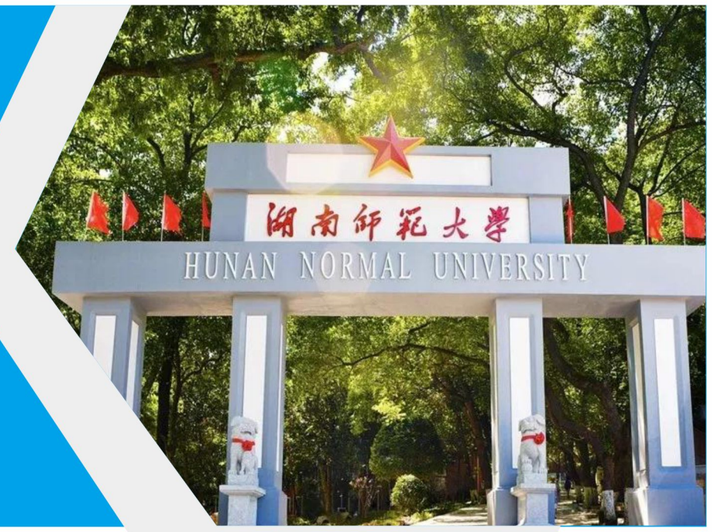
仁爱精勤

湖南師範大學
HUNAN NORMAL UNIVERSITY

中国现场统计研究会高维数据统计分会

第七届学术研讨会

程 序 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：湖南师范大学数学与统计学院

2023年5月19-21日 长沙

目 录 CONTENTS

会议委员会

程序委员会

主席

王启华

委员

艾明要　林华珍　唐年胜　王兆军　张日权

朱利平　邹国华

组织委员会

主席

郭水霞

委员

杨　晶　刘思晟　赵　维　耿旖旎　周杰明

陈　旭　张墨力　李　娟

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七届学术研讨会，本着圆满完成此次会议宗旨，现将有关事宜宣告如下：

一、酒店安排

长沙茉莉花国际酒店（长沙市岳麓区金星中路 528 号）

长沙金麓国际大酒店（长沙市岳麓区岳麓大道 311 号）

二、就餐安排

日期	就餐时间	就餐地点
5 月 19 日	晚餐 18: 00-19: 30	茉莉花国际酒店
5 月 20 日	早餐 06: 30-08: 00	所住酒店
	午餐 12: 20-13: 30	茉莉花国际酒店
	晚宴 18: 00-19: 30	茉莉花国际酒店
5 月 21 日	早餐 06: 30-08: 15	所住酒店
	午餐 12: 10-14: 00	茉莉花国际酒店

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，会务组将及时提前通知。

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。

3、 会议期间，请遵守疫情防控相关政策，如无特殊情况，请不要离会单独活动。

4、 会务组联系人：

杨晶：13548785657;　　刘思晟：18974166797

日程简表

日期	时间	事项	地点
5月19日（星期五）	09: 00-20: 00	会议报到、注册	茉莉花国际酒店
	18: 00-19: 30	晚餐	茉莉花国际酒店
5月20日（星期六）	06: 30-08: 00	早餐	所住酒店
	08: 10	乘车前往湖南师大附中报告厅	所住酒店门口
	08: 45-09: 30	开幕式、合影	湖南师大附中报告厅（道德讲堂）
	09: 30-10: 15	大会特邀报告 1	
	10: 15-11: 00	大会特邀报告 2	
	11: 00-11: 15	茶歇	
	11: 15-12: 00	大会特邀报告 3	
	12: 00	乘车回酒店用餐	湖南师大附中
	12: 20-13: 30	午餐	茉莉花国际酒店
	13: 30-17: 40	第一分会场报告	15 楼茉莉 1 厅
	13: 30-17: 40	第二分会场报告	15 楼茉莉 3 厅
	13: 30-17: 40	第三分会场报告	15 楼贵宾厅
	15: 20-15: 40	茶歇	茉莉花国际酒店
	18: 00-19: 30	晚宴	茉莉花国际酒店
5月21日（星期日）	06: 30-08: 15	早餐	所住酒店
	08: 15-12: 10	第一分会场报告	15 楼茉莉 1 厅
	08: 15-12: 10	第二分会场报告	15 楼茉莉 3 厅
	08: 15-12: 10	第三分会场报告	15 楼贵宾厅
	10: 05-10: 20	茶歇	茉莉花国际酒店
	12: 10-14: 00	午餐	茉莉花国际酒店
	14: 00-	自由交流、离会	

会议议程

5 月 19 日

会议报到 9: 00-20: 00

晚餐 18: 00-19: 30 **地点：**茉莉花国际酒店

5 月 20 日 上午

乘车前往湖南师大附中报告厅 8:10

开幕式及合影 8:45-9:30 **地点：**湖南师大附中报告厅（道德讲堂）

主持人：郭水霞 副院长 湖南师范大学数学与统计学院

湖南师范大学副校长 谢资清教授 致辞

中国现场统计研究会理事长 郭建华教授 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 **地点：**湖南师大附中

大会特邀报告 9:30-12:00 **地点：**湖南师大附中报告厅（道德讲堂）

大会特邀报告 1 9:30-10:15

报告题目: Statistical Inference on a Four-regime Segmented Regression Model

报告人：陈松蹊 北京大学

主持人：王启华 中国科学院数学与系统科学研究院

大会特邀报告 2 10:15-11:00

报告题目：Transfer Learning: Optimality and Adaptive Algorithms

报告人：Tony Cai 宾夕法尼亚大学

主持人：林华珍 西南财经大学

茶歇 时间: 11:00-11:15

大会特邀报告 3 11:15-12:00

报告题目：Large-scale Detection of Differential Sparsity Structure

报告人：邹长亮 南开大学

主持人：邹国华 首都师范大学

乘车回酒店用餐 12:00

午餐 12:20-13:30 **地点：**茉莉花国际酒店

5 月 20 日 下午

分会场报告（20.1.1-20.3.8） 13: 30-17: 40 **地点：**茉莉花国际酒店 15 楼

第一分会场（20.1.1-20.1.4） **地点: 15 楼茉莉 1 厅**

主持人: 张立新 浙江大学

1、邀请报告（20.1.1） 13:30-14:05

报告题目：Multi-Consensus Decentralized Primal-dual Fixed Point Algorithm for Distributed Learning

报告人: 刘卫东 上海交通大学

2、报告（20.1.2） 14:05-14:30

报告题目：Communication-efficient Distributed Estimation of Partially Linear Additive Models for Large-scale Data

报告人: 王磊 南开大学

3、报告（20.1.3） 14:30-14:55

报告题目: A Two-stage Optimal Subsampling Estimation for Missing Data Problems with Large-scale Data

报告人: 苏苗苗 北京邮电大学

4、**报告（20.1.4）** 14:55-15:20

报告题目: Estimation and Inference for Ultra-high Dimensional Quasi-likelihood Models Based on Data Splitting

报告人: 蒋学军 南方科技大学

第二分会场（20.2.1-20.2.4）　　地点: 15 楼茉莉 3 厅

主持人: 石坚　中科院数学与系统科学研究院

1、邀请报告（20.2.1） 13:30-14:05

报告题目: Model-free Test and Measure for Partial Mean Dependence

报告人: 钟威　　厦门大学

2、报告（20.2.2） 14:05-14:30

报告题目: Estimation and Order Selection for Multivariate Exponentional Power Mixture Models

报告人: 冯峥晖　哈尔滨工业大学（深圳）

3、报告（20.2.3） 14:30-14:55

报告题目: Tensor t-distribution and Tensor Response Regression

报告人: 王宁　　北京师范大学

4、报告（20.2.4） 14:55-15:20

报告题目: A Weighted Iterative Projection Estimation Procedure for Robust Tensor Factor Model with Tucker Decomposition

报告人: 胡雪梅　　重庆工商大学

第三分会场（20.3.1-20.3.4）　　地点: 15 楼贵宾厅

主持人：张忠占　北京工业大学

1、邀请报告（20.3.1） 13:30-14:05

报告题目: High Dimensional (Mean) Independence Tests Based on Rank Indices in the Presence of Heterogeneity

报告人: 朱利平　　中国人民大学

2、报告（20.3.2） 14:05-14:30

报告题目: Score Function-based Tests for Ultrahigh-Dimensional Linear Models

报告人: 郭旭　　北京师范大学

3、报告（20.3.3） 14:30-14:55

报告题目: Detecting communities in attributed networks through bi-direction penalized clustering and its application

报告人：杨虎　　中央财经大学

4、报告（20.3.4） 14:55-15:20

报告题目：New Tests for High-Dimensional Two-sample Mean Problems with Consideration of Correlation Structure

报告人：杨松山　　中国人民大学

茶歇 时间: 15:20-15:40

第一分会场（20.1.5-20.1.8）　　地点：15 楼茉莉 1 厅

主持人：艾明要　北京大学

1、邀请报告（20.1.5） 15:40-16:15

报告题目：深度非线性降维

报告人：於州　　华东师范大学

2、邀请报告（20.1.6） 16:15-16:50

报告题目：Regulation-incorporated Network-based Heterogeneity Analysis

报告人：张庆昭　　厦门大学

3、报告（20.1.7） 16:50-17:15

报告题目：非线性期望与统计机器学习

报告人：严晓东　　山东大学

4、报告（20.1.8） 17:15-17:40

报告题目：Robust Transfer Learning in High-dimensional GLM via γ-divergence

报告人：徐福芝　　厦门大学

第二分会场（20.2.5-20.2.8）　　地点：15 楼茉莉 3 厅

主持人：薛留根　河南大学

1、邀请报告（20.2.5） 15:40-16:15

报告题目：Tensor Model Averaging

报告人：张新雨　　中国科学院

2、邀请报告（20.2.6） 16:15-16:50

报告题目：Introducing the Specificity Score: Measuring the Credibility of Causal Associations

报告人：苗旺　　北京大学

3、报告（20.2.7） 16:50-17:15

报告题目：Rank-based Greedy Model Averaging for High-dimensional Survival Data

报告人：贺百花　　中国科学技术大学

4、报告（20.2.8） 17:15-17:40

报告题目：Time-varying Complete Subset Averaging in a Data-rich Environment

报告人：李海奇　　湖南大学

第三分会场（20.3.5-20.3.8） **地点：15 楼贵宾厅**

主持人：林路　山东大学

1、邀请报告（20.3.5） 15:40-16:15

报告题目：Statistical Inferences for Complex Dependence of Multimodal Imaging Data

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

2、邀请报告（20.3.6） 16:15-16:50

报告题目：The Aggregation-heterogeneity Trade-off in Federated Learning

报告人：林伟　　北京大学

3、报告（20.3.7） 16:50-17:15

报告题目：Sequential Data Integration under Dataset Shift

报告人：盛赢　　中国科学院数学与系统科学研究院

4、报告（20.3.8） 17:15-17:40

报告题目：Federated Survival Analysis via Data Augmentation Using Multi-task Variational Autoencoder

报告人：王洪　　中南大学

晚餐： 18:00-19:30　　地点：茉莉花国际酒店

5 月 21 日 上午

分会场报告（21.1.1-21.3.8） 8: 15-12: 10 **地点：**茉莉花国际酒店 15 楼

第一分会场（21.1.1-21.1.4） **地点: 15 楼茉莉 1 厅**

主持人: 朱仲义 复旦大学

1、邀请报告（21.1.1） 8:15-8:50

报告题目: Feature Screening for Clustering Analysis

报告人: 席瑞斌 北京大学

2、报告（21.1.2） 8:50-9:15

报告题目: Generalization Ability of Wide Neural Networks

报告人: 赖建发 清华大学

3、报告（21.1.3） 9:15-9:40

报告题目: Citation Counts Prediction of Statistical Publications Based on Multi-layer Academic Networks via Neural Network Model

