimators in Semiparametric EV Model with Asymptotically Almost Negative Associated Errors and Missing Responses

**报告人：** 葛梅梅 安徽大学

**报告摘要：** In this paper, we consider the semiparametric errors-in-variables (EV) model  $y_i = \xi_i \beta + g(t_i)$ ,  $x_i = \xi_i + \mu_i$ ,  $1 \leq i \leq n$ , where the response variables  $y_i$  are missing,  $(\xi_i, t_i)$  are the design points,  $\xi_i$  are the potential variables observed with measurement errors  $\mu_i$ , and random errors  $\varepsilon_i$  are asymptotically almost negative associated (AANA, in short). The unknown slope parameter  $\beta$  and nonparametric component  $g(\cdot)$  are estimated by three different methods, whose moment consistency and complete consistency are presented under AANA errors and missing responses.



## 第四分会场

分会场报告 5 (邀请报告)

**报告题目:** Lasso-adjusted Treatment Effect Estimation under Covariate-adaptive Randomization

**报告人:** 马维 中国人民大学

**报告摘要:** We consider the problem of estimating and inferring treatment effects in randomized experiments. In practice, stratified randomization, or more generally, covariate-adaptive randomization, is routinely used in the design stage to balance treatment allocations with respect to a few variables that are most relevant to the outcomes. Then, regression is performed in the analysis stage to adjust the remaining imbalances to yield more efficient treatment effect estimators. Building upon and unifying recent results obtained for ordinary-least-squares adjusted estimators under covariate-adaptive randomization, this paper presents a general theory of regression adjustment that allows for model mis-specification and the presence of a large number of baseline covariates. We exemplify the theory on two lasso-adjusted treatment effect estimators, both of which are optimal in their respective classes. In addition, nonparametric consistent variance estimators are proposed to facilitate valid inferences, which work irrespective of the specific randomization methods used. The robustness and improved efficiency of the proposed estimators are demonstrated through numerical studies.

-----

分会场报告 6

**报告题目:** Kernel Cox Partially Linear Regression: Building Predictive Models for Cancer Patients' Survival

**报告人:** 荣耀华 北京工业大学

**报告摘要:** Wide heterogeneity exists in cancer patients' survival, ranging from a few months to several decades. To accurately predict clinical outcomes, it is vital to build an accurate predictive model that relates patients' molecular profiles with patients' survival. With complex relationships between survival and high-dimensional molecular predictors, it is challenging to conduct non-parametric modeling and irrelevant predictors removing simultaneously. In this paper, we build a kernel Cox proportional hazards semi-parametric model and propose a novel regularized garrotized kernel machine (RegGKM) method to fit the model. We use the kernel machine method to describe the complex relationship between survival and predictors, while automatically removing irrelevant parametric and non-parametric predictors through a LASSO penalty. An efficient high-dimensional algorithm is developed for the proposed method. Comparison with other competing methods in simulation shows that the proposed method always has better predictive accuracy. We apply this method to analyze a multiple myeloma dataset and predict patients' death burden based on their





gene expressions. Our results can help classify patients into groups with different death risks, facilitating treatment for better clinical outcomes.

---

分会场报告 7

**报告题目：** Accelerating the Elastic Net Penalized Cox Proportional Hazards Regression Using Safe Screening

报告人：王洪 中南大学

**报告摘要：** The elastic net penalty which includes both the  $L_1$  and  $L_2$  penalties is a widely applied in survival analysis for simultaneous estimation and feature selection. Although many efforts have been devoted to its efficient implementation, its application to high dimensional or large sample sized survival data still poses significant computational challenges. In this paper, we present fast and effective safe screening rules for Cox proportional hazard model with (adaptive) elastic net penalties to identify the zero coefficients in the solution vector, which may lead to a substantial reduction in the number of features before the final optimization procedure. Theoretically, we prove that the algorithm with our screening rules is able to guarantee identical results with the original algorithms. Numerically, we demonstrate that our screening rule improves the efficiency of elastic net Cox proportional hazard model with a significant gain on a variety of simulated and real scenarios without any loss of accuracy.

---

分会场报告 8

**报告题目：** Stochastic Alternating Structure-adapted Proximal Gradient Descent Method with Variance Reduction for Nonconvex

报告人：贾泽慧 南京信息工程大学

**报告摘要：** Group testing has been used extensively to reduce the testing time and the screening costs in epidemiological studies involving low-prevalence diseases. This testing strategy works by first combining specimens (e.g., blood, urine, swabs, etc.) from several individuals to form a pool and then testing the pooled specimen for infection. When the endpoint of interest is a time-to-event outcome, for example, the time until infection or disease, and pools are measured only once, the resulting data are called group-tested current status data (Petito and Jewell, 2016). In this paper, we propose a new type of regression analysis for these data using a semiparametric probit model, an alternative to the proportional hazards model used in survival analysis. A sieve maximum likelihood estimation approach is developed that approximates the model's nonparametric nuisance function with logarithmic monotone splines. To facilitate sieve estimation, we develop a highly efficient expectation-maximization algorithm. The asymptotic properties of the resulting estimators are investigated by using empirical process techniques and sieve estimation theory. Numerical results from simulation studies suggest our proposed method performs nominally, even when pools



are possibly misclassified due to assay error, and can outperform individual testing when the number of assays (tests) is fixed. We illustrate our work by estimating a time to-event regression model for chlamydial infection using group testing data from a large public health laboratory in Iowa.



4月14日 08:15-10:05

### 第一分会场

分会场报告1 (邀请报告)

**报告题目:** Feature Screening for Cluster Analysis Using Quasi-likelihood

报告人: 席瑞斌 北京大学

**报告摘要:** We consider distribution-free feature screening method for ultrahigh dimensional clustering analysis. We propose to screen the variables by investing homogeneity of the variables. Cluster-irrelevant variables should have the homogenous means across different clusters, while cluster-relevant variables should have heterogenous means. The homogeneity is tested by comparing the quasi-likelihoods under the heterogeneous and homogenous model. A MM-test statistic that utilizes the Majorization-Minimization (MM) algorithm is used to maximize the difference between the quasi-likelihood under the heterogeneous and homogenous model. We show that when the number of MM updates approaches infinity at an appropriate rate, MM-test statistic diverges under the heterogeneous model while remains sufficiently small under the homogeneous model with high probability. By setting a suitable threshold, our screening procedure can achieve the consistency in selection property. Our method is computationally efficient, can accurately identify cluster-irrelevant features and significantly improve cluster accuracy, as demonstrated in our simulation and real-data studies.

分会场报告2

**报告题目:** Group Feature Screening via Maximum Information Coefficient (MIC) for Ultra-high Dimensional Multi-Classification

报告人: 陈婷婷 桂林理工大学

**报告摘要:** Data with a group structure in ultra-high dimensional data are becoming increasingly common, particularly in the analysis of genomic data and economic data, where a large amount of information can be lost by univariate filtering methods. Addressing a limitation of existing feature selection techniques, which primarily focus on individual predictors and may not be suitable for structured predictors such as grouped variables, we present a novel method for group screening in classification models. Based on the Maximum Information Coefficient (MIC), our approach is model-free. Compared to conventional methods, it offers several advantages. On the one hand, it identifies important predictors in structured data sets, including grouped variables, in an efficient and accurate manner. On the other hand, our approach shows secure screening properties and consistency of the rankings under certain regularity conditions. The performance of our process is evaluated through simulation studies and real data analyses, which demonstrate the effectiveness of our process in terms of screening efficiency and classification accuracy, including for limited sample size.

分会场报告3

**报告题目:** Partitioning and Aggregating Cross-tissue and Tissue-specific Genetic Effects in Identifying Gene-trait Associations

**报告人:** 宋爽 清华大学

**报告摘要:** Transcriptome-wide association studies (TWAS) have shown great promises in extending GWAS loci to a functional understanding of disease mechanisms. In an effort to fully unleash the TWAS and GWAS information, we propose MTWAS, a statistical framework that partitions and aggregates cross-tissue and tissue-specific genetic effects in identifying gene-trait associations. Different from previous methods, we introduce a non-parametric imputation strategy to augment the inaccessible tissues, which allows for barren conditions, such as complex interactions and non-linear expression data structure across tissues. We further classify eQTLs into cross-tissue eQTLs and tissue-specific eQTLs via a step-wise procedure based on the extended Bayesian information criterion, which was consistent under high-dimensional settings. We have shown that MTWAS significantly improves the imputation accuracy across all 47 tissues in the GTEx dataset and 13 immune cells and 2 activation conditions in the DICE dataset, compared with other single-tissue and multi-tissue methods, such as PrediXcan, TIGAR, and UTMOST. MTWAS also identifies more predictable genes that can be replicated with independent studies. Applications to 84 UKBB GWAS studies have provided novel insights into disease etiology.

---

分会报告 4

**报告题目:** Knockoff-based Statistics for the Identification of Putative Causal Genes in Genetic Studies

**报告人:** 马诗洋 上海交通大学

**报告摘要:** Gene-based tests are important tools for elucidating the genetic basis of complex traits. Despite substantial recent efforts in this direction, the existing tests are still limited, owing to low power and detection of false-positive signals due to the confounding effects of linkage disequilibrium. In this paper, we describe a gene-based test that attempts to address these limitations by incorporating data on long-range chromatin interactions, several recent technical advances for region-based testing, and the knockoff framework for synthetic genotype generation. Through extensive simulations and applications to multiple diseases and traits, we show that the proposed test increases the power over state-of-the-art gene-based tests and provides a narrower focus on the possible causal genes involved at a locus. We apply BIGKnock to the UK Biobank data with 405,296 participants for multiple binary and quantitative traits, and show that relative to conventional gene-based tests, BIGKnock produces smaller sets of significant genes that contain the causal gene(s) with high probability.



## 第二分会场

分会场报告 1 (邀请报告)

**报告题目:** Graphical Models for Temporal Point Processes

报告人: 林伟 北京大学

**报告摘要:** In the era of data-driven research, the analysis of event sequence datasets, such as spike train data and electronic health records, has gained significant importance. Understanding the interaction between different types of events, particularly the exciting and inhibitory dynamics, is of great interest. To formalize this interaction, we propose the use of graphical models of temporal point processes and introduce an interpretable parametrization. Under this parametrization, a point process is locally conditionally dependent on another point process only if the former is directly excited or inhibited by the latter. However, existing methods for parameter estimation, such as maximum likelihood estimation and least squares estimation, are not directly applicable to the nonlinear and high-dimensional case. To address this challenge, we propose a wide class of nodewise weighted  $M$ -estimators and extend them to high-dimensional settings using  $\ell_1$ -regularization. Our proposed estimators can approximate the maximum likelihood estimator and the least squares estimator through a sequence of iterations, and we derive the convergence rate of our estimator. Additionally, we introduce a weighted bias-correction method to construct confidence intervals and perform simultaneous hypothesis tests for edges in the graph. To demonstrate the statistical and computational efficiency of our estimator, we conduct extensive simulations and analyze a spike train dataset from an optogenetics study.

-----

分会场报告 2

**报告题目:** Mathematical Imaging: From Variational Modelling to Deep Learning for Images

报告人: 马倩婷 南京信息工程大学

**报告摘要:** In this talk, we will learn about some of mathematical problems, about variational models for image analysis and inverse imaging problems as well as recent advances where such mathematical models are complemented and replaced by deep neural networks. The talk is furnished with applications to art restoration and image enhancement.

-----

分会场报告 3

**报告题目:** Grouped Heterogeneous Gaussian Graphical Models for High-dimensional Clustered Data

报告人: 曾鑫 厦门大学

**报告摘要:** Clustered data-based analysis has been extensively conducted in various studies. Recent research has demonstrated that a network-based heterogeneity analysis, which adopts a system perspective and incorporates the interconnections among variables while considering heterogeneity between subgroups, can provide more informative results compared to approaches based on simpler statistics. Moreover,

incorporating grouping in analysis can better delineate the sources of heterogeneity and enable more flexible modeling for clustered data. In this article, we introduce a novel approach called the grouped heterogeneous Gaussian graphical models (Grouped-HGGM) for analyzing high-dimensional clustered data with a network structure. Our approach assumes that clusters can be divided into distinct groups, and any heterogeneity across clusters is captured through the cluster-wise mixture probabilities. Unlike most previous approaches which assume that the number of subgroups is known in advance, an appealing feature of our method is the automatic determination of the number of subgroups and sparse estimation by a fusion technique. Consistency properties are rigorously established, and an effective computational algorithm is developed. Extensive simulations demonstrate the practical superiority of the proposed approach over closely related alternatives. In the analysis of breast cancer data, the proposed approach identifies heterogeneity structures different from the alternatives.

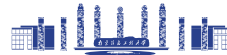
---

分会场报告 4

**报告题目:** Asymptotic Properties of a Multi-colour Random Reinforced Urn Model with Multiple-drawing and Random Addition

报告人: 杨丽 东北师范大学

**报告摘要:** In this paper, we consider a multi-colour, multiple-drawing random reinforced urn model. The replacement matrix is a diagonal random matrix, and at each stage, we allow the number of balls drawn to be not just one, but a random variable. We obtain the limits of the proportions of various colours of balls in the urn and derive their corresponding exact convergence rates. Furthermore, we propose strong consistency estimators for the limit of the means of the diagonal random variables of the replacement matrix and, based on this, we address the hypothesis testing problem for the expected payoffs for each arm of a multi-armed bandit.



## 第三分会场

分会场报告 1 (邀请报告)

**报告题目:** Navigating Challenges in Classification and Outlier Detection: A Remedy Based on Semi-parametric Density Ratio Models

**报告人:** 刘玉坤 华东师范大学

**报告摘要:** The goal of classification is to assign categorical labels to unlabeled test data based on patterns and relationships learned from a labeled training dataset. However, this task becomes challenging when the training data and test data exhibit distributional mismatches. The unlabeled test data follow a finite mixture model, which is not identifiable without any model assumptions. In this paper, we propose to model the test data using a finite semi-parametric mixture model under density ratio model and construct a semi-parametric likelihood prediction set (SPLPS) for the labels in the test data. The goal of our approach is to optimize the out-of-sample performance, with the aim of including the correct class and detecting outliers as often as possible. This approach has the potential to enhance the robustness and effectiveness of classification models when dealing with training and test data with different distributions. Our method circumvents a stringent assumption of separation between training data and outliers, which is required by Guan and Tibshirani (2022) but is often violated by commonly used distributions. We prove the asymptotic consistency and normalities of our parameter estimators and the asymptotic optimality of the proposed SPLPS. We validate our methods by analyzing four real-world datasets.

-----

分会场报告 2

**报告题目:** Bidirectional Efficient Non-convex Adaptive Federated Learning

**报告人:** 杨光 南京审计大学

**报告摘要:** We propose two new strategies within the framework of federated learning: LA (Lazy Aggregation) and AA (Accelerated Aggregation). The LA strategy reduces the costs of communication and computation by adaptively skipping the gradient, and the AA accelerates the computation and reduces the communication cost by adaptively accruing the gradient computation. Based on these novel strategies and compression techniques, we propose two new algorithms: FedBLACA and FedBACA, to reduce bidirectional communication costs. We give a proof of client (full or partial) participation in correlation under non-convex setting. In the nonconvex stochastic optimization full client participation setting, our proposed FedBLACA and FedBACA algorithms achieve the same convergence rate  $O(\sqrt{TKM})$  as its noncompact counterpart. We have demonstrated through extensive experiments that our protocol achieves efficient training in non-convex environments and is robust to large amounts of device, partial participation, and unbalanced data.

## 分会场报告 3

**报告题目:** The Optimality of Kernel Classifiers in Sobolev Space

报告人: 赖建发 清华大学

**报告摘要:** Kernel methods are widely used in machine learning, especially for classification problems. However, the theoretical analysis of kernel classification is still limited. This paper investigates the statistical performances of kernel classifiers. With some mild assumptions on the conditional probability  $\eta(x) = \mathbb{P}(Y = 1 | X = x)$ , we derive an upper bound on the classification excess risk of a kernel classifier using recent advances in the theory of kernel regression. We also obtain a minimax lower bound for Sobolev spaces, which shows the optimality of the proposed classifier. Our theoretical results can be extended to the generalization error of overparameterized neural network classifiers. To make our theoretical results more applicable in realistic settings, we also propose a simple method to estimate the interpolation smoothness of  $2\eta(x) - 1$  and apply the method to real datasets.

-----

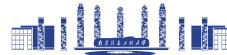
## 分会场报告 4

**报告题目:** Solving Saddle Point Problems: A Landscape of Primal-dual Algorithm with Larger Stepsizes

报告人: 姜帆 南京信息工程大学

**报告摘要:** We consider a class of saddle point problems frequently arising in the areas of image processing and machine learning. In this paper, we propose a simple primal-dual algorithm, which embeds a general proximal term induced with a positive definite matrix into one subproblem. It is remarkable that our algorithm enjoys larger stepsizes than many existing state-of-the-art primal-dual-like algorithms due to our relaxed convergence-guaranteeing condition. Moreover, our algorithm includes the well-known primal-dual hybrid gradient method as its special case, while it is also of possible benefit to deriving partially linearized primal-dual algorithms. Finally, we show that our algorithm is able to deal with multi-block separable saddle point problems. In particular, an application to a multi-block separable minimization problem with linear constraints yields a parallel algorithm. Some computational results sufficiently support the promising improvement brought by our relaxed requirement.