报告人: 潘蕊 中央财经大学

4、**报告（21.1.4）** 9:40-10:05

报告题目: Dimension Reduction for Covariates in Network Data

报告人: 赵俊龙 北京师范大学

第二分会场（21.2.1-21.2.4） **地点: 15 楼茉莉 3 厅**

主持人: 刘小惠 江西财经大学

1、邀请报告（21.2.1） 8:15-8:50

报告题目: Tuning-parameter-free Propensity Score Matching Approach for Causal Inference Under Shape Restriction

报告人: 刘玉坤 华东师范大学

2、报告（21.2.2） 8:50-9:15

报告题目: Novel Statistical Methods for High-dimensional Microbiome Data Analysis

报告人: 占翔 北京大学

3、报告（21.2.3） 9:15-9:40

报告题目: High-dimensional Response Growth Curve Modeling for Longitudinal Neuroimaging Analysis

报告人：王璐　　中南大学

4、报告（21.2.4） 9:40-10:05

报告题目：Generalized Liquid Association Analysis for Multimodal Data Integration

报告人：曾靖　　中国科学技术大学

第三分会场（21.3.1-21.3.4） **地点：15 楼贵宾厅**

主持人：程维虎　北京工业大学

1、邀请报告（21.3.1） 8:15-8:50

报告题目：Group-orthogonal Subsampling for Big Data Linear Mixed Models

报告人：孙法省　　东北师范大学

2、报告（21.3.2） 8:50-9:15

报告题目：Debiased and Thresholded Ridge Regression for Linear Nodels with Heteroskedastic and Correlated Errors

报告人：张云翼　　香港中文大学（深圳）

3、报告（21.3.3） 9:15-9:40

报告题目：Sketched Ridgeless Linear Regression: The Statistical Role of Sketching

报告人：曾奕程　　深圳市大数据研究院

4、报告（21.3.4） 9:40-10:05

报告题目：Classified Generalized Linear Mixed Model Prediction Incorporating Pseudo-prior Information

报告人：马海强　　江西财经大学

茶歇 时间: 10:05-10:20

第一分会场（21.1.5-21.1.8） **地点：15 楼茉莉 1 厅**

主持人：张日权　上海对外经贸大学

1、邀请报告（21.1.5） 10:20-10:55

报告题目：广义离群随机矩阵的极限理论、统计推断和交叉应用

报告人：姜丹丹　　西安交通大学

2、报告（21.1.6） 10:55-11:20

报告题目: Weighted Residual Empirical Processes, Martingale Transformations, and Model Checking for Regressions

报告人: 谭发龙 湖南大学

3、报告（21.1.7） 11:20-11:45

报告题目: Regularized t Distribution: Definition, Properties and Application

报告人: 胡宗良 深圳大学

4、报告（21.1.8） 11:45-12:10

报告题目: Randomization-based Joint Central Limit Theorem and Efficient Covariate Adjustment in Randomized Block 2K Factorial Experiments

报告人: 杨玥含 中央财经大学

第二分会场（21.2.5-21.2.8） **地点: 15 楼茉莉 3 厅**

主持人: 岳荣先 上海师范大学

1、邀请报告（21.2.5） 10:20-10:55

报告题目: One-way or Two-way Factor Model for Matrix Sequences?

报告人: 孔新兵 南京审计大学

2、报告（21.2.6） 10:55-11:20

报告题目: 基于结构化随机森林的地震波初至拾取方法

报告人: 张春霞 西安交通大学

3、报告（21.2.7） 11:20-11:45

报告题目: Functional Additive Expectile Regression in the Reproducing Kernel Hilbert Space

报告人: 刘育孜 江西财经大学

4、报告（21.2.8） 11:45-12:10

报告题目: A General Framework to Select Tuning Parameter for Nonparametric Derivative Estimation.

报告人: 刘思晟 湖南师范大学

第三分会场（21.3.5-21.3.8） **地点: 15 楼贵宾厅**

主持人: 张崇岐 广州大学

1、<u>邀请报告（21.3.5）</u> 10:20-10:55

报告题目: Heterogeneity-aware Clustered Distributed Learning for Multi-source Data Analysis

报告人: 方匡南 厦门大学

2、报告（21.3.6） 10:55-11:20

报告题目: Distribution-free Simultaneous Prediction Bands for Clustered Data with Missing Responses

报告人: 唐炎林 华东师范大学

3、报告（21.3.7） 11:20-11:45

报告题目: Renewable Quantile Regression for Streaming Data Sets

报告人: 姜荣 上海第二工业大学

4、报告（21.3.8） 11:45-12:10

报告题目: Smooth-threshold GMM Estimation of Higher-order Semiparametric Spatial Autoregressive Model with Increasing Dimension

报告人: 杨晶 湖南师范大学

午餐: 12:10-14:00 地点：茉莉花国际酒店

自由交流、离会: 14:00-

报告题目和摘要（按报告顺序）

大会特邀报告 1: Statistical Inference on a Four-Regime Segmented Regression Model

报告人：陈松蹊 教授　北京大学

摘要：Segmented regression models are attractive for their flexibility and interp retability as compared to the global parametric and the nonparametric models, and yet are challenging in both estimation and inference. We consider a four-re gime segmented model for temporally dependent data with two segmenting bou ndaries depending on multivariate covariates with non-diminishing boundary eff ects. A mixed integer quadratic programming algorithm is formulated to facilita te the least square estimation to both the regression and the boundary coefficie nts. The rates of convergence and the asymptotic distributions of the least squa re estimators are obtained, which shows differential convergence rates and limit ing distributions between the regression and the boundary coefficients. Estimati on and testing for degenerated segmented models with less than four segments are also considered with a testing procedure to decide if neighboring segments can be merged. Numerical simulations and a case study on air pollution in B eijing are conducted to demonstrate the proposed model and the inference resul ts. In particular, it shows that the segmented models with three or four regime s are suitable for the modeling of the meteorological effects on the PM2.5 con centration.

大会特邀报告 2: Transfer Learning: Optimality and Adaptive Algorithms

报告人：Tony Cai　Professor　宾夕法尼亚大学

摘要：Human learners have the natural ability to use knowledge gained in one setting for learning in a different but related setting. This ability to transfer knowledge from one task to another is essential for effective learning. However, in statistical learning, most procedures are designed to solve one single task, or to learn one single distribution, based on observations from the same setting. In this talk, we discuss statistical transfer learning in various settings under the posterior drift model, which is a general framework and arises in many practical problems. The results show that significant benefit of incorporating data from the source distributions for learning under the target distribution.

大会特邀报告 3: Large-scale Detection of Differential Sparsity Structure

报告人：邹长亮 教授　南开大学

摘要： Two-sample multiple testing has a wide range of applications. Most of the literature considers simultaneous tests of equality of parameters. This work takes a different perspective and investigates the null hypotheses that the two support sets are equal. This formulation of the testing problem is motivated by the fact that in many applications where the two parameter vectors being compared are both sparse, we might be more concerned about the detection of differential sparsity structures rather than the difference in parameter magnitudes. A general approach to problems of this type is developed via a novel double thresholding (DT) filter. The DT filter first constructs a sequence of pairs of ranking statistics that fulfill global symmetry properties, and then chooses two data-driven thresholds along the ranking to simultaneously control the false discovery rate (FDR) and maximize the number of rejections. Several applications of the methodology are given, including tests for large-scale correlation matrices, high-dimensional linear models and Gaussian graphical models.

分会场邀请报告（20.1.1）：Multi-Consensus Decentralized Primal-dual Fixed Point Algorithm for Distributed Learning

报告人：刘卫东　　上海交通大学

摘要： Decentralized distributed learning has recently attracted significant attention in many applications in machine learning and signal processing. To solve a decentralized optimization with regularization, we propose a Multi-consensus Decentralized Primal-Dual Fixed Point (MD-PDFP) algorithm. We apply multiple consensus steps with the gradient tracking technique to extend the primal-dual fixed point method over a network. The communication complexities of our procedure are given under certain conditions. Moreover, we show that our algorithm is consistent under general conditions and it enjoys global linear convergence under strong convexity. With some particular choices of regularizations, our algorithm can be applied to decentralized machine learning applications. Finally, several numerical experiments and real data analyses are conducted to demonstrate the effectiveness of the proposed algorithm.

分会场报告（20.1.2）：Communication-efficient Distributed Estimation of Partially Linear Additive Models for Large-scale Data

报告人：王磊　　南开大学

摘要： Distributed estimation for parametric models has drawn attention in modern statistical learning, but few studies focus on semiparametric models. In this paper, we propose two communication-efficient distributed estimators for partially linear additive models with high-dimensional covariates. The commonly used B-spline basis functions are first applied to approximate the nonparametric functions and then we construct a profiled communication-efficient surrogate loss function with Lasso penalty based on one local machine solving the final optimization problem. Further, to reduce the effect of local machines and improve the stability of the algorithm, a profiled gradient-enhanced loss estimator is derived. The resulting two estimators and their theoretical convergence rates for both parametric and nonparametric components are established. The finite-sample performance of the proposed estimators is studied through simulations and an application to appliances energy prediction data set is also presented.

分会场报告（20.1.3）: A Two-Stage Optimal Subsampling Estimation for Missing Data Problems with Large-Scale Data

报告人：苏苗苗　　北京邮电大学

摘要： Subsampling is useful to downsize data volumes and speed up calculations for large-scale data and is well studied with completely observed data. In the presence of missing data, computation is more challenging and subsampling becomes more crucial and complex. However, there is still a lack of studies on subsampling for missing data problems. This paper fills the gap by studying the subsampling method for a widely used missing data estimator, the augmented inverse probability weighting (AIPW) estimator. The response mean estimation problem with missing responses is discussed for illustration. A two-stage subsampling method is proposed via Poisson sampling framework. A small subsample of expected size n1 is used in the first stage to estimate the parameters in the propensity score and the outcome regression models, while a larger subsample of expected size n2 is used in the computationally simple second stage to calculate the final estimator. An attractive property of the resulting estimator is that its convergence rate is n_{2}^{−1/2} rather than n_{1}^{−1/2} when

both the propensity score and the outcome regression functions are correctly specified. The rate n_{2}^{−1/2} is still attainable for some important cases if only one of the two functions is correctly specified. This indicates that using a small subsample in the computationally complex first stage can reduce the computational burden with little impact on the statistical accuracy. The asymptotic normality of the resulting estimator is established and the optimal subsampling probability is derived by minimizing the asymptotic variance of the resulting estimator. Simulations and a real data analysis were conducted to demonstrate the empirical performance of the resulting estimator.

分会场报告（20.1.4）：Estimation and Inference for Ultra-high Dimensional Quasi-likelihood Models Based on Data Splitting

报告人：蒋学军 南方科技大学

摘要： In this article, we develop a valid weighted estimation and inference framework for ultra-high dimensional quasi-likelihood models. The weighted estimator is obtained by minimizing the variance function. We split the full data into two subsets, conduct the model selection on one subset and compute the maximum quasi-likelihood estimator on the other subset. Then we aggregate the two estimators with the optimal weighted matrixes to form the final weighted estimator. With the weighted estimator, we construct the confidence intervals for the group components of the regression vector, and the Wald test for the linear structure of the group components. Theoretically, we establish the asymptotic normality of the weighted estimator, and the asymptotic χ^2-distribution of the corresponding Wald test without assuming model selection consistency. Advantages of the proposed tests are highlighted via theoretical and empirical comparison to some competitive tests, which guarantees that our proposed estimation and inference framework is locally optimal. In addition, when the selection consistency is achieved, we prove that the proposed Wald test is asymptotically identically distributed as the oracle tests in the sense that it knows the support of regression vector. Extensive simulations demonstrate more favorable finite sample performance of the proposed tests. An application to the breast cancer data illustrates the use of our proposed methodology.