## 第四分会场

分会场报告 1 (邀请报告)

**报告题目:** A Random Graph-based Autoregressive Model for Networked Time Series**报告人:** 吴未迟 清华大学

**报告摘要:** Contemporary time series data often feature objects connected by a social network that naturally induces temporal dependence involving connected neighbours. The network vector autoregressive model is useful for describing the influence of linked neighbours, while its recent generalizations aim to separate influence and homophily. Existing approaches, however, require either correct specification of a time series model or accurate estimation of a network model or both, and rely exclusively on least-squares for parameter estimation. This paper proposes a new autoregressive model incorporating a flexible form for latent variables used to depict homophily. We develop a first-order differencing method for the estimation of influence requiring only the influence part of the model to be correctly specified. When the homophily part is correctly specified admitting a semiparametric form, we leverage and generalize the recent notion of neighbour smoothing for parameter estimation, bypassing the need to specify the generative mechanism of the network. We develop new theory to show that all the estimated parameters are consistent and asymptotically normal. The efficacy of our approach is confirmed via extensive simulations and an analysis of a social media dataset.

-----  
分会场报告 2**报告题目:** Statistical Inference of High-dimensional Vector Autoregressive Time Series with Non-i.i.d. Innovations**报告人:** 张云翼 香港中文大学 (深圳)

**报告摘要:** Independent or i.i.d. innovations is an essential assumption in the literature for analyzing a vector time series. However, this assumption is either too restrictive for a real-life time series to satisfy or is hard to verify through a hypothesis test. This paper performs statistical inference on a sparse high-dimensional vector autoregressive time series, allowing its white noise innovations to be dependent, even non-stationary. To achieve this goal, it adopts a post-selection estimator to fit the vector autoregressive model and derives the asymptotic distribution of the post-selection estimator. The innovations in the autoregressive time series are not assumed to be independent, thus making the covariance matrices of the autoregressive coefficient estimators complex and difficult to estimate. Our work develops a bootstrap algorithm to facilitate practitioners in performing statistical inference without having to engage in sophisticated calculations. Simulations and real-life data experiments reveal the validity of the proposed methods and theoretical results.

Real-life data is rarely considered to exactly satisfy an autoregressive model with independent or i.i.d. innovations, so our work should better reflect the reality compared to the literature that assumes i.i.d. innovations.

分会场报告 3

**报告题目:** Privacy-preserving Parametric Inference for Spatial Autoregressive Model

**报告人:** 宋允全 中国石油大学 (华东)

**报告摘要:** Spatial regression models are important tools in dealing with spatially dependent data and are widely used in many fields such as spatial econometric and regional science. When the spatial data contains sensitive information, the privacy of the data will be compromised along with the release of the analysis if appropriate privacy-preserving measures are not in place. In this paper, we study the privacy-preserving parametric inference for the spatial autoregressive model and propose corresponding differentially private algorithm. We construct a differentially private spatial autoregression framework that takes graph data into account. We improve the functional mechanism to be more accurate under the same degree of privacy protection. Theoretical analysis establishes both the privacy guarantees of the algorithm and the asymptotic normality of the estimation. Simulation and real data studies show improvements of our approach.

-----  
分会场报告 4

**报告题目:** 空间分位面板自回归模型的贝叶斯 Elastic Net 变量选择及其应用

**报告人:** 于卓熙 辽宁大学

**报告摘要:** 本文将 Elastic Net 惩罚方法应用于贝叶斯空间分位面板自回归模型分析, 依据非对称 Laplace 分布推导各参数的满条件后验分布, 通过 MCMC 算法给出参数后验估计及变量选择。在解释变量存在不同相关程度时, 将此方法与 LASSO 分位数回归方法和自适应 LASSO 分位数回归方法进行比较, 并在不同随机扰动项假设和不同样本容量下探究贝叶斯 Elastic Net 分位数回归方法的优势, 结果表明贝叶斯 Elastic Net 分位数回归方法更加适用于解释变量存在相关性和维度较高情形。在实证部分, 本文选取互联网金融、互联网工业、互联网保险和互联网医疗四个概念板块的上市公司经济增加值 (EVA) 作为响应变量, 选取营业利润率等 15 个解释变量, 构建贝叶斯 Elastic Net 空间分位面板自回归模型。实证结果表明总资产收益率、存货周转率、长期资本负债率、已获利息倍数和总资产增长率对 EVA 的影响较大, Elastic Net 惩罚方法剔除了其余 10 个变量。



4月14日 10:20-12:10

### 第一分会场

分会场报告5（邀请报告）

**报告题目：**Model-free Variable Selection in High Dimension via Constrained Kernel Regression

**报告人：**李长城 大连理工大学

**报告摘要：** We propose a non-parametric and model-free variable selection approach called constrained kernel regression. Instead of relying on model-based loss functions, the proposed constrained kernel regression is developed based on conditional independence relationship measured by the conditional distance covariance/correlation. The conditional distance covariance/correlation is approximated by the kernel method. And the constrained kernel regression coefficient vector is defined to be the vector satisfying the conditional independence constraints. By varying the tuning parameter in the conditional independence constraints, the proposed approach provides a solution path. And we establish that with appropriate tuning parameters the proposed approach can consistently identify the true important predictor set under high-dimensional model-free settings. The advantage of the proposed method is further shown by various numerical studies. More specifically, the proposed method outperforms existing model-based in the presence of model misspecification and has better or comparable performance with existing methods when their models are correctly specified.

-----

分会场报告6

**报告题目：**On Selection of Semiparametric Spatial Regression Models with Partial Differential Equation Regularization

**报告人：**刘静 中国矿业大学

**报告摘要：** This study introduces an innovative spatial regression framework that integrates variable selection and spatial anisotropy through partial differential equations (PDEs) analysis. The purpose is to identify influential variables in the parameter space, facilitating a more comprehensive understanding of the relationships between covariates and the spatial outcomes. We incorporate spatial anisotropy into the model using PDEs to represent the underlying spatial processes. We have provided assessments of the consistency of the estimates and their asymptotic normality. We conduct a series of simulations and apply our proposed approach to PM2.5 dataset, demonstrating that our method effectively identifies relevant variables in the parameter space and accurately characterizes the spatial anisotropy.

## 分会场报告 7

**报告题目：** A Two-stage Variable Selection Based on the Adjusted Adaptive Lasso in High-dimensional Cancer Classification

**报告人：** 蒋青青 桂林理工大学

**报告摘要：** Variable selection methods are often used in high-dimensional datasets. Nevertheless, when there is severe multicollinearity among variables, a class of unimportant variables may be mistakenly selected due to their high marginal correlation with some important variables, which is the problem of false positives in variable selection. Therefore, this paper combines the ideas of the screening method with the adaptive lasso and then proposes a two-stage method to deal with this issue. The first stage is the screening stage, in which the maximum marginal likelihood estimates for each variable are used to screen those variables with a high marginal correlation to the samples. The second stage is the selection stage, where an index about the degree of correlation between variables and the sample is added to adjust the initial penalty weight to mitigate the impact of variable selection due to multicollinearity. Afterward, the results of the simulation experiment revealed that our proposed method was able to lower the false-positive rate during variable selection significantly. Finally, three publicly accessible datasets for gene expression were used to test our proposed method. The results demonstrate the method's efficiency in terms of prediction performance as well as its consistency in variable selection. Moreover, it is theoretically proven that our method has Oracle properties in variable selection.

---

## 分会场报告 8

**报告题目：** fastCCLasso: Fast and Efficient Algorithm for Estimating Correlation Matrix from Compositional Data

**报告人：** 张琄 首都师范大学

**报告摘要：** The composition and structure of microbial communities on the body surface are closely related to human health. The interaction relationship among microbes can help us understand the formation of the microecological environment and the biological mechanism by which microorganisms influence host health. With the help of high-throughput sequencing technologies, microbial abundances in a natural environment can be directly measured without the isolation of microorganisms in culture. Sequencing experiments in microbiome studies can measure the relative abundances of microbes, which is called compositional data. Although there are already many methods for correlation analysis for compositional data, the computation time or accuracy still needs to be improved for current microbiome studies. We develop a fast and efficient algorithm, called fastCCLasso, based on a penalized weighted least squares for inferring the correlation structure of microbes from compositional data in microbiome studies. We perform a large number of numerical experiments and the simulation results show that fastCCLasso outperforms its competitors in edge detection for inferring the correlation network. We also apply fastCCLasso for estimating microbial networks in microbiome studies and fastCCLasso provides a conservative network with a comparable false discovery counts that are derived from shuffled data.



## 第二分会场

分会场报告 5 (邀请报告)

**报告题目:** Causal Inference in Randomized Experiments for Dyadic Data with Interference

报告人: 苗旺 北京大学

**报告摘要:** Estimating the global average treatment effect on a network could be considerably biased due to spillover effects in the presence of unknown network interference. We consider novel dyadic outcomes in the presence of interference. Such outcomes are common in many social network sources, such as forwarding a message or sharing a link. We first establish the causal inference framework for dyadic outcome in network interference, which is of particular interest in online experimentation in many social media companies. Then we show that the unbiased estimator for the global average treatment effect based on the conventional outcomes does not exist in most cases. We provide subsequently unbiased estimators based on dyadic outcomes for randomized experiments. We derive the variance bound of the proposed estimators and provide a variance estimator for quantifying the estimation uncertainty. We also illustrate with a variety of numerical experiments.