分会场邀请报告（20.2.1）：Model-free Test and Measure for Partial Mean Dependence

报告人：钟威 厦门大学

摘要：It is of importance to investigate the significance of a subset of covariates W for the response Y given covariates Z in regression modeling. To this end, we propose a significance test for partial mean independence based on deep neural networks and data splitting. The test statistic converges to the standard chi-squared distribution under the null hypothesis while it converges to a normal distribution under the alternative hypothesis. We suggest a powerful ensemble algorithm based on multiple data splitting to enhance the testing power. If the null hypothesis is rejected, we propose a partial Generalized Measure of Correlation (pGMC) to measure the partial mean dependence of Y given W after controlling for the nonlinear effect of Z. We present the theoretical properties of the pGMC and establish the asymptotic normality of its estimator with the optimal root-N converge rate. Furthermore, the valid confidence interval for the pGMC is also derived. As an important special case when there is no conditional covariates Z, we consider a new test of overall significance of covariates for the response in a model-free setting. Numerical studies and real data analysis are conducted to compare with existing approaches and to illustrate the validity of our procedures.

分会场报告（20.2.2）：Estimation and Order Selection for Multivariate Exponential Power Mixture Models

报告人：冯峥晖 哈尔滨工业大学（深圳）

摘要：For multivariate non-Gaussian density estimation and approximation, in this paper, we consider to use multivariate exponential power mixture models. We use the penalized-likelihood method with generalized EM algorithm to estimate locations, scale matrices, shape parameters and mixing probabilities. Order selection is achieved simultaneously. Properties of the estimated order have been derived. Although we mainly focus on the unconstrained scale matrix type in multivariate exponential power mixture models, three more parsimonious type of scale matrix have also been considered. Performance based on simulation and real data analysis imply the parsimony of the exponential power mixture models, and verify the properties on order selection.

分会场报告（20.2.3）: Tensor t-distribution and Tensor Response Regression

报告人: 王宁 北京师范大学

摘 要： In recent years, promising statistical modeling approaches to tensor data analysis have been rapidly developed. Traditional multivariate analysis tools, such as multivariate regression and discriminant analysis, are now generalized from modeling random vectors and matrices to higher-order random tensors (a.k.a. array-valued random objects). Equipped with tensor algebra and high-dimensional computation techniques, concise and interpretable statistical models and estimation procedures prevail in many applications. One of the biggest challenges to statistical tensor models is the non-Gaussian nature of many real-world data. Unfortunately, existing approaches are either restricted to normality or implicitly using least squares type objective functions that are computationally efficient but sensitive to data contamination. Motivated by this, we propose a simple tensor t-distribution that is, unlike existing matrix t-distributions, compatible with tensor operators and reshaping of the data. We then study the tensor response regression with tensor t-error, and develop penalized estimation and hypothesis testing under this t-modeling approach. A novel one-step estimation algorithm is developed for penalized non-convex optimization and is proven to converge to the global optimum. We study the asymptotic relative efficiency of various estimators under this model and establish the oracle properties in variable selection and near-optimal asymptotic efficiency. Extensive numerical studies show the encouraging performance of the one-step estimator.

分会场报告（20.2.4）: A Weighted Iterative Projection Estimation Procedure for Robust Tensor Factor Model with Tucker Decomposition

报告人: 胡雪梅 重庆工商大学

摘 要： Tensor Factor Models(TFM) have been proposed as appealing dimension reduction tools for tensor-valued time series. In this paper we construct a robust TFM with Tucker decomposition by replacing the least square loss by the modified Huber loss whose tail function substituted by an exponential squared loss, propose a robust weighted iterative projection estimation procedure for loading matrices, latent dynamic tensor factors and signal parts, and develop their asymptotic consistent properties and faster convergence rates. Both two simulation experiment examples

and a real example show that the proposed robust procedure outperforms the existing methods in terms of robust estimation and tensor reconstruction.

分会场邀请报告（20.3.1）：High Dimensional (Mean) Independence Tests Based on Rank Indices in the Presence of Heterogeneity

报告人：朱利平 中国人民大学

摘 要： In the big data era, how to deal with heterogeneous observations is an inevitable and important issue. We consider testing (mean) independence in the presence of heterogeneity. To be precise, in the first part of the talk, we consider testing for the effects of high-dimensional covariates on the response. In the second part of the talk, we propose three tests to test independence between two high-dimensional random vectors based on the rank-based indices. To simultaneously accommodate such heteroscedasticity and high dimensionality, we propose novel tests based on aggregations, requiring no prior information on the specific form of models. Our proposed test statistics are scale-invariance, tuning-free and convenient to implement. We further study the asymptotic relative efficiency of our proposed test with respect to the state-of-art universal tests.

分会场报告（20.3.2）：Score Function-based Tests for Ultrahigh-Dimensional Linear Models

报告人：郭旭 北京师范大学

摘要：To sufficiently exploit the model structure under the null hypothesis such that the conditions on the whole model can be mild, this paper investigates score function-based tests to check the significance of an ultrahigh-dimensional sub-vector of the model coefficients when the nuisance parameter vector is also ultrahigh-dimensional in linear models. We first reanalyze and extend a recently proposed score function-based test to derive, under weaker conditions, its limiting distributions under the null and local alternative hypotheses. As it may fail to work when the correlation between testing covariates and nuisance covariates is high, we propose an orthogonalized score function-based test with two merits: debiasing to make the non-degenerate error term degenerate and reducing the asymptotic variance to enhance the power performance. Simulations evaluate the finite-sample performances of the proposed tests, and a real data analysis illustrates its application.

分会场报告（20.3.3）：Detecting communities in attributed networks through bi-direction penalized clustering and its application

报告人：杨虎 中央财经大学

摘 要： How to exploit the heterogeneous information in attributed networks to improve the performance of community detection is a hot research topic. Recently, network embedding methods have been extended to perform community detection on attributed networks, especially state-of-art deep learning methods (such as VGAE-based) translating attributed networks to vectorized data and enabling traditional machine learning methods to deal with attributed networks. It have been proven to be an effective strategy. However, previous methods perform community detection with assumptions about the dimension of embedding and the number of communities. It can limit the effectiveness and applicability of a clustering method if the dimension of embedding or the number of communities is assumed incorrectly, the resulting clustering may be suboptimal or even misleading. Thus, this study fuses VGAE-based methods and a bi-direction penalized clustering algorithm (BiPClust) for community detection and addresses the challenges. To the best of our knowledge, this study is the first to combine these techniques with the following advantages: it solves the issues of dimension selection and the number of communities determination through optimizing penalized clustering automatically; it provides both the computational algorithm and the statistical theorems as a proof to show that BiPClust can counteract the effects of redundant embedding, and determine the unknown number of communities; it is applied for community detection on benchmark datasets and syndicated investment networks in China. Our theoretical analysis suggests that BiPClust can effectively address the challenges of dimension selection and determination of the number of clusters. Furthermore, experimental results demonstrate that BiPClust outperforms other methods, further supporting our theoretical analysis.

分会场报告（20.3.4）：New Tests for High-Dimensional Two-sample Mean Problems with Consideration of Correlation Structure

报告人：杨松山 中国人民大学

摘要： This paper proposes a test statistic for two sample mean testing problems for high dimensional data by assuming the linear structure on high dimensional precision

matrices. A new precision matrix estimation method considering its linear structure is first proposed, and the regularization method is implemented to select the true basis matrices that can further reduce the approximation error. Then the test statistic is constructed by imposing the estimation of the precision matrix. The proposed test is valid for both the low dimensional setting and high dimensional setting even if the dimension of the data is greater than the sample size. The limiting null distributions of the proposed test statistic under both null distribution and alternative distribution are derived. Extensive simulations are conducted for estimating the precision matrix and testing difference of the high dimensional mean vector. Simulation results show that the proposed estimation method enjoy low estimation error for the precision matrix and the regularization method is able to efficiently select the important basis matrix. The testing method performs well compared with the existing methods especially when the elements of the vector have unequal variances. A real data example is then provided to demonstrate the potential of the proposed method in real world applications.

分会场邀请报告（20.1.5）：深度非线性降维

报告人：於州 华东师范大学

摘要：充分降维方法在过去三十年间得到了长足的发展。以切片逆回归、切片平均方差估计为代表的线性充分降维方法计算便捷应用广泛。我们在此将讨论如何基于深度神经网络将线性充分降维推广至非线性充分降维，并探究充分降维中心域估计的无偏性和收敛性以及收敛速度。我们将通过模拟试验和实际数据验证新方法的有效性。

分会场邀请报告（20.1.6）：Regulation-incorporated Network-based Heterogeneity Analysis

报告人：张庆昭 厦门大学

摘要： Gene expression-based heterogeneity analysis has been extensively conducted. In recent studies, it has been shown that network-based analysis, which takes a system perspective and accommodates the interconnections among genes, can be more informative than that based on simpler statistics. Gene expressions are highly regulated. Incorporating regulations in analysis can better delineate the ``sources'' of gene expression effects. Although conditional network analysis can somewhat serve

this purpose, it does render enough attention to the regulation relationships. In this article, significantly advancing from the existing heterogeneity analyses based only on gene expression networks, conditional gene expression network analyses, and regression-based heterogeneity analyses, we propose heterogeneity analysis based on gene expression networks (after accounting for or ``removing'' regulation effects) as well as regulations of gene expressions. A high-dimensional penalized fusion approach is proposed, which can determine the number of sample groups and parameter values (for all sample groups) in a single step. An effective computational algorithm is proposed. It is rigorously proved that the proposed approach enjoys the estimation, selection, and grouping consistency properties under high-dimensional settings. Extensive simulations demonstrate practical superiority of the proposed approach over closely related alternatives. In the analysis of two breast cancer datasets, the proposed approach identifies heterogeneity and gene network structures different from the alternatives and with sound biological implications.

分会场报告（20.1.7）：非线性期望与统计机器学习

报告人：严晓东 山东大学

摘要：山东大学非线性期望团队在风险研究领域已取得了一批原创性的、具有世界一流水平的研究成果，特别是 G-期望理论成为新一代金融风险研究工具。最近，山东大学非线性期望团队团队基于 G-期望理论研究了双臂及多臂机器人策略博弈行为并建立了非线性极限理论，在国际上产生了很大的影响。本报告考虑这一原创理论成果的统计科学问题以及提出新的统计机器学习方法，其中包括（1）双臂老虎机下的最优分布和悖论；（2）针对大数据下的双样本检验；（3）提出序贯 bootstrap 方法进行相似性假设检验和单样本检验的一般框架；（4）最后为了考虑一般化的模型，我们提出了基于随机梯度下降法的参数检验。值得注意的是，新的统计机器学习方法充分结合“知识”属于知识推理与数据驱动的统计学习方法，并且新的方法解决的假设检验问题与传统方法相比获得了更高的势。

分会场报告（20.1.8）：Robust Transfer Learning in High-dimensional GLM via γ-divergence

报告人：徐福芝 厦门大学

摘要：Outlying observations and even data contamination is often occurred in practice

due to high-dimensional sparsity. Robustness against outliers and contamination based on the divergence has been widely adopted. With the rapid growth in the volumes of high-dimensional data, learning from multiple sources of evidence is desired. Transfer learning can improve the performance of target models by transferring information from source datasets. Yet, multiple sources of information, introducing outlying observations and even contamination, may lead to biased estimation and misleading inference. In this study, we propose a robust transfer learning approach based on the minimum γ-divergence under a generalized linear model(GLM) framework for high-dimensional data. Using an algorithm-free transferable source detection scheme, the proposed approach identifies informative sources and avoids negative transfer of learning. We rigorously establish the consistency properties and estimation bounds under high dimensionality. Computational algorithm is developed based on proximal gradient descent for transferring and debiasing steps. Simulation demonstrates superior and competitive performance of the proposed approach in selection and prediction/classification. Analysis of genetic data on breast cancer and glioblastoma confirms its practical usefulness.