-----

分会场报告 6

**报告题目:** Multivariate Spatiotemporal Models with Low Rank Coefficient Matrix

报告人: 蒲丹 西南财经大学

**报告摘要:** Multivariate spatiotemporal data arise frequently in practical applications, often involving complex dependencies across cross-sectional units, time points and multivariate variables. In the literature, few studies jointly model the dependence in three dimensions. To simultaneously model the cross-sectional, dynamic and cross-variable dependence, we propose a multivariate reduced-rank spatiotemporal model. By imposing the low-rank assumption on the spatial influence matrix, the proposed model achieves substantial dimension reduction and has a nice interpretation, especially for financial data. Due to the innate endogeneity, we propose the quasi-maximum likelihood estimator (QMLE) to estimate the unknown parameters. A ridge-type ratio estimator is also developed to determine the rank of the spatial influence matrix. We establish the asymptotic distribution of the QMLE and the rank selection consistency of the ridge-type ratio estimator. The proposed methodology is further illustrated via extensive simulation studies and two applications to a stock market dataset and an air pollution dataset.



分会场报告 7

**报告题目:** Correcting Spot Power Variation Estimator via Edgeworth Expansion

**报告人:** 何丽丹 南京信息工程大学

**报告摘要:** In this report, we propose an estimator of power spot volatility of order  $p$  through Edgeworth expansion. We provide a precise description of how to compute the expansion and the first four cumulants are given in an explicit form. We also construct feasible confidence intervals (one-sided and two-sided) for the  $p$ th power spot volatility estimator by using Edgeworth expansion. A Monte Carlo simulation study shows that the confidence intervals and probability density curve based on Edgeworth expansion perform better than the conventional case based on Normal approximation.

-----

分会场报告 8

**报告题目:** Generalization Error Curves for Analytic Spectral Algorithms under Power-law Decay

**报告人:** 李易诚 清华大学

**报告摘要:** The generalization error curve of certain kernel regression method aims at determining the exact order of generalization error with various source condition, noise level and choice of the regularization parameter rather than the minimax rate. In this work, under mild assumptions, we rigorously provide a full characterization of the generalization error curves of the kernel gradient descent method (and a large class of analytic spectral algorithms) in kernel regression. Consequently, we could sharpen the near inconsistency of kernel interpolation and clarify the saturation effects of kernel regression algorithms with higher qualification, etc. Thanks to the neural tangent kernel theory, these results greatly improve our understanding of the generalization behavior of training the wide neural networks. A novel technical contribution, the analytic functional argument, might be of independent interest.



## 第三分会场

分会场报告 5 (邀请报告)

**报告题目:** Construction of High-dimensional Designs for Computer Experiments

报告人: 何煦 中国科学院数学与系统科学研究院

**报告摘要:** Space-filling designs that possess high separation distance are useful for computer experiments. We propose a novel method to construct high-dimensional high-separation distance designs. The construction involves taking the Kronecker product of sub-Hadamard matrices and rotation. In addition to possessing better separation distance than most existing types of space-filling designs, our newly proposed designs enjoy orthogonality and projection uniformity and are more flexible in the numbers of runs and factors than that from most algebraic constructions. From numerical results, such designs are excellent in Gaussian process emulation of high-dimensional computer experiments. An R package on design construction is available online.

-----

分会场报告 6

**报告题目:** Gaussian Variational Approximation with Composite Likelihood for Crossed Random Effect Models

报告人: 徐礼柏 苏州大学

**报告摘要:** Composite likelihood usually ignores dependencies among response components, while variational approximation to likelihood ignores dependencies among parameter components. What both methods have in common is that they break the dependence of random effects essentially. In this paper, we derive a Gaussian variational approximation to the composite log-likelihood function for Poisson and Gamma regression models with crossed random effects. We show consistency and asymptotic normality of the estimates derived from this approximation and support this theory with some simulation studies. The approach is computationally faster than a Gaussian variational approximation to the full log-likelihood function.

-----

分会场报告 7

**报告题目:** Influence Exploration of Estimating Procedure on Adequacy Test of Semiparametric Censored Regression Models

报告人: 孙志华 中国科学院大学

**报告摘要:** In this work, the adequacy check of partially linear (PL) models is investigated when the response variable is randomly right-censored. A weighted least squares (WLS) method is first developed to estimate the null hypothetical model. Then, an empirical-process-based test with a linear-indicator (LI) weighting function is proposed, and the asymptotic properties of the test statistic are rigorously proved. It is shown that the proposed test possesses several desiring merits: consistency, computationally expediency, the effect of dimension reduction, being free of distribution assumption, and the ability of detecting Pitman alternative hypothesis. This study also mainly focuses on exploring the influence of the estimating method on the model checking procedure. An interesting but long-neglected fact is uncovered that



sometime the estimating procedures are extremely important for model checking methods. Simulation studies and real-data analyses are presented to illustrate the performances of the proposed method.

---

分会场报告 8

**报告题目:** Simultaneous Variable Selection and Estimation for Survival Data via the Gaussian Seamless- $L_0$  Penalty

**报告人:** 刘自力 中南大学

**报告摘要:** We propose a new simultaneous variable selection and estimation procedure with the Gaussian seamless- $L_0$  (GSELO) penalty for Cox proportional hazard model and additive hazards model. The GSELO procedure shows good potential to improve the existing variable selection methods by taking strength from both best subset selection (BSS) and regularization. In addition, we develop an iterative algorithm to implement the proposed procedure in a computationally efficient way. Theoretically, we establish the convergence properties of the algorithm and asymptotic theoretical properties of the proposed procedure. Since parameter tuning is crucial to the performance of the GSELO procedure, we also propose an extended Bayesian information criteria (EBIC) parameter selector for the GSELO procedure. Simulated and real data studies have demonstrated the prediction performance and effectiveness of the proposed method over several state-of-the-art methods.





## 第四分会场

分会场报告 5（邀请报告）

**报告题目：**度量空间回归的随机森林方法及理论

**报告人：**於州 华东师范大学

**报告摘要：**当代大数据愈发呈现复杂性状，往往表现为复杂的非欧几里得结构，比如大脑功能连接研究中的对称正定矩阵可视为黎曼流形结构数据。为了更好的对这一类非欧数据进行统计建模，我们借鉴近代标志性机器学习方法随机森林的内蕴统计思想，进一步发展了度量空间回归的随机森林方法。基于无限阶  $U$  过程理论我们建立了该方法的相合性、收敛速度以及渐近正态性。而现有欧式空间的随机森林方法及理论可以视为新方法和理论的一个特殊情形。模拟试验以及实际非欧数据应用验证了该方法的有效性。

分会场报告 6

**报告题目：**儿科临床大数据分析面临的机遇与挑战

**报告人：**徐铤明 重庆医科大学附属儿童医院

**报告摘要：**临床数据既是对病人诊疗过程的详实记录，更是通过数字和文本等形式对医生临床经验与智慧的客观反应。在本报告中，报告人将基于重庆医科大学附属儿童医院千万级规模临床数据实践经验，通过几个实例说明海量临床数据中所蕴含的巨大价值（科研与应用转化），以及在数据处理和统计分析过程中所面临的挑战。

分会场报告 7

**报告题目：**大模型的成功与高维数据算法存在的问题

**报告人：**夏慧异 池州学院

**报告摘要：**很多计算科学研究者认为以 ChatGPT 为代表的大模型的成功是“蛮力出奇迹”，但报告人的研究表明这种说法是错误的，“大模型的成功”是“科学出奇迹”，产生“蛮力出奇迹”观点的原因是现行的高维数据算法存在问题。本报告首先介绍高维数据算法存在的问题，2004 年，美国科学院院士 Efron 等人给出了 lasso 的快速算法 Lars，使得梯度下降算法、坐标下降算法等算法成为 lasso 的快速算法；2010 年，报告人开始研究 lasso，发现这些 lasso 的快速算法都存在问题。其次发现大模型的成功原因就是梯度下降算法等 lasso 的快速算法在大模型的情况下，可以解决 lasso 问题，大模型的成功是科学出奇迹。最后指出由于美国科学院院士 Efron 等人的错误，使得学界普遍认为中小模型就可以解决实际问题，认为大模型是不科学的，使得学界研究方向出了问题。2014 年，如果中国学界能认真对待报告人指出美国科学院院士的错误，肯定会发现中小模型解决实际问题的方法不对，大模型的成功有可能会出现在中国，而不是在美国。由于美国科学院院士的错误，使得很多研究还在无效空转，因此纠错工作迫在眉睫。



分会场报告 8

**报告题目:** Quantile Generalized Measures of Correlation

**报告人:** 张新瑜 北京师范大学

**报告摘要:** In this paper, we introduce a quantile Generalized Measure of Correlation (GMC) to describe nonlinear quantile relationship between response variable and predictors. The introduced correlation takes values between zero and one. It is zero if and only if the conditional quantile function is equal to the unconditional quantile. We also introduce a quantile partial Generalized Measure of Correlation. Estimators of these correlations are developed. Notably by adopting machine learning methods, our estimation procedures allow the dimension of predictors very large. Under mild conditions, we establish the estimators' consistency. For construction of confidence interval, we adopt sample splitting and show that the corresponding estimators are asymptotic normal. We also consider composite quantile GMC by integrating information from different quantile levels. Numerical studies are conducted to illustrate our methods. Moreover, we apply our methods to analyze genome-wide association study data from Carworth Farms White mice.



## 通讯录（按工作单位首字母顺序）

序号	姓名	工作单位	电子邮箱
1	林希虹	Harvard University	xlin@hsph.harvard.edu
2	李凡群	安徽财经大学	lfq006@163.com
3	李琴	安徽财经大学	liqin_math@163.com
4	刘文锐	安徽财经大学	1169628364@qq.com
5	孙军	安徽财经大学	sunj@aufe.edu.cn
6	王文雅	安徽财经大学	3236813661@qq.com
7	吴厚武	安徽财经大学	3542813735@qq.com
8	邢雪	安徽财经大学	2725539520@qq.com
9	方红燕	安徽大学	fanhoy@ustc.edu.cn
10	葛梅梅	安徽大学	784812678@qq.com
11	唐徐飞	安徽大学	Xufei_tang@163.com
12	王冰	安徽大学	bingwang@ahu.edu.cn
13	王丽娜	安徽大学	bingwang@ahu.edu.cn
14	王苗苗	安徽大学	mmwang2020@163.com
15	王星惠	安徽大学	wangxinghui@ahu.edu.cn
16	王学军	安徽大学	wxjahdx2000@126.com
17	王燕	安徽大学	wyahdx2017@126.com
18	吴捷	安徽大学	jiewu@ahu.edu.cn
19	席梦梅	安徽大学	ximengmei123@163.com
20	杨文志	安徽大学	wzyang@ahu.edu.cn
21	张正龙	安徽大学	15567710520@163.com
22	周厚林	安徽大学	Houlin972022@163.com
23	朱汉兵	安徽大学	zhuhbecnu@163.com
24	曹明响	安徽师范大学	caomingx@163.com
25	何道江	安徽师范大学	djhe@ahnu.edu.cn
26	贺磊	安徽师范大学	lhstat@163.com
27	时慧君	安徽师范大学	huijunshi2021@163.com
28	许凯	安徽师范大学	tjxxukai@163.com



29	胡学平	安庆师范大学	hxprob@163.com
30	高艺丹	北方民族大学	574578988@qq.com
31	黄夕	北方民族大学	Huangxi202107@163.com
32	李萌	北方民族大学	2305375730@qq.com
33	毛浩阳	北方民族大学	1003976972@qq.com
34	申菊梅	北方民族大学	shenjumei2839@163.com
35	王涵	北方民族大学	2388317384@qq.com
36	王小刚	北方民族大学	10266770@qq.com
37	郭爽	北华大学	309192145@qq.com
38	孙佩宇	北华大学	704046867@qq.com
39	艾明要	北京大学	myai@pku.edu.cn
40	陈松蹊	北京大学	csx@gsm.pku.edu.cn
41	林伟	北京大学	weilin@math.pku.edu.cn
42	苗旺	北京大学	mwfy@pku.edu.cn
43	邱宇谋	北京大学	qiuyumou@math.pku.edu.cn
44	席瑞斌	北京大学	ruibinxi@math.pku.edu.cn
45	王学丽	北京工商大学	xlwang@bftu.edu.cn
46	陈巍璐	北京工业大学	chenweilu@bjut.edu.cn
47	程维虎	北京工业大学	chengweihu@bjut.edu.cn
48	荣耀华	北京工业大学	rongyaohua@bjut.edu.cn
49	吴密霞	北京工业大学	wumixia@bjut.edu.cn
50	谢琨	北京工业大学	xieli@bjut.edu.cn
51	薛留根	北京工业大学	lgxue@bjut.edu.cn
52	李高荣	北京师范大学	ligaorong@bnu.edu.cn
53	李望成	北京师范大学	2290726821@qq.com
54	饶昊宸	北京师范大学	rhc_bnu_stat@mail.bnu.edu.cn
55	史延美	北京师范大学	13165150979@163.com
56	张新瑜	北京师范大学	zhangxinyu1@mail.bnu.edu.cn
57	赵俊龙	北京师范大学	zhaojunlong928@126.com
58	朱力行	北京师范大学	lzhu@bnu.edu.cn
59	梁丰	北京师范大学珠海校区	lfstats@163.com
60	张晓晨	北京师范大学珠海校区	zhangxiaochen@bnu.edu.cn



61	苏苗苗	北京邮电大学	smm@bupt.edu.cn
62	郝江锋	巢湖学院	1219223965@qq.com
63	彭维才	巢湖学院	56833965@qq.com
64	吴焱	池州学院	wuyi8702@163.com
65	夏慧异	池州学院	xiayz_88@163.com
66	邓新	滁州学院	Tzdx0120@163.com
67	郑伟	楚雄师范学院	zw@cxtc.edu.cn
68	李长城	大连理工大学	lichangcheng@dlut.edu.cn
69	李江豪	东北师范大学	lijh458@nenu.edu.cn
70	杨丽	东北师范大学	yangl245@nenu.edu.cn
71	马少沛	对外经济贸易大学	mshp@uibe.edu.cn
72	熊巍	对外经济贸易大学	xiongwei@uibe.edu.cn
73	叶鹏	对外经济贸易大学	yepeng@amss.ac.cn
74	高小燕	甘肃农业大学	452285288@qq.com
75	梁玥	甘肃农业大学	105090990@qq.com
76	马鑫	甘肃农业大学	maxin263@126.com
77	秦丽娟	甘肃农业大学	qinlj@gsau.edu.cn
78	周彦秋	广西科技大学	zhouyanqiu@163.com
79	张崇岐	广州大学	cqzhang@gzhu.edu.cn
80	胡尧	贵州大学	yhul@gzu.edu.cn
81	陈婷婷	桂林理工大学	ctt991116@163.com
82	蒋青青	桂林理工大学	jqqq312425@163.com
83	陆军	国防科技大学	junlu@nudt.edu.cn
84	马晓雨	国防科技大学	maxiaoyu@nudt.edu.cn
85	赵振宇	国防科技大学	zhaoyu@mail.ustc.edu.cn
86	姚梅	合肥工业大学	ymnancy@sina.com
87	解俊山	河南大学	junshan@henu.edu.cn
88	刘中强	河南理工大学	zhongqiang@hpu.edu.cn
89	王照良	河南理工大学	wangzhaoliang@hpu.edu.cn
90	陈文	湖南大学	chenwen@cupk.edu.cn
91	贺姿	湖南大学	3335226423@qq.com
92	胡月	湖南大学	575673268@qq.com



93	刘杰	湖南大学	JieL000@hnu.edu.cn
94	汤山	湖南大学	3462735397@qq.com
95	程学伟	湖南师范大学	xwcheng@hunnu.edu.cn
96	杨晶	湖南师范大学	yj_stst@163.com
97	田苗	华东理工大学	tianmiao@ecust.edu.cn
98	湛自奇	华东师范大学	zqchen@fem.ecnu.edu.cn
99	刘玉坤	华东师范大学	ykliu@sfs.ecnu.edu.cn
100	汤银才	华东师范大学	yctang@stat.ecnu.edu.cn
101	王光辉	华东师范大学	ghwang.nk@gmail.com
102	王小舟	华东师范大学	xzawang@sfs.ecnu.edu.cn
103	於州	华东师范大学	zyu@stat.ecnu.edu.cn
104	赵伟	华东师范大学/《统计理论及其应用（英文）》编辑部	wzhao@xb.ecnu.edu.cn
105	王涛	淮阴师范学院	skywangtao2007@126.com
106	梁文娟	黄山学院	wjliang0627@163.com
107	陈望学	吉首大学	chenwangxue2015@163.com
108	黄晓雯	暨南大学	1149179660@qq.com
109	邹航	暨南大学	zouhang19980304@163.com
110	陈雪平	江苏理工学院	chenxueping@jsut.edu.cn
111	陈慧	江苏师范大学	chenhui@163.com
112	李建波	江苏师范大学	lijianbo@jsnu.edu.cn
113	任鹏程	江苏师范大学	renpengcheng12@jsnu.edu.cn
114	赵鹏	江苏师范大学	zhaop@jsnu.edu.cn
115	胡天英	江西财经大学	1141994328@qq.com
116	刘显慧	江西财经大学	liuxh109@126.com
117	谭祥勇	江西财经大学	xiangyongtan@126.com
118	王勇	江西财经大学	ywang@jxufe.edu.cn
119	邓文丽	江西师范大学	wldfudan@126.com
120	冯卓熙	江西师范大学	864056872@qq.com
121	滕雯雯	江西师范大学	1140622460@qq.com
122	任苏灵	兰州财经大学	rensuling2009@163.com
123	张继红	兰州城市学院	zhzhzhangjihong@163.com