分会场邀请报告（20.2.5）：Tensor Model Averaging

报告人：张新雨 中国科学院

摘要： Tensor has a wide range of applications in neuroimaging, data mining, digital marketing, etc. It is difficult to perform simple linear regression directly on tensor data due to its large number of parameters. CANDECOMP/PARAFAC (CP) decomposition can represent a tensor as a sum of several rank-1 tensors, thus effectively reducing the number of parameters. However, we do not know which rank to use for regression in practice, because it is an NP-complete problem to solve the rank of a given tensor. In this article, we develop a model averaging method on CP tensor regression to solve this problem by weighting the estimators from models of different ranks. Simulations and empirical applications show that the proposed method has a promising performance.

分会场邀请报告（20.2.6）：Introducing the Specificity Score: Measuring the Credibility of Causal Associations

报告人：苗旺 北京大学

摘 要： Most applied science is concerned with uncovering causal relationships. Observational studies offer an important source of data for causal inference. A major challenge for observational studies is the unmeasured confounding. A variety of methods have been proposed for adjustment for unmeasured confounding, including instrumental variable and proximal inference. However, these auxiliary variables approaches often rely on complete and exact knowledge about the exclusion of certain causal effects, which is sometimes problematic and the difficulty is amplified when multiple treatments and outcomes are concerned because multiple auxiliary variables are necessitated. In this paper, I establish a framework for specificity analysis of causal effects in observational studies in the presence of unmeasured confounding, including a score measuring the specificity and credibility of causal associations, a test of the null hypothesis of no causal effects with controlled type I error, and an identification and estimation method of causal effects. While only requiring a rough idea on the extent of the causal associations in view, the specificity analyses do not invoke exact knowledge about which causal associations are active or null. The specificity analyses admit joint causal discovery with multiple treatments and multiple outcomes, which is particularly desirable in gene expressions studies, EHR studies, etc. This framework extends previous auxiliary variables approaches and has potential application in Mendelian randomization and selection bias adjustment. Simulations are used for illustration and an application to a mouse obesity dataset detects potential active effects of genes on clinical traits that are relevant to metabolic syndrome.

分会场报告（20.2.7）: Rank-based Greedy Model Averaging for High-dimensional Survival Data

报告人：贺百花 中国科学技术大学

摘要： Model averaging is an effective way to enhance prediction accuracy. However, most previous works focus on parametric models in low-dimensional settings with completely observed responses. To attain an accurate prediction for the risk effect of survival data with highdimensional predictors, we propose a novel method: rank-based nonparametric greedy (RNG) model averaging. Specifically, adopting the

nonparametric transformation model with splitting predictors as working models, we doubly utilize the smooth concordance index function to derive the candidate predictions and optimal model weights. The final prediction is achieved by weighted averaging all the candidates. Our approach is flexible, computationally efficient, and robust against model misspecification, as it neither requires the correctness of a joint model nor involves nonparametric estimation. We further adopt the greedy algorithm for high dimensions. Theoretically, we derive an asymptotic error bound for the optimal weights under some mild conditions. In addition, the summation of weights assigned to the correct candidate submodels is proven to approach one in probability when there are correct models included among the candidate submodels. Extensive numerical studies are carried out using both simulated and real datasets to show the proposed approach's robust performance compared to the existing regularization approaches.

分会场报告（20.2.8）：Time-varying Complete Subset Averaging in a Data-rich Environment

报告人：李海奇 湖南大学

摘要：This study proposes a novel time-varying complete subset averaging (TVCSA) method in a data-rich environment. The dimension of regressors is allowed to diverge with an increase in sample size. The optimal time-varying subset size is obtained by minimizing the local leave-one-out cross-validation (LOOCV) criterion. The proposed TVCSA estimator is consistent and asymptotically follows a normal distribution. Moreover, it is asymptotically optimal in terms of minimizing the local squared error loss and associated risk criterion. In particular, the proposed uniform-weighting TVCSA estimator is shown to be superior to the optimal-weighting averaging estimator. Monte Carlo simulations show favorable evidence for the proposed TVCSA method relative to other popular alternatives. The applications on U.S. equity premium and inflation rate forecasting demonstrate that our method has better forecasting performance than competing alternatives.

分会场邀请报告（20.3.5）：Statistical Inferences for Complex Dependence of Multimodal Imaging Data

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

摘要：Statistical analysis of multimodal imaging data is a challenging task, since the data involves high-dimensionality, strong spatial correlations and complex data structures. In this article, we propose rigorous statistical testing procedures for making inferences on the complex dependence of multimodal imaging data. Motivated by the analysis of multi-task fMRI data in the Human Connectome Project (HCP) study, we particularly address three hypothesis testing problems: (a) testing independence among imaging modalities over brain regions, (b) testing independence between brain regions within imaging modalities, and (c) testing independence between brain regions across different modalities. Considering a general form for all the three tests, we develop a global testing procedure and a multiple testing procedure controlling the false discovery rate. We study theoretical properties of the proposed tests and develop a computationally efficient distributed algorithm. The proposed methods and theory are general and relevant for many statistical problems of testing independence structure among the components of high-dimensional random vectors with arbitrary dependence structures. We also illustrate our proposed methods via extensive simulations and analysis of five task fMRI contrast maps in the HCP study.

分会场邀请报告（20.3.6）: The Aggregation-Heterogeneity Trade-off in Federated Learning

报告人：林伟 北京大学

摘要：Conventional wisdom in machine learning holds that the more data you train your model on, the better the model can perform. Accordingly, a plethora of federated learning methods have been developed to aggregate as many local samples as possible. Contrary to this belief, this paper shows that aggregation of more data is not necessarily beneficial in the presence of heterogeneity, and reveals a fundamental trade-off between aggregation and heterogeneity in federated learning. We consider a general family of weighted M-estimators that interpolate between FedAvg and the local estimator, in which an aggregation rule is determined by the weights of local samples. We derive an upper bound for the estimation error of the weighted M-estimators, which decomposes into a bias term induced by heterogeneity and a

variance term influenced by aggregation. A measure of heterogeneity, the federated smoothness β, is introduced to simplify the general result. As an important consequence, the optimal aggregation rule for each local device is to aggregate only its $\lfloor K^{2\beta/(2\beta+1)}/(n/\sigma^2)^{1/(2\beta+1)}\rfloor\vee 1$ closest neighbors among the K devices, where n is the local sample size and σ^2 is the noise variance. Moreover, we show that our estimator, termed FedKNN, attains the minimax optimal rate over a certain parameter space characterized by β. This optimal procedure depends crucially on the neighboring structure among devices in terms of the proximity of local parameters. Finally, we prove that without such prior knowledge no estimator can achieve a convergence rate faster than $O(\sigma^2/n)$ and hence adaptation is impossible.

分会场报告（20.3.7）: Sequential Data Integration under Dataset Shift

报告人: 盛赢 中国科学院数学与系统科学研究院

摘 要： With the rapidly increasing availability of large-scale and high-velocity streaming data, efficient algorithms that can process data in batches without requiring expensive storage and computation resources have drawn considerable attention. An emerging challenge in developing efficient batch processing techniques is dataset shift, where the joint distribution of the collected data varies across batches. If not recognized and addressed properly, dataset shift often leads to erroneous statistical inferences when integrating data from different batches. In this paper, two shift-adjusted estimation procedures are developed for updated estimation of the parameter in the presence of dataset shift. Under prior probability shift, we can obtain parameter estimation and assess the degree of dataset shift simultaneously. We study the asymptotic properties of the proposed estimators and evaluate their performance in numerical studies. This is a joint work with Jing Qin and Chiung-Yu Huang.

分会场报告（20.3.8）: Federated Survival Analysis via Data Augmentation Using Multi-task Variational Autoencoder

报告人: 王洪 中南大学

摘 要： Survival models finds vast applications in biomedical studies. However, survival data used to train these models are usually distributed, censored and facing a growing concern for data privacy. In addition to these issues, it is less commonly

recognized that survival times are usually tailed. In this study, we attempt to tackle such challenges via a novel federated learning scheme. The proposed scheme aims to mitigate the censoring and tailed data problems via data augmentation using multi-task variational autoencoder(MVAE). Experimental results from extensive simulated and real world survival datasets have demonstrated the effectiveness of the proposed methodology with possible deployments at the server or the clients.

分会场邀请报告（21.1.1）：Feature Screening for Clustering Analysis

报告人：席瑞斌 北京大学

摘要： We consider feature screening for ultrahigh dimensional clustering analyses. Based on the observation that the marginal distribution of any given feature is a mixture of its conditional distributions in different clusters, we propose to screen clustering features by independently evaluating the homogeneity of each feature's mixture distribution. Important clustering-relevant features have heterogeneous components in their mixture distributions and unimportant features have homogeneous components. The well-known EM-test statistic is used to evaluate the homogeneity. Under general parametric settings, we establish the tail probability bounds of the EM-test statistic for the homogeneous and heterogeneous features, and further show that the proposed screening procedure can achieve the sure independent screening and even the consistency in selection properties. Limiting distribution of the EM-test statistic is also obtained for general parametric distributions. The proposed method is computationally efficient, can accurately screen for important clustering-relevant features and help to significantly improve clustering, as demonstrated in our extensive simulation and real data analyses.

分会场报告（21.1.2）：Generalization Ability of Wide Neural Networks

报告人：赖建发 清华大学

摘要： We perform a study on the generalization ability of the wide two-layer ReLU neural network over $\mathbb{R}$. We first establish some spectral properties of the neural tangent kernel (NTK): $a)$ K_{d}, the NTK defined on $\mathbb{R}^{d}$, is positive definite; $b)$ $\lambda_{i}(K_{1})$, the i-th largest eigenvalue of K_{1}, is proportional to i^{-2}. We then show that: $i)$ when the width $m\rightarrow\infty$, the neural network kernel (NNK) uniformly converges to the

NTK; $ii)$ the minimax rate of regression over the RKHS associated to K_{1} is $n^{-2/3}$; $iii)$ if one adopts the early stopping strategy in training a wide neural network, the resulting neural network achieves the minimax rate; $iv)$ if one trains the neural network till overfitted, the resulting neural network can not generalize well. Finally, we provide an explanation to reconcile our theory and the widely observed `benign overfitting phenomenon'.

分会场报告（21.1.3）: Citation Counts Prediction of Statistical Publications Based on Multi-Layer Academic Networks via Neural Network Model

报告人：潘蕊 中央财经大学

摘要： Citation counts is a crucial factor for evaluating the quality of a research paper. Therefore, it is of vital importance to accurately predict citation counts and to explore the mechanism underlying citations. In this paper, we mainly focus on the prediction of citation counts in the field of statistics. We collect 55,024 academic papers published in 43 statistics journals from 2001 to 2018. Our main contributions are as follows. First, we invest considerable effort in collecting and cleaning a high-quality dataset. Second, multi-layer networks are constructed from different perspectives, including journal network, author citation network, co-citation network, co-authorship network, and keyword co-occurrence network. In addition, we extract 77 factors for citation counts prediction, including 22 traditional factors and 55 network-related factors. Third, to address the issues of zero-inflated and over-dispersed citation counts, a neural network model is designed to achieve high prediction accuracy. Furthermore, we adopt a leave-one-feature-out approach to investigate the importance of factors. The results show that our neural network model outperforms the other methods. This study provides a useful guide for researchers to predict citation counts and can be easily extended to other research fields.

分会场报告（21.1.4）: Dimension Reduction for Covariates in Network Data

报告人：赵俊龙 北京师范大学

摘要： A problem of major interest in network data analysis is to explain the strength of connections using context information. To achieve this, we introduce a novel approach, called network supervised dimension reduction, in which covariates are projected onto low-dimensional spaces to reveal the linkage pattern without assuming

a model. We propose a new loss function for estimating the parameters in the resulting linear projection, based on the notion that closer proximity in the low-dimension projection corresponds to stronger connections. Interestingly, the convergence rate of our estimator is found to depend on a network effect factor, which is the smallest number that can partition a graph in a manner similar to the graph colouring problem. Our method has interesting connections to principal component analysis and linear discriminant analysis, which we exploit for clustering and community detection. The proposed approach is further illustrated by numerical experiments and analysis of a pulsar candidates dataset from astronomy.

分会场邀请报告（21.2.1）：Tuning-parameter-free Propensity Score Matching Approach for Causal Inference Under Shape Restriction

报告人：刘玉坤 华东师范大学

摘要： Propensity score matching (PSM) is a pseudo-experimental method that uses statistical techniques to construct an artificial control group by matching each treated unit with one or more untreated units of similar characteristics. To date, the problem of determining the optimal number of matches per unit, which plays an important role in PSM, has not been adequately addressed. We propose a tuning-parameter-free PSM method based on the nonparametric maximum-likelihood estimation of the propensity score under the monotonicity constraint. The estimated propensity score is piecewise constant, and therefore automatically groups data. Hence, our proposal is free of tuning parameters. The proposed estimator is asymptotically semiparametric efficient when the covariate is univariate or the outcome and the propensity score depend on the covariate in the same direction. We conclude that matching methods based on the propensity score alone cannot, in general, be efficient.

分会场报告（21.2.2）：Novel Statistical Methods for High-dimensional Microbiome Data Analysis

报告人：占翔 北京大学

摘要： Advancement in next generation high-throughput sequencing technologies —such as genomics, transcriptomics, proteomics, metabolomics and metagenomi cs—allows characterization of the human omics profile at an extraordinarily det ailed molecular level. Among the fields of omics studies, a very popular mode

of analysis is the association analysis, which tries to establish associative relationships between omics features and disease outcomes as the first step to study the underlying biological omics mechanism of the disease. Despite its popularity, the field of omics association studies, however, has not yet reached enough maturity for making the leap from omics survey to rational omics-based personalized therapeutics. One primary limitation to leverage this large body of omics sequencing data is computational and statistical challenges, including high-dimensionality, sparse data structure, relatively small effect size or sample size and complex dependence/correlation structure among omics features. Taking microbiome and metagenomics data as examples, in this talk, we discuss some recent statistical methods to combat these challenges in microbiome association analysis. Our proposed methods are both powerful and robust, while maintaining both statistical rigor and biological relevance. Using comprehensive numerical simulation studies, we will show that the proposed methods are superior than existing counterparts in literature. We will also demonstrate the potential usefulness of our methods by applications to several real data sets.

分会场报告（21.2.3）: High-dimensional Response Growth Curve Modeling for Longitudinal Neuroimaging Analysis

报告人: 王璐 中南大学

摘 要： There is increasing interest in modeling high-dimensional longitudinal outcomes in applications such as developmental neuroimaging research. Growth curve model offers a useful tool to capture both dynamic changes of outcomes over time within each individual, as well as the mean growth pattern across individuals. However, when the number of outcomes is large, it becomes challenging and often infeasible to tackle the large covariance matrix of the random effects involved in the model. In this article, we propose a high-dimensional response growth curve model, with three novel components: a low-rank factor model structure that substantially reduces the number of parameters in the large covariance matrix, a re-parameterization formulation coupled with a sparsity penalty that selects important fixed and random effect terms, and a computational trick that turns the inversion of a large matrix into the inversion of a stack of small matrices and thus considerably speeds up the computation. We develop an efficient expectation-maximization type estimation algorithm, and demonstrate the competitive performance of the proposed method through both simulations and a longitudinal study of brain structural

connectivity in association with human immunodeficiency virus.

分会场报告（21.2.4）: Generalized Liquid Association Analysis for Multimodal Data Integration

报告人：曾靖 中国科学技术大学

摘要： Multimodal data are now prevailing in scientific research. One of the central questions in multimodal integrative analysis is to understand how two data modalities associate and interact with each other given another modality or demographic variables. The problem can be formulated as studying the associations among three sets of random variables, a question that has received relatively less attention in the literature. In this article, we propose a novel generalized liquid association analysis method, which offers a new and unique angle to this important class of problems of studying three-way associations. We extend the notion of liquid association from the univariate setting to the sparse, multivariate, and high-dimensional setting. We establish a population dimension reduction model, transform the problem to sparse Tucker decomposition of a three-way tensor, and develop a higher-order orthogonal iteration algorithm for parameter estimation. We derive the nonasymptotic error bound and asymptotic consistency of the proposed estimator, while allowing the variable dimensions to be larger than and diverge with the sample size. We demonstrate the efficacy of the method through both simulations and a multimodal neuroimaging application for Alzheimer's disease research. Supplementary materials for this article are available online.

分会场邀请报告（21.3.1）: Group-orthogonal Subsampling for Big Data Linear Mixed Models

报告人：孙法省 东北师范大学

摘要： Linear mixed model is a popular and common modeling method in statistical analysis. It is computationally difficult to obtain parameter estimates in linear mixed model for big data. The current subsampling methods are mainly aimed at the situation where the data is independent, without considering the correlation within the data. We provide some theoretical results on information matrix for linear mixed model. Based on these findings, an optimal subsampling method for linear mixed model is proposed, which maximizes the determinant of the variance-covariance

matrix of the subsampling estimator. Besides, the proposed subsampling procedure is also optimal under A-optimality criterion, which minimizes the trace of the variance-covariance matrix of the subsampling estimator. Furthermore, asymptotic property of the subsampling estimator is established. Numerical examples based on both simulated and real data are provided to illustrate the proposed subsampling method.

分会场报告（21.3.2）: Debiased and Thresholded Ridge Regression for Linear Models with Heteroskedastic and Correlated Errors

报告人：张云翼 香港中文大学（深圳）

摘 要： High-dimensional linear models with independent errors have been well-studied. However, statistical inference on a high-dimensional linear model with heteroskedastic, dependent (and possibly non-stationary) errors is still a novel topic. Under such complex assumptions, the paper at hand introduces a debiased and thresholded ridge regression estimator that is consistent, and is able to recover the model sparsity. Moreover, we derive a Gaussian approximation theorem for the estimator, and apply a dependent wild bootstrap algorithm to construct simultaneous confidence interval and hypothesis tests for linear combinations of parameters. Numerical experiments with both real and simulated data show that the proposed estimator has good finite sample performance. Of independent interest is the development of a new class of heteroscedastic, (weakly) dependent, and non-stationary random variables that can be used as a general model for regression errors.

分会场报告（21.3.3）: Sketched Ridgeless Linear Regression: The Statistical Role of Sketching

报告人：曾奕程 深圳市大数据研究院

摘要: Overparametrization often helps improve the generalization performance. We propose a dual view of overparametrization suggesting that sketching, as a popular technique of reducing effective sample size, may also help generalize. Motivated by this dual view, we characterize two out-of-sample prediction risks of the sketched ridgeless least square estimator in the proportional regime with comparable m, n and p, where m is the sketching size, n the sample size, and p the feature dimensionality. Our

results reveal the statistical role of sketching. Specifically, sketching does not always hurt the generalization performance, and may actually help improve it in some cases. We identify the optimal sketching sizes that minimize the out-of-sample prediction risks, and find that the optimally sketched estimator has stabler risk curves that eliminates the peaks of those for the full-sample estimator. We then propose a practical procedure to empirically identify the optimal sketching size. Numerical studies strongly support our theory.

分会场报告（21.3.4）：Classified Generalized Linear Mixed Model Prediction Incorporating Pseudo-prior Information

报告人：马海强 江西财经大学

摘 要： We develop a method of classified mixed model prediction based on generalized linear mixed models that incorporate pseudo-prior information to improve prediction accuracy. We establish consistency of the proposed method both in terms of prediction of the true mixed effect of interest and in terms of correctly identifying the potential class corresponding to the new observations if such a class matching one of the training data classes exists. Empirical results, including simulation studies and real-data validation, fully support the theoretical findings.

分会场邀请报告（21.1.5）：广义离群随机矩阵的极限理论、统计推断和交叉应用

报告人：姜丹丹 西安交通大学

摘要：从构建基础极限理论、创新统计推断方法、拓展交叉领域应用三方面深入地研究了广义离群随机矩阵。首先，在变量维数和样本容量成比例趋于无穷的条件下，推导了广义离群随机矩阵特征值的一阶和二阶极限理论，即特征值的相变和谱统计量的中心极限定理。其次，基于上述极限理论提出离群特征值的一致估计、噪声方差的修正估计、离群特征值个数检验、以及广义离群模型中部分最小根的相等检验等统计推断新方法。最后，无线通信、电力等领域的数据分析中大多会涉及离群随机矩阵模型，如信道矩阵、电力系统同步相量测量数据矩阵等对应的协方差矩阵。因此我们将统计推断新方法应用于这些领域，解决大规模多输入多输出系统中信道容量的估计问题、大规模电力系统的故障快速检测问题等。实验研究表明，所提出的极限理论和统计推断方法放宽了之前一些对角化或分块对角化假设的严格条件，并在不假设正态性总体的情况下将工作扩展到更广泛的

范围；在交叉领域的应用中所提出的方法与现有方法相比，为无线通信、电力等领域的一些实际问题分析提供了更低耗的统计方法和计算方法，从而达到在实际领域中针对海量数据进行快速分析的目的。

分会场报告（21.1.6）：Weighted Residual Empirical Processes, Martingale Transformations, and Model Checking for Regressions

报告人：谭发龙 湖南大学

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 then study the convergence of weighted residual empirical processes and their martingale transformation under the null and alternative hypotheses in the diverging dimension setting. 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 the 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 asymptoticallydistribution-freetestsbasedonmartingaletransformations.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.

分会场报告（21.1.7）：Regularized t Distribution: Definition, Properties and Application

报告人：胡宗良 深圳大学

摘要：For gene expression data analysis, an important task is to identify genes that are differentially expressed between two or more groups. Nevertheless, as biological experiments are often measured with a relatively small number of samples, how to accurately estimate the variances of gene expression becomes a challenging issue. To tackle this problem, we introduce a regularized t distribution and derive its statistical

properties including the probability density function and the moment generating function. The noncentral regularized t distribution is also introduced for computing the statistical power of hypothesis testing. For practical applications, we apply the regularized t distribution to establish the null distribution of the regularized t statistic, and then formulate it as a regularized t-test for detecting the differentially expressed genes. Simulation studies and real data analysis show that our regularized t-test performs much better than the Bayesian t-test in the "limma" package, in particular when the sample sizes are small.