124	王歆旸	辽宁大学	wxy_math@163.com
125	熊玮	辽宁大学	xiongwei16@mails.jlu.edu.cn
126	姚雪	辽宁大学	2011014866@qq.com
127	于卓熙	辽宁大学	yzx8170561@163.com
128	刘群	闽南师范大学	604011685@qq.com
129	赵洋	南昌大学	yangz@ncu.edu.cn
130	陈海清	南京财经大学	chenhq@nufe.edu.cn
131	孟慧珍	南京财经大学	menghz2001@163.com
132	钱永淇	南京财经大学	1330788406@qq.com
133	王平平	南京财经大学	catherinewpp@126.com
134	吴星月	南京财经大学	wuxingyuedaytime@126.com
135	潘茂林	南京大学	mlpan@nju.edu.cn
136	陈鹏	南京航空航天大学	chenpengmath@nuaa.edu.cn
137	黄振生	南京理工大学	stahzs@126.com
138	袁亮亮	南京理工大学	llyuan_dlut@163.com
139	温阳俊	南京农业大学	wenyangjun@njau.edu.cn
140	张瑾	南京农业大学	zhangjin@njau.edu.cn
141	付金玉	南京审计大学	fujinyu163@163.com
142	郝红霞	南京审计大学	hxhao@nau.edu.cn
143	黄性芳	南京审计大学	xfhuang@nau.edu.cn
144	孔新兵	南京审计大学	xinbingkong@126.com
145	汪红霞	南京审计大学	hxwang@nau.edu.cn
146	杨光	南京审计大学	guangyang_nau@163.com
147	岳莉莉	南京审计大学	yueliliz@163.com
148	赵彦勇	南京审计大学	yyzhao@nau.edu.cn
149	周兴才	南京审计大学	xczhou@nau.edu.cn
150	王明泉	南京师范大学	wmqhhh@126.com
151	曹春正	南京信息工程大学	caochunzheng@nuist.edu.cn
152	陈金凤	南京信息工程大学	1982784675@qq.com
153	陈特	南京信息工程大学	1803578610@qq.com
154	陈文珠	南京信息工程大学	202212150025@nuist.edu.cn
155	董文正	南京信息工程大学	228700970@qq.com



156	杜莹莹	南京信息工程大学	820172084@qq.com
157	付春蕊	南京信息工程大学	2012086059@qq.com
158	公颜笑	南京信息工程大学	2459276767@qq.com
159	何丽丹	南京信息工程大学	hld@nuist.edu.cn
160	化理想	南京信息工程大学	1137212753@qq.com
161	黄丽	南京信息工程大学	13103970135@163.com
162	贾冰	南京信息工程大学	001745@nuist.edu.cn
163	贾泽慧	南京信息工程大学	jiazehui90@126.com
164	姜帆	南京信息工程大学	15905154902@163.com
165	靳雨红	南京信息工程大学	rainredj@163.com
166	来鹏	南京信息工程大学	laipeng@nuist.edu.cn
167	兰天涯	南京信息工程大学	echo1592022@163.com
168	刘文军	南京信息工程大学	wjliu@nuist.edu.cn
169	吕广迎	南京信息工程大学	gylvmaths@nuist.edu.cn
170	马倩婷	南京信息工程大学	qtma@nuist.edu.cn
171	宋凤丽	南京信息工程大学	songfl@nuist.edu.cn
172	田汶鑫	南京信息工程大学	tianwx2905@163.com
173	童凤茹	南京信息工程大学	1429609381@qq.com
174	王博慧	南京信息工程大学	2625633837@qq.com
175	王辉	南京信息工程大学	h_wang@nuist.edu.cn
176	王洲	南京信息工程大学	516889159@qq.com
177	卫雨婷	南京信息工程大学	003839@nuist.edu.cn
178	武业文	南京信息工程大学	ywwu@nuist.edu.cn
179	杨建伟	南京信息工程大学	yjianw@nuist.edu.cn
180	杨洲木	南京信息工程大学	001170@nuist.edu.cn
181	张婷	南京信息工程大学	zhangtstat@nuist.edu.cn
182	张毅	南京信息工程大学	zhangy2016@nuist.edu.cn
183	张宇荣	南京信息工程大学	zhangyurong0105@163.com
184	张泽文	南京信息工程大学	20211215010@nuist.edu.cn
185	周胜	南京信息工程大学	782092209@qq.com
186	朱连华	南京信息工程大学	ahualian@126.com
187	朱妍萱	南京信息工程大学	z83310062@163.com





188	羊忱	南京邮电大学	xiaoxiandada@outlook.com
189	张晶	南京邮电大学	s20082112095@163.com
190	李忠华	南开大学	zli@nankai.edu.cn
191	赵为华	南通大学	zhaowhstat@163.com
192	王延新	宁波工程学院	yxwang@nbut.edu.cn
193	徐晨	鹏城国家实验室	xuch01@pcl.ac.cn
194	陈占寿	青海师范大学	chenzhsh@qhnu.edu.cn
195	刁丽	青海师范大学	xilixili2008@163.com
196	赖建发	清华大学	244259931@qq.com
197	李易诚	清华大学	liyc22@mails.tsinghua.edu.cn
198	林乾	清华大学	qianlin@tsinghua.edu.cn
199	宋爽	清华大学	songs15@tsinghua.org.cn
200	吴未迟	清华大学	wuweichi@tsinghua.edu.cn
201	陈晓林	曲阜师范大学	xlchen@amss.ac.cn
202	王少新	曲阜师范大学	shxwang@qfnu.edu.cn
203	曾鑫	厦门大学	xmuzengxin@163.com
204	刘婧媛	厦门大学	Jingyuan@xmu.edu.cn
205	王中雷	厦门大学	wangzl@xnu.edu.cn
206	张庆昭	厦门大学	zhangqingzhao@amss.ac.cn
207	胡琴琴	山东大学	qqhu@sdu.edu.cn
208	林路	山东大学	linlu@sdu.edu.cn
209	倪培培	山东大学	2725381623@qq.com
210	严晓东	山东大学	yanxiaodong@sdu.edu.cn
211	赵艳艳	山东大学	yyzhao824@163.com
212	李华鹏	山西大同大学	lihuaa1_0@163.com
213	张秀珍	山西大同大学	zhangxiuzhen132@163.com
214	罗瑞苗	山西大学	luorm955@sxu.edu.cn
215	陈夏	陕西师范大学	xchen80@snnu.edu.cn
216	张耀武	上海财经大学	zhang.yaowu@mail.shufe.edu.cn
217	姜荣	上海第二工业大学	jiangrong@sspu.edu.cn
218	罗琳	上海第二工业大学	luolin@sspu.edu.cn
219	赵岩斌	上海第二工业大学	zhaoyb553@nenu.edu.cn



220	李睿	上海对外经贸大学	lirui@suibe.edu.cn
221	刘关福	上海对外经贸大学	liuguanfu07@163.com
222	邹玉叶	上海海事大学	zouyuye@shmtu.edu.cn
223	刘卫东	上海交通大学	weidongl@sjtu.edu.cn
224	马诗洋	上海交通大学	mashiyang1991@sjtu.edu.cn
225	毛晓军	上海交通大学	maoxj@sjtu.edu.cn
226	牟荣吉	上海交通大学	rjmu@sjtu.edu.cn
227	庞道琳	上海交通大学	pangdaolin@sjtu.edu.cn
228	任好洁	上海交通大学	haojieren@sjtu.edu.cn
229	吴鑑洪	上海师范大学	wujianhong@shnu.edu.cn
230	周彦	深圳大学	zhouy1016@szu.edu.cn
231	蔡婷婷	首都师范大学	2220501029@cnu.edu.cn
232	方李君	首都师范大学	fang_lijun@163.com
233	张坤	首都师范大学	zhangshen2010@126.com
234	邹国华	首都师范大学	ghzou@amss.ac.cn
235	刘洋	苏州大学	liuyangecnu@163.com
236	徐礼柏	苏州大学	lxbu@suda.edu.cn
237	刘国风	天津财经大学	tjliuguofeng@163.com
238	秦珊珊	天津财经大学	qinss@tjufe.edu.cn
239	李姝妤	同济大学	2523664567@qq.com
240	廖雨苇	同济大学	475665360@qq.com
241	于晓戈	同济大学	1424163588@qq.com
242	赵湘茹	同济大学	2311752@tongji.edu.cn
243	周叶青	同济大学	zhouyeqing@tongji.edu.cn
244	黄辉林	温州大学	30683894@qq.com
245	吴艳	西安电子科技大学	yanwu@mail.xidian.edu.cn
246	杨有龙	西安电子科技大学	ylyang@mail.xidian.edu.cn
247	惠雨馨	西安工程大学	xyx980321@163.com
248	刘持金	西安交通大学	824364470@qq.com
249	刘懿	西安交通大学	yiyi0305@stu.xjtu.edu.cn
250	施蒙洁	西安交通大学	1208368566@qq.com
251	张誉腾	西安交通大学	zytnlnlzn@outlook.com