分会场报告（21.1.8）: Randomization-based Joint Central Limit Theorem and Efficient Covariate Adjustment in Randomized Block 2K Factorial Experiments

报告人: 杨玥含 中央财经大学

摘 要： Randomized block factorial experiments are widely used in industrial engineering, clinical trials, and social science. Researchers often use a linear model and analysis of covariance to analyze experimental results; however, limited studies have addressed the validity and robustness of the resulting inferences because assumptions for a linear model might not be justified by randomization in randomized block factorial experiments. In this article, we establish a new finite population joint central limit theorem for usual (unadjusted) factorial effect estimators in randomized block 2K factorial experiments. Our theorem is obtained under a randomization-based inference framework, making use of an extension of the vector form of the Wald–Wolfowitz–Hoeffding theorem for a linear rank statistic. It is robust to model misspecification, numbers of blocks, block sizes, and propensity scores across blocks. To improve the estimation and inference efficiency, we propose four covariate adjustment methods. We show that under mild conditions, the resulting covariate-adjusted factorial effect estimators are consistent, jointly asymptotically normal, and generally more efficient than the unadjusted estimator. In addition, we propose Neyman-type conservative estimators for the asymptotic covariances to facilitate valid inferences. Simulation studies and a clinical trial data analysis demonstrate the benefits of the covariate adjustment methods. Supplementary materials for this article are available online.

分会场邀请报告（21.2.5）：One-way or Two-way Factor Model for Matrix Sequences?

报告人：孔新兵 南京审计大学

摘要： In this talk, we investigate the issue of determining the dimensions of row and column factor spaces in matrix-valued data. Exploiting the eigen-gap in the spectrum of sample second moment matrices of the data, we propose a family of randomised tests to check whether a one-way or two-way factor structure exists or not. Our tests do not require any arbitrary thresholding on the eigenvalues, and can be applied with (virtually) no restrictions on the relative rate of divergence of the cross-sections to the sample sizes as they pass to infinity. Although tests are based on a randomization which does not vanish asymptotically, we propose a de-randomized, strong (based on the Law of the Iterated Logarithm) decision rule to choose in favour or against the presence of common factors. We use the proposed tests and decision rule in two ways. We further cast our individual tests in a sequential procedure whose output is an estimate of the number of common factors. Our tests are built on two variants of the sample second moment matrix of the data: one based on a row (or column) flattened version of the matrix-valued sequence, and one based on a projection-based method. Our simulations show that both procedures work well in large samples and, in small samples, the one based on the projection method delivers a superior performance compared to existing methods in virtually all cases considered.

分会场报告（21.2.6）：基于结构化随机森林的地震波初至拾取方法

报告人：张春霞 西安交通大学

摘要：初至拾取是地球物理勘探中的重要环节，但随着野外采集技术的发展和进步，尤其是面临十分复杂的地表条件时，获得满意的拾取结果仍是一个挑战。随着人工智能技术的迅速发展，实现地震波初至拾取的智能化是 AI+物探技术领域的一个研究热点。本报告中，我们将介绍基于结构化随机森林的初至拾取方法，该方法考虑了相邻道集数据之间的相关性，并设计了合适的数据正则化技术和拾取结果的后期处理策略，在模拟数据和实际工区数据上的试验结果表明，新方法可达到与人工拾取结果基本相当的自动拾取，从而为智能化初至拾取提供了一种新途径。

分会场报告（21.2.7）：Functional Additive Expectile Regression in the Reproducing Kernel Hilbert Space

报告人：刘育孜 江西财经大学

摘要： In the literature, the functional additive regression model has received much attention. Most studies currently conducted, however, only estimate the mean function, which may not adequately capture the heteroscedasticity and/or asymmetries of the model errors. In light of this, we extend functional additive regression models to their expectile counterparts and obtain an upper bound on the convergence rate of its regularized estimator under mild conditions. To demonstrate its finite sample performance, a few simulation experiments, and a real data example are provided.

分会场报告（21.2.8）：A General Framework to Select Tuning Parameter for Nonparametric Derivative Estimation

报告人：刘思晟 湖南师范大学

摘要： Derivative estimation plays an important role in many practical applications. In this paper, a general framework is proposed to select tuning parameter for the nonparametric derivative estimation. The new framework enlarges the scope of the generalized Cp criterion Charnigo et al. (2011) by replacing the empirical derivative via any other linear nonparametric smoother. The framework is flexible since it can work for any nonparametric derivative estimation that are linear in observed responses. We provide the theoretical support of the proposed tuning parameter selection method and justify it through empirical studies. The simulation results showed that our proposed framework works well for different nonparametric derivative estimation methods. The practical application of the proposed framework is demonstrated in the study of the age effect on hippocampal gray matter volume in healthy adults from the IXI dataset.

分会场邀请报告（21.3.5）：Heterogeneity-aware Clustered Distributed Learning for Multi-source Data Analysis

报告人：方匡南 厦门大学

摘要： In diverse fields ranging from finance to omics, it is increasingly common that

data is distributed and with multiple individual sources (referred to as “clients” in some studies). Integrating raw data, although powerful, is often not feasible, for example, when there are considerations on privacy protection. Distributed learning techniques have been developed to integrate summary statistics as opposed to raw data. In many of the existing distributed learning studies, it is stringently assumed that all the clients have the same model. To accommodate data heterogeneity, some federated learning methods allow for client-specific models. In this article, we consider the scenario that clients form clusters, those in the same cluster have the same model, and different clusters have different models. Further considering the clustering structure can lead to a better understanding of the “interconnections” among clients and reduce the number of parameters. To this end, we develop a novel penalization approach. Specifically, group penalization is imposed for regularized estimation and selection of important variables, and fusion penalization is imposed to automatically cluster clients. An effective ADMM algorithm is developed, and the estimation, selection, and clustering consistency properties are established under mild conditions. Simulation and data analysis further demonstrate the practical utility and superiority of the proposed approach.

分会场报告（20.3.6）：Distribution-free Simultaneous Prediction Bands for Clustered Data with Missing Responses

报告人：唐炎林 华东师范大学

摘要： Prediction plays a crucial role in decision-making across various fields and industries, as it enables us to anticipate future events or outcomes based on observed data and knowledge. However, existing methods often rely on strong modeling assumptions and thus are prone to model misspecification errors. In this work, we construct simultaneous prediction for missing clustered data based on the idea of conformal inference, without any assumptions about the model distribution and within-group dependency structures. These covariate-dependent predictions assess the overall pattern and magnitude of global associations between the responses and covariates, and provide coverage guarantees in finite samples. In particular, compared to alternative methods, our methods yield the smallest prediction regions by converging to the highest density set, and fully adapts to complex error distributions by approximating conditional coverage. Simulations show the excellent finite-sample behavior of our methods in comparison to naive alternatives. The practical use of our methods is demonstrated in two case studies on serum cholesterol and CD4+ cells.

分会场报告（21.3.7）：Renewable Quantile Regression for Streaming Data Sets

报告人：姜荣 上海第二工业大学

摘要： Online updating is an important statistical method for the analysis of big data arriving in streams due to its ability to break the storage barrier and the computational barrier under certain circumstances. The quantile regression, as a widely used regression model in many fields, faces challenges in model fitting and variable selection with big data arriving in streams. Chen et al. (2019, Annals of Statistics) has proposed a quantile regression method for streaming data, but a strong additional condition is required. In this paper, renewable optimized objective functions for regression parameter estimation and variable selection in a quantile regression are proposed. The proposed methods are illustrated using current data and the summary statistics of historical data. Theoretically, the proposed statistics are shown to have the same asymptotic distributions as the standard version computed on an entire data stream with the data batches pooled into one data set, without additional condition. Both simulations and data analysis are conducted to illustrate the finite sample performance of the proposed methods.

分会场报告（20.3.8）：Smooth-threshold GMM Estimation of Higher-order Semiparametric Spatial Autoregressive Model with Increasing Dimension

报告人：杨晶 湖南师范大学

摘要： The generalized method of moments (GMM) is a particularly popular method that has been widely used in econometric modeling, in terms of its computation simplicity and estimation efficiency. In this paper, to capture the spatial dependence as well as heterogeneous effects of the impacts of some regressors, we concentrate on the GMM estimation and variable selection of a higher-order spatial autoregressive model with partially linear varying coefficients and diverging number of parameters. Specifically, with the varying coefficient functions being approximated by some basis functions, the GMM estimation of the considered model is firstly studied and then a novel and convenient smooth-threshold GMM procedure is constructed for variable selection. Under some regularity conditions, we establish the asymptotic normality and oracle property of the resulting estimators. This is not only the first attempting at studying the model based on the series GMM estimation, but also the first considering variable selection based on the smooth-threshold estimating equations. Extensive

numerical simulations are conducted to confirm the theories and to demonstrate the finite sample performance of the proposed method. A real data analysis is followed for application.

通讯录（按工作单位首字母顺序）

排序	姓名	工作单位	电子邮箱
1	李凡群	安徽财经大学	lfq006@163.com
2	孙军	安徽财经大学	sunj@aufe.edu.cn
3	张孔生	安徽财经大学	zks155@163.com
4	赵明涛	安徽财经大学	zhmingtao@outlook.com
5	方红燕	安徽大学	fanhoy@ustc.edu.cn
6	李良雪	安徽大学	1972567984@qq.com
7	李振龙	安徽大学	1244980910@qq.com
8	毛军军	安徽大学	maojunjun@ahu.edu.cn
9	苏元龙	安徽大学	2943502583@qq.com
10	孙孝东	安徽大学	644451601@qq.com
11	王学军	安徽大学	wxjahdx2000@126.com
12	吴捷	安徽大学	wu12jie@mail.ustc.edu.cn
13	吴涛	安徽大学	wutao@ahu.edu.cn
14	徐东横	安徽大学	2335757628@qq.com
15	徐猛猛	安徽大学	1607533850@qq.com
16	杨小乐	安徽大学	yangxl0305@126.com
17	张正龙	安徽大学	a22201055@stu.ahu.edu.cn
18	朱汉兵	安徽大学	zhuhbecnu@163.com
19	曹明响	安徽师范大学	caomingx@163.com
20	何道江	安徽师范大学	djheahnu@163.com
21	贺磊	安徽师范大学	lhstat@163.com
22	胡学平	安庆师范大学	hxprob@163.com
23	马耀兰	北方民族大学	ylma@nmu.edu.cn
24	艾明要	北京大学	myai@pku.edu.cn
25	陈松蹊	北京大学	csx@gsm.pku.edu.cn
26	林伟	北京大学	weilin@math.pku.edu.cn
27	苗旺	北京大学	mwfy@pku.edu.cn
28	王惠远	北京大学	huiyuan.wang@pku.edu.cn
29	席瑞斌	北京大学	ruibinxi@math.pku.edu.cn
30	占翔	北京大学	zhanxiang@bicmr.pku.edu.cn
31	王学丽	北京工商大学	xlwang@btbu.edu.cn
32	边江园	北京工业大学	809604702@qq.com
33	陈巍璐	北京工业大学	chenweilu@bjut.edu.cn
34	程维虎	北京工业大学	chengweihu@bjut.edu.cn
35	邓欣依	北京工业大学	xinyi@bjut.edu.cn
36	葛娇	北京工业大学	gejiao@emails.bjut.edu.cn

37	蒋媛媛	北京工业大学	jiangyy@bjut.edu.cn
38	闵慧	北京工业大学	huimin@bjut.edu.cn
39	田镇滔	北京工业大学	zttien@emails.bjut.edu.cn
40	谢琍	北京工业大学	xieli@bjut.edu.cn
41	张忠占	北京工业大学	zzhang@bjut.edu.cn
42	曾辉	北京化工大学	2021700057@buct.edu.cn
43	杜斌	北京师范大学	201931011012@mail.bnu.edu.cn
44	郭旭	北京师范大学	xustat12@bnu.edu.cn
45	贾吉航	北京师范大学	202221011011@mail.bnu.edu.cn
46	王宁	北京师范大学	ningwangbnu@bnu.edu.cn
47	王中含	北京师范大学	202221011016@mail.bnu.edu.cn
48	赵俊龙	北京师范大学	zhaojunlong928@126.com
49	郑胜彬	北京师范大学	a1261110123@163.com
50	郑泽靖	北京师范大学	201911011160@mail.bnu.edu.cn
51	杜心灵	北京师范大学珠海分校	du_xinling2008@163.com
52	王潇逸	北京师范大学珠海校区	wangxy059@nenu.edu.cn
53	苑飞	北京师范大学珠海分校	Yuanfei@bnuz.edu.cn
54	周妮文	北京师范大学珠海校区	nwzhou@bnu.edu.cn
55	苏苗苗	北京邮电大学	smm@bupt.edu.cn
56	Tony Cai	宾夕法尼亚大学	tcai@wharton.upenn.edu
57	许婧榆	宾夕法尼亚州立大学	jfx5090@psu.edu
58	李应求	长沙理工大学	liyq-2001@163.com
59	赵乃非	长沙理工大学	zhaonf@csust.edu.cn
60	朱恩文	长沙理工大学	zhuenwen@csust.edu.cn
61	王巍	池州学院	573198607@qq.com
62	吴燚	池州学院	wuyi8702@163.com
63	夏慧异	池州学院	xiayz_88@163.com
64	邓新	滁州学院	Tzdx0120@163.com
65	丁洋	滁州学院	676762218@qq.com
66	陈银钧	重庆大学	Chenyinjun295@163.com
67	李晓妍	重庆大学	xyanli2016@163.com
68	明浩	重庆大学	mh@cqu.edu.cn
69	谭作洵	重庆大学	tzx@cqu.edu.cn
70	吴小飞	重庆大学	xfwu1016@163.com
71	夏小超	重庆大学	xxc@cqu.edu.cn
72	徐建文	重庆大学	xjw@cqu.edu.cn
73	尹燕青	重庆大学	yinyq@cqu.edu.cn
74	运瑶瑶	重庆大学	202206021026@stu.cqu.edu.cn

75	胡雪梅	重庆工商大学	huxuem@163.com
76	王斯	重庆工商大学	Wangsi_ctbu@163.com
77	郭建华	东北师范大学	jhguo@nenu.edu.cn
78	李江豪	东北师范大学	lijh458@nenu.edu.cn
79	李伟荣	东北师范大学	2373543276@qq.com
80	孙法省	东北师范大学	sunfs359@nenu.edu.cn
81	魏雅诗	东北师范大学	weiys958@nenu.edu.cn
82	武前永	东北师范大学	wuqy923@nenu.edu.cn
83	杨丽	东北师范大学	yangl245@nenu.edu.cn
84	郝美玲	对外经济贸易大学	meilinghao@uibe.edu.cn
85	熊巍	对外经济贸易大学	xiongwei@uibe.edu.cn
86	陈瑞娟	福建农林大学	2254810547@qq.com
87	赖雨婷	福建农林大学	2607376498@qq.com
88	黎德元	复旦大学	deyuanli@fudan.edu.cn
89	唐一苇	复旦大学	ywtang22@m.fudan.edu.cn
90	朱仲义	复旦大学	zhuzy@fudan.edu.cn
91	李乃医	广东海洋大学	Linaiyi1979@163.com
92	汪国新	广西大学	wgx183@163.com
93	张涛	广西科技大学	taozhangmaths@126.com
94	周彦秋	广西科技大学	zhouyanqiunihao@163.com
95	张军舰	广西师范大学	jjzhang@gxnu.edu.cn
96	张崇岐	广州大学	cqzhang@gzhu.edu.cn
97	张兴发	广州大学	xingfazhang@gzhu.edu.cn
98	邓靖涵	贵州大学	ddd1422237756@icloud.com
99	胡怀青	贵州大学	2276291017@qq.com
100	李发成	贵州大学	1060276975@qq.com
101	刘惠篮	贵州大学	liuhuilan69@163.com
102	马俊杰	贵州大学	majunjie1029@126.com
103	王飞云	贵州大学	wangfeiyun98@163.com
104	昝思吕	贵州大学	1264624752@qq.com
105	张梨	贵州大学	169038080@qq.com
106	黄海午	桂林航天工业学院	23470202@qq.com
107	冯峥晖	哈尔滨工业大学（深圳）	fengzhenghui@hit.edu.cn
108	黄倩	合肥工业大学	hq@hfut.edu.cn
109	姚梅	合肥工业大学	ymnancy@sina.com
110	李宁	合肥学院	lining@hfuu.edu.cn
111	牛勇	合肥学院	niuyong@hfuu.edu.cn
112	陈花萍	河南大学	chenhp0107@henu.edu.cn

113	解俊山	河南大学	junshan@henu.edu.cn
114	辛欣	河南大学	xinxin@henu.edu.cn
115	薛留根	河南大学	lgxue@bjut.edu.cn
116	郑晨	河南大学	hedazc@126.com
117	刘中强	河南理工大学	zhongqiang@hpu.edu.cn
118	马学思	河南理工大学	maxuesi@hpu.edu.cn
119	王照良	河南理工大学	wangzhaoliang@hpu.edu.cn
120	张素婷	河南理工大学	2370430659@qq.com
121	张旭阳	河南理工大学	xuyangzhang0927@163.com
122	张耀峰	湖北经济学院	53718953@qq.com
123	黄新	湖南城市学院	729314661@ qq.com
124	李海奇	湖南大学	lihaiqi00@hnu.edu.cn
125	谭发龙	湖南大学	falongtan@hnu.edu.cn
126	程慧慧	华北水利水电大学	chenghuihui@ncwu.edu.cn
127	马悦	华北水利水电大学	mayue@ncwu.edu.cn
128	庄智涛	华北水利水电大学	zhuangzhitao@ncwu.edu.cn
129	蔡舒雅	华东师范大学	sycai@xb.ecnu.edu.cn
130	刘玉坤	华东师范大学	ykliu@sfs.ecnu.edu.cn
131	汤银才	华东师范大学	yctang@stat.ecnu.edu.cn
132	唐炎林	华东师范大学	yltang@fem.ecnu.edu.cn
133	王光辉	华东师范大学	ghwang.nk@gmail.com
134	王小舟	华东师范大学	xzwang@sfs.ecnu.edu.cn
135	於州	华东师范大学	zyu@stat.ecnu.edu.cn
136	赵伟	华东师范大学	wzhao@xb.ecnu.edu.cn
137	唐风琴	淮北师范大学	tfq05@163.com
138	王涛	淮阴师范学院	skywangtao2007@126.com
139	王露	黄淮学院	Lwang008@163.com
140	郭翔宇	吉林大学	970738215@qq.com
141	徐悦	吉林大学	972388216@qq.com
142	张佳玥	吉林大学	2996143363@qq.com
143	赵世舜	吉林大学	zhaoss@jlu.edu.cn
144	朱复康	吉林大学	zfk8010@163.com
145	陈望学	吉首大学	chenwangxue2015@163.com
146	邹航	暨南大学	1357103682@qq.com
147	陈慧	江苏师范大学	chenhuiah@163.com
148	李传权	江西财经大学	lichuanquan@jxufe.edu.cn
149	刘庆	江西财经大学	liuqing@jxufe.edu.cn
150	刘小惠	江西财经大学	liuxiaohui@jxufe.edu.cn

151	刘艺媛	江西财经大学	yliuga@connect.ust.hk
152	刘育孜	江西财经大学	liuyuzi@jxufe.edu.cn
153	马海强	江西财经大学	mhqtc118@163.com
154	彭凌	江西财经大学	pengling@jxufe.edu.cn
155	谭祥勇	江西财经大学	xiangyongtan@126.com
156	王平平	江西财经大学	1200400842@jxufe.edu.cn
157	徐慧植	江西财经大学	xuhuizhi@126.com
158	蔡奥丽	江西师范大学	2910893107@qq.com
159	曹汝逸	江西师范大学	3151814498@qq.com
160	陈国武	江西师范大学	66933180@qq.com
161	黄根帆	江西师范大学	792351857@qq.com
162	刘蔚	江西师范大学	2508859207@qq.com
163	王琛琛	江西师范大学	1349991328@qq.com
164	王兆辉	江西师范大学	136894181@qq.com
165	温利民	江西师范大学	wlmjxnu@163.com
166	余鑫	江西师范大学	282477071@qq.com
167	曾剑锋	江西师范大学	jf_zeng@126.com
168	张奉洋	江西师范大学	2356707587@qq.com
169	朱鸿祥	江西师范大学	2503246350@qq.com
170	黄玉婷	兰州财经大学	502613186@qq.com
171	马蓉	兰州财经大学	marong@lzufe.edu.cn
172	牛成英	兰州财经大学	ncy2007815@163.com
173	司颖华	兰州财经大学	siyinghua1789@126.com
174	田艳丽	兰州财经大学	epii02@126.com
175	肖强	兰州财经大学	xiaoqiang1789@126.com
176	赵煜	兰州财经大学	zhaoyulzcc@163.com
177	倪小禾	兰州大学	540494934@qq.com
178	刘群	闽南师范大学	604011685@qq.com
179	刘宣	南昌师范学院	liuxuanyg@163.com
180	蒋学军	南方科技大学	jiangxj@sustech.edu.cn
181	田国梁	南方科技大学	tiangl@sustech.edu.cn
182	陈鹏	南京航空航天大学	chenpengmath@nuaa.edu.cn
183	余迁	南京航空航天大学	qyumath@nuaa.edu.cn
184	温阳俊	南京农业大学	wenyangjun@njau.edu.cn
185	张瑾	南京农业大学	zhangjin@njau.edu.cn
186	蔡雄	南京审计大学	caix2016@163.com
187	孔新兵	南京审计大学	xinbingkong@126.com
188	岳莉莉	南京审计大学	yuelilicz@163.com

189	曹春正	南京信息工程大学	caochunzheng@nuist.edu.cn
190	曹志杰	南京信息工程大学	Czj112463@163.com
191	陈文珠	南京信息工程大学	202212150025@nuist.edu.cn
192	杜莹莹	南京信息工程大学	820172084@qq.com
193	付春蕊	南京信息工程大学	fcr990112@163.com
194	化理想	南京信息工程大学	hlx000425@163.com
195	来鹏	南京信息工程大学	laipeng@nuist.edu.cn
196	宋凤丽	南京信息工程大学	songfl@nuist.edu.cn
197	田汶鑫	南京信息工程大学	tianwx2905@163.com
198	汪子同	南京信息工程大学	2115008755@qq.com
199	王博慧	南京信息工程大学	2625633837@qq.com
200	王黎明	南京信息工程大学	wanglimingnuist@163.com
201	吴香华	南京信息工程大学	wuxianghua@nuist.edu.cn
202	张婷	南京信息工程大学	zhangtstat@nuist.edu.cn
203	张泽文	南京信息工程大学	1971268991@qq.com
204	周胜	南京信息工程大学	7782092209@qq.com
205	陈大川	南开大学	dchen@nankai.edu.cn
206	邓路	南开大学	denglu014@mail.nankai.edu.cn
207	梁万丰	南开大学	983341161@qq.com
208	王磊	南开大学	lwangstat@nankai.edu.cn
209	邹长亮	南开大学	nk.chlzou@gmail.com
210	许腾腾	南通大学	txu@ntu.edu.cn
211	赵为华	南通大学	zhaowhstat@163.com
212	王延新	宁波工程学院	yxwang@nbut.edu.cn
213	沈亮	青岛理工大学	shenliang@qut.edu.cn
214	陈占寿	青海师范大学	chenzhanshou@126.com
215	娘毛措	青海师范大学	2199459589@qq.com
216	马健琦	青海师范大学	670797504@qq.com
217	陈锐	清华大学	chenrui@mail.tsinghua.edu.cn
218	赖建发	清华大学	jianfalai@mail.tsinghua.edu.cn
219	卢伟灏	清华大学	luwh19@mails.tsinghua.edu.cn
220	卢鑫	清华大学	lux20@mails.tsinghua.edu.cn
221	田松涛	清华大学	tst20@mails.tsinghua.edu.cn
222	徐曼芸	清华大学	xumy20@mails.tsinghua.edu.cn
223	高振龙	曲阜师范大学	gzlkygz@163.com
224	王少新	曲阜师范大学	shxwang@qfnu.edu.cn
225	初一	山东财经大学	20198527@sdufe.edu.cn
226	林路	山东大学	linlu@sdu.edu.cn