252	朱学虎	西安交通大学	zhuxuehu@xjtu.edu.cn
253	王丹	西北大学	wangdan@nwu.edu.cn
254	夏志明	西北大学	statxzm@nwu.edu.cn
255	田玉柱	西北师范大学	pole1999@163.com
256	蒲丹	西南财经大学	pudan666@163.com
257	张佛德	西南财经大学	fredzh@swufe.edu.cn
258	常晋源	西南财经大学、中国科学院数学与系统科学研究院	changjinyuan@swufe.edu.cn
259	黄磊	西南交通大学	yellones@163.com
260	张云翼	香港中文大学（深圳）	zhangyunyi@cuhk.edu.cn
261	李翠霞	徐州工程学院	79419383@qq.com
262	李万斌	盐城师范学院	liwanbinc@163.com
263	杨晓洁	云南财经大学	yangxiaojie@amss.ac.cn
264	王涛	云南师范大学	wtaokm@263.net
265	李晗	长春大学	lihancu@163.com
266	李岩	长春大学	lyan2020@126.com
267	王文钊	长春工业大学	794364352@qq.com
268	徐悦	长春工业大学	972388216@qq.com
269	杨凯	长春工业大学	yangkai@ccut.edu.cn
270	赵乃非	长沙理工大学	zhaonf@csust.edu.cn
271	罗季	浙江财经大学	20000018@zufe.edu.cn
272	何尹	浙江大学	heyin__i@163.com
273	贾小可	浙江大学	jiaxiaoke2021@outlook.com
274	金寅	浙江大学	12035022@zju.edu.cn
275	骆威	浙江大学	wluostat@outlook.com
276	叶舟夫	浙江大学	zhoufuyue@zju.edu.cn
277	蔡利	浙江工商大学	caili16@126.com
278	王江峰	浙江工商大学	wjf2929@163.com
279	郑顺平	浙江工商大学	zsp12052022@163.com
280	张立新	浙江工商大学、浙江大学	stazlx@zju.edu.cn
281	刘祎	中国海洋大学	liuyi@ouc.edu.cn
282	曾靖	中国科学技术大学	zengjxl@ustc.edu.cn



283	丁浩	中国科学技术大学	dinghao@ustc.edu.cn
284	胡彦林	中国科学技术大学	hyl11@mail.ustc.edu.cn
285	李亚光	中国科学技术大学	liyg@ustc.edu.cn
286	卢超	中国科学技术大学	luchao2022@ustc.edu.cn
287	杨青	中国科学技术大学	yangq@ustc.edu.cn
288	郑泽敏	中国科学技术大学	zhengzm@ustc.edu.cn
289	张三国	中国科学院大学	sgzhang@ucas.ac.cn
290	张正军	中国科学院大学经济与管理学院、中国科学院预测科学中心	zhangzhengjun@ucas.ac.cn
291	李轩宇	中国科学院大学数学科学学院	LixuanYu22@mails.ucas.ac.cn
292	孙志华	中国科学院大学数学科学学院	sunzh@ucas.ac.cn
293	王孟浩	中国科学院大学数学科学学院	1171635231@qq.com
294	何煦	中国科学院数学与系统科学研究院	hexu@amss.ac.cn
295	李冬雨	中国科学院数学与系统科学研究院	lidongyu@amss.ac.cn
296	刘子辰	中国科学院数学与系统科学研究院	liuzichen22@mails.ucas.ac.cn
297	吕福润	中国科学院数学与系统科学研究院	1467738506@qq.com
298	潘越峰	中国科学院数学与系统科学研究院	panyuefeng16@mails.ucas.ac.cn
299	盛赢	中国科学院数学与系统科学研究院	shengying@amss.ac.cn
300	石坚	中国科学院数学与系统科学研究院	jshi@iss.ac.cn
301	王启华	中国科学院数学与系统科学研究院	qhwan@amss.ac.cn
302	张新雨	中国科学院数学与系统科学研究院	xinyu@amss.ac.cn
303	李再兴	中国矿业大学（北京）	lzxgscas@126.com
304	刘静	中国矿业大学（北京）	1548588374@qq.com
305	黄辉	中国人民大学	huangh89@ruc.edu.cn
306	李杰	中国人民大学	lijie_stat@ruc.edu.cn
307	李伟	中国人民大学	weilistat@ruc.edu.cn
308	马维	中国人民大学	wma16@foxmail.com
309	陈璇	中国石油大学（华东）	1255469767@qq.com
310	侯一鸣	中国石油大学（华东）	1059457471@qq.com
311	李洁	中国石油大学（华东）	1617424663@qq.com
312	宋允全	中国石油大学（华东）	syqfly1980@163.com
313	王元凤	中国石油大学（华东）	wyeartha@163.com



314	王之鉴	中国石油大学（华东）	893737924@qq.com
315	杨睿	中国石油大学（华东）	yr_1232001@163.com
316	刘自力	中南大学	zili_liu@csu.edu.cn
317	王洪	中南大学	wh@csu.edu.cn
318	盖玉洁	中央财经大学	yujie_gai@126.com
319	胡雪梅	重庆工商大学	huxuem@163.com
320	廖书	重庆工商大学	8621978@qq.com
321	秦湘斌	重庆工商大学	qinxiangbin@ctbu.edu.cn
322	杨炜明	重庆工商大学	yangweiming@ctbu.edu.cn
323	徐铎明	重庆医科大学附属儿童医院	ximing@hospital.cqmu.edu.cn



## 南京信息工程大学 ▶▶

南京信息工程大学是国家“双一流”建设高校，是江苏高水平大学建设高峰计划A类建设高校。学校办学特色鲜明，大气科学入选国家“世界一流学科”建设学科，在教育部最新学科评估中取得优异成绩。气象学为国家重点学科，地球科学、工程学、计算机科学、环境科学与生态学、化学、农业科学、材料科学、社会科学总论、数学等9个学科跻身ESI学科排名全球前1%，其中地球科学和计算机科学进入ESI全球排名前1%。拥有大气科学、环境科学与工程、信息与通信工程、管理科学与工程、数学、科学技术史、计算机科学与技术7个一级学科博士学位授权点，25个一级学科硕士学位授权点、21个硕士专业学位授权点。学校现有“大气科学”“环境科学与工程”“数学”“管理科学与工程”“计算机科学与技术”“信息与通信工程”6个博士后科研流动站，80个本科专业分布于理、工、文、管、经、法、农、艺、教9个学科领域。

学校秉承“艰苦朴素、勤奋好学、追求真理、自强不息”的优良校风，恪守“明德格物、立己达人”的校训，坚持以人才培养为中心，以培养拔尖精英人才、创新创业人才、国际化人才为导向，不断深化教育教学改革，构建了特色鲜明的人才培养体系。建校以来，已培养各类毕业生20余万人，校友中涌现出一批两院院士、部委领导、央企高管、战略专家、国际组织官员等杰出人才，众多校友成为中国乃至世界气象行业的业务骨干和科研精英，学校被誉为“气象人才的摇篮”。

展望“十四五”，南京信息工程大学将深入学习贯彻习近平新时代中国特色社会主义思想，宣传落实党的二十大精神，全面贯彻党的教育方针，以立德树人为根本，以内涵发展、高质量发展为永恒主题，以服务支撑国家和地方创新驱动发展为战略导向，聚焦“一流特色研究型大学”战略目标，坚持“开放、协同、特色”发展理念，秉承“笃行以生为本、厚植大学精神”办学宗旨，“面向行业、面向地方、面向国际”，踔厉奋发，笃行不怠，着力培养德智体美劳全面发展的社会主义建设者和接班人，全面提升学校核心竞争力和综合办学实力，为全面建设社会主义现代化国家、实现中华民族伟大复兴的中国梦提供有力支撑、作出更大贡献。





## 数学与统计学院 ▶▶

数学学科是我校大气科学一流学科建设的重要支撑，已有 60 多年的发展历史。数学与统计学院 1978 年招收首届数学本科生，现设有数学系、信息与计算科学系、统计系和公共数学教学部等机构，建有江苏省应用数学中心、江苏省系统建模与数据分析国际合作联合实验室、



江苏省统计科学研究基地、外国专家工作室、实验教学与实践教育示范中心等省部级平台，同时参与共建国家应用数学中心。

学院建立了完整的本、硕、博培养体系，拥有“数学”一级学科博士点与硕士点、“统计学”一级学科硕士点及“应用统计”专业学位硕士点，设有“数学”博士后流动站；数学学科入选“十四五”江苏省重点学科、江苏高校优势学科建设工程四期项目立项学科，进入 ESI 学科排名全球前 1%。

目前学院拥有全时教职工 110 余人，其中专业系教师 82 人（含外籍教师 6 人），高级职称 58 人。具有博士学位占比 98%，85% 以上的专任教师具有海外工作或学习经历。教师队伍中拥有法兰西科学院院士、国际杰青计划、澳门青年学者计划、江苏省双创人才、江苏省“六大高峰”人才、江苏省“青蓝工程”中青年学术带头人、江苏省产业教授等各类国家及省部级人才 60 余人次。

学院设有数学与应用数学、信息与计算科学、应用统计学 3 个本科专业。“信息与计算科学”和“应用统计学”为国家一流专业建设点；“数学与应用数学”为江苏省一流专业建设点，入选江苏省数学拔尖学生培养计划 2.0 基地。学院培养了包括国家特聘专家、国家优青、欧美终身教授、谢义炳青年气象科技奖、联合国世界气象组织青年专业官员在内的一批有着重要影响的杰出人才。

欢迎浏览数学与统计学院官方网站  
(<https://math.nuist.edu.cn/main.htm>) 以及关注“NUIST 数声朗朗”公众号，获取更多关于学院的最新资讯。

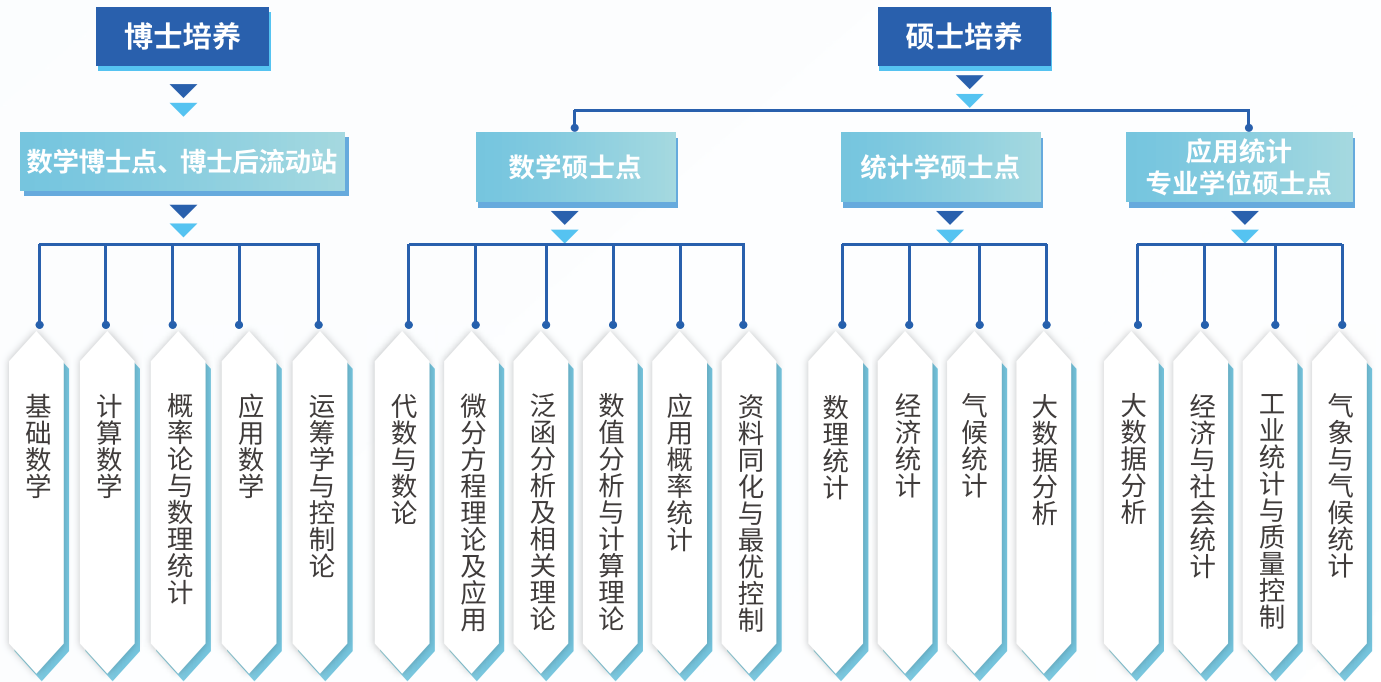


数学与统计学院微信公众号

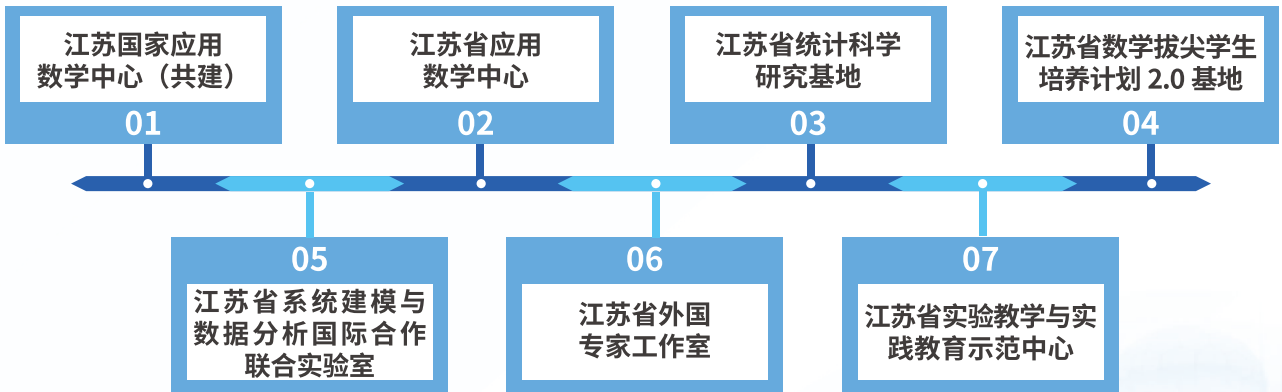


## 基本情况

### 学科体系



### 平台资源



### 教学资源

#### 国家级课程



#### 省级教材 5 部

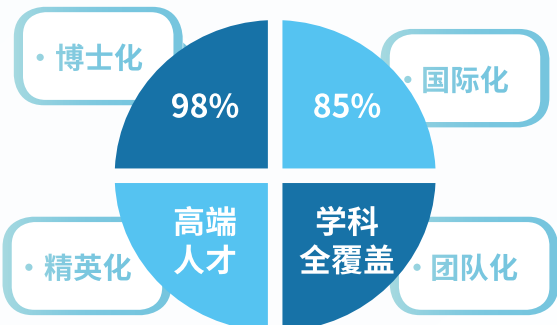






### 师资队伍

专业系教师 82 人，高级职称 58 人，占比 70.1%



### 国际化

- 不同国别的全时外籍教师 6 名
- 国际杰青计划科学家 3 名
- 国家自然科学基金外青项目 3 项、国合项目 2 项
- 科技部外专项目 2022、2023 年获批 6 项
- 江苏省科技厅外国专家工作室 1 项
- 江苏省系统建模与数据分析国际合作联合实验室 1 项
- 海外半年以上工作、学习经历者占比 85%

### 校企合作

- 江苏省“产业教授” 5 名
- 江苏省科技厅企业科技副总 2 名
- 江苏省科协首席专家（工程师） 1 名

### 高层次人才

- 法兰西科学院院士 1 名
- 国家重大人才工程 A 类人才 1 名
- 澳门青年学者计划 1 名
- 江苏省“双创计划”各类人才 9 名
- 江苏省“六大人才高峰”人才 3 名
- 江苏省“333 工程人才” 2 名
- 中国气象局教学团队 2 个
- 江苏青蓝工程优秀教学团队 1 个
- 江苏青蓝工程中青年学术带头人 4 名
- 江苏青蓝工程优秀青年骨干教师 8 名
- 江苏省工业与应用数学青年奖 2 名

### 数学学科

基础数学	动力系统与气象应用	计算数学	概率统计与气候诊断
环与代数表示论、拓扑学、微分方程、调和分析。 <b>特色:</b> 在罗巴代数、Domain 理论、随机微分方程等方面研究在国际上具有一定影响	生物动力系统、空间磁动力系统、卫星资料同化。 <b>特色:</b> 解决气象、海洋、生物、医学等领域的关键数学问题，形成数学为主导的应用技术。	高精度算法的设计与分析、大数据基础算法、信息处理与应用。 <b>特色:</b> 将微分方程数值解、模式识别、人工智能等应用于雷达、卫星数据处理与医学诊断。	统计学习理论与方法、复杂数据统计建模与推断、统计气候及其应用。 <b>特色:</b> 研究高维、函数型复杂数据的统计推断，以及气候诊断与预测问题驱动的现代统计方法。

### 统计学学科

数理统计	经济统计	气候统计	大数据分析
统计建模、推断和诊断的理论、方法及其应用研究。 <b>特色:</b> 研究纵向数据、函数型数据等复杂数据的统计推断理论和统计诊断理论和方法，并应用于辅助精准医疗；研究高维复杂数据的降维和特征筛选方法，并应用于文本分类、基因筛选。	经济、金融、管理科学等领域的现代统计方法和应用研究。 <b>特色:</b> 开展气象、环境领域统计调查、经济评估与预测研究，研究成果应用于雾霾经济损失评估、开展统计质量控制研究，为行业 and 部门提供管理咨询服务。	气候中的极值理论、关联分析、统计物理、统计诊断及预测研究。 <b>特色:</b> 研究多重关联分析和极值理论，并应用于极端气候诊断与预测；研究基于统计物理的台风气候动力学，并应用于我国台风活动的季节预测。	统计机器学习与气象资料的数据同化方法研究。 <b>特色:</b> 研究统计模式识别、遥感图像处理，并应用于气象灾害监测预警；研究卫星遥感气象资料的数据同化方法，并应用于台风、强降水预报。