227	严晓东	山东大学	yanxiaodong@sdu.edu.cn
228	胡琴琴	山东大学（威海）	qqhu@sdu.edu.cn
229	李华鹏	山西大同大学	lihuaa1_0@163.com
230	张秀珍	山西大同大学	zhangxiuzhen132@163.com
231	罗瑞苗	山西大学	luorm955@sxu.edu.cn
232	史建红	山西师范大学	shijh70@163.com
233	肖菊霞	山西师范大学	xiaojuxia2426@126.com
234	余平	山西师范大学	ypxye@sina.com
235	张耀武	上海财经大学	zhang.yaowu@mail.shufe.edu.cn
236	姜荣	上海第二工业大学	jiangrong@sspu.edu.cn
237	柴海涛	上海对外经贸大学	cht0816@163.com
238	董平	上海对外经贸大学	dongping35@126.com
239	范彩云	上海对外经贸大学	fancaiyun2010@163.com
240	范燕	上海对外经贸大学	fanyan212@126.com
241	方艳	上海对外经贸大学	yiffanyfang@163.com
242	柯蓉	上海对外经贸大学	kerong1996@126.com
243	李莉	上海对外经贸大学	lilycb163@163.com
244	李睿	上海对外经贸大学	lirui@suibe.edu.cn
245	刘关福	上海对外经贸大学	liuguanfu07@163.com
246	张建新	上海对外经贸大学	jxzhang@suibe.edu.cn
247	张日权	上海对外经贸大学	zhangriquan@163.com
248	赵霞	上海对外经贸大学	zhaoxia-w@163.com
249	邹欣悦	上海对外经贸大学	xinyuezou@suibe.edu.cn
250	孙章爽	上海工程技术大学	18284513677@163.com
251	姚若愚	上海工程技术大学	yaoyry@163.com
252	周雨婷	上海工程技术大学	edith725@163.com
253	范国良	上海海事大学	hfutfgl@163.com
254	邹玉叶	上海海事大学	zouyuye@shmtu.edu.cn
255	刘卫东	上海交通大学	weidongl@sjtu.edu.cn
256	马诗洋	上海交通大学	mashiyang1991@sjtu.edu.cn
257	毛晓军	上海交通大学	maoxj@sjtu.edu.cn
258	牟荣吉	上海交通大学	rjmu@sjtu.edu.cn
259	任好洁	上海交通大学	haojieren@sjtu.edu.cn
260	王成	上海交通大学	chengwang@sjtu.edu.cn
261	王涛	上海交通大学	neowangtao@sjtu.edu.cn
262	吴瑞佳	上海交通大学	rjwu@sjtu.edu.cn
263	严嘉伟	上海交通大学	yjw1998@sjtu.edu.cn
264	俞章盛	上海交通大学	yuzhangsheng@sjtu.edu.cn

265	郭燕	上海师范大学	guoyan_shnu@163.com
266	彭丽洁	上海师范大学	penglij@163.com
267	吴鑑洪	上海师范大学	wujianhong@shnu.edu.cn
268	许佩蓉	上海师范大学	prxu@shnu.edu.cn
269	岳荣先	上海师范大学	yue2@shnu.edu.cn
270	周睿超	上海师范大学	ruichaozhou@126.com
271	胡宗良	深圳大学	zlhu@szu.edu.cn
272	林炳清	深圳大学	bqlin@szu.edu.cn
273	周彦	深圳大学	zhouy1016@szu.edu.cn
274	曾奕程	深圳市大数据研究院	statzyc@sribd.cn
275	蔡婷婷	首都师范大学	2220501029@cnu.edu.cn
276	郭文雯	首都师范大学	wwguo@cnu.edu.cn
277	金力	首都师范大学	jinl@cnu.edu.cn
278	唐豪	首都师范大学	13269112155@139.com
279	邹国华	首都师范大学	ghzou@amss.ac.cn
280	李体政	西安建筑科技大学	tizhengli@xauat.edu.cn
281	姜丹丹	西安交通大学	jiangdd@xjtu.edu.cn
282	李星祥	西安交通大学	lxxwlm2013@stu.xjtu.edu.cn
283	喻达磊	西安交通大学	yudalei@126.com
284	张春霞	西安交通大学	cxzhang@mail.xjtu.edu.cn
285	常晋源	西南财经大学、中国科学院数学与系	changjinyuan@swufe.edu.cn
286	林华珍	西南财经大学	linhz@swufe.edu.cn
287	张佛德	西南财经大学	fredzh@swufe.edu.cn
288	方匡南	厦门大学	xmufkn@163.com
289	姜佳佳	厦门大学	jiangjjxmu@qq.com
290	李晶茂	厦门大学	mr_ljm@foxmail.com
291	王淳林	厦门大学	wangc@xmu.edu.cn
292	徐福芝	厦门大学	1668274338@qq.com
293	张庆昭	厦门大学	qzzhang@xmu.edu.cn
294	钟威	厦门大学	wzhong@xmu.edu.cn
295	张云翼	香港中文大学（深圳）	zhangyunyi@cunk.edu.cn
296	李翠霞	徐州工程学院	79419383@qq.com
297	赵建强	徐州工程学院	zjqghh@xzit.edu.cn
298	吴月华	约克大学	wuyh@yorku.ca
299	陈飞	云南财经大学	chen_fei_77@hotmail.com
300	费宇	云南财经大学	feiyu.km@qq.com
300	赵慧	云南大学	zhaohuistat@ynu.edu.cn
301	周建军	云南大学	jjzhou@ynu.edu.cn

303	王亮	云南师范大学	liang610112@163.com
304	王秋平	肇庆学院	qp.wang@mails.ccnu.edu.cn
305	梁忠琦	浙大城市学院	2945155436@qq.com
306	李康强	浙江大学	11935023@zju.edu.cn
307	李翔	浙江大学	12135003@zju.edu.cn
308	骆威	浙江大学	weiluo@zju.edu.cn
309	张立新	浙江大学	stazlx@zju.edu.cn
310	蔡利	浙江工商大学	caili16@126.com
311	刘俊皓	浙江工商大学	1667294467@qq.com
312	覃水丹	浙江工商大学	shuidanqin@163.com
313	王炳兴	浙江工商大学	wangbingxing@163.com
314	王江峰	浙江工商大学	wjf2929@163.com
315	吴温慧	浙江工商大学	wuwenhui2888@163.com
316	徐安察	浙江工商大学	xuancha2011@aliyun.com
317	庄亮亮	浙江工商大学	a641292753@qq.com
318	刘祎	中国海洋大学	liuyi@ouc.edu.cn
319	曾靖	中国科学技术大学	zengjxl@ustc.edu.cn
320	贺百花	中国科学技术大学	baihua@ustc.edu.cn
321	胡彦林	中国科学技术大学	hyl11@mail.ustc.edu.cn
322	金百锁	中国科学技术大学	jbs@ustc.edu.cn
323	吕凌峰	中国科学技术大学	lief1997@mail.ustc.edu.cn
324	马梦园	中国科学技术大学	mmy@mail.ustc.edu.cn
325	缪柏其	中国科学技术大学	bqmiao@ustc.edu.cn
326	杜金烨	中国科学院数学与系统科学研究院	dujinye14@mails.ucas.ac.cn
327	刘子辰	中国科学院数学与系统科学研究院	liuzichen22@mails.ucas.ac.cn
328	骆文娟	中国科学院数学与系统科学研究院	luowenjuan@amss.ac.cn
329	潘越峰	中国科学院数学与系统科学研究院	panyuefeng@amss.ac.cn
330	盛赢	中国科学院数学与系统科学研究院	shengying@amss.ac.cn
331	石坚	中国科学院数学与系统科学研究院	jshi@iss.ac.cn
332	王启华	中国科学院数学与系统科学研究院	qhwang@amss.ac.cn
333	张新雨	中国科学院数学与系统科学研究院	xinyu@amss.ac.cn
334	韩威	中国科学院大学	hanwei19@mails.ucas.ac.cn
335	李惠超	中国科学院大学	2842067866@qq.com
336	李轩宇	中国科学院大学	lixuanyu22@mails.ucas.ac.cn
337	王瑞跃	中国科学院大学	wangruiyue21@mails.ucas.ac.cn
338	熊逸飞	中国科学院大学	night@loli.gs
339	张丰川	中国科学院大学	zhangfengchuan21@mails.ucas.ac.cn
340	张三国	中国科学院大学	sgzhang@ucas.ac.cn

341	祝国隽	中国科学院大学	861855561@qq.com
342	李再兴	中国矿业大学（北京）	lzxgscas@126.com
343	夏思薇	中国民用航空飞行学院	xsw@cafuc.edu.cn
344	王璐	中国农业大学	lwang@cau.edu.cn
345	田茂再	中国人民大学	mztian@ruc.edu.cn
346	杨松山	中国人民大学	yangss@ruc.edu.cn
347	朱利平	中国人民大学	zhu.liping@ruc.edu.cn
348	陈文	中国石油大学（北京）克拉玛依校区	chenwen@cupk.edu.cn
349	宋允全	中国石油大学（华东）	syqfly1980@163.com
350	王之鉴	中国石油大学（华东）	893737924@qq.com
351	詹珉珉	中国石油大学（华东）	z13695562072@163.com
352	张晓娣	中国石油大学（华东）	zhxd0258@163.com
353	刘斌霞	中南财经政法大学	lbx@stu.zuel.edu.cn
354	邵酩媛	中南财经政法大学	shaomingyuan111@163.com
355	于金钊	中南财经政法大学	yjz@stu.zuel.edu.cn
356	张瑜	中南财经政法大学	zhangyu@stu.zuel.edu.cn
357	程学伟	中南大学	xwcheng@csu.edu.cn
358	刘自力	中南大学	Zili_liu@csu.edu.cn
359	王洪	中南大学	wh@csu.edu.cn
360	王璐	中南大学	wanglu_stat@csu.edu.cn
361	罗敬	中南民族大学	stajluo@scue.edu.cn
362	汪宝彬	中南民族大学	wbb1818@126.com
363	王浩铭	中山大学	183527233@qq.com
364	潘蕊	中央财经大学	panrui_cufe@126.com
365	杨虎	中央财经大学	hu.yang@cufe.edu.cn
366	杨玥含	中央财经大学	yyh@cufe.edu.cn
367	郭水霞	湖南师范大学	guoshuixia75@163.com
368	周杰明	湖南师范大学	jmzhou@hunnu.edu
369	杨晶	湖南师范大学	yj@hunnu.edu.cn
370	赵维	湖南师范大学	18182105361@163.com
371	刘思晟	湖南师范大学	ssl1989@hunnu.edu.cn
372	耿旖旎	湖南师范大学	nicki3669@outlook.com
373	陈旭	湖南师范大学	chenxu981388@hunnu.edu.cn
374	欧辉	湖南师范大学	bt_huiou@hunnu.edu.cn
375	彭旭辉	湖南师范大学	xhpeng@hunnu.edu.cn
376	吴拥政	湖南师范大学	wuyzh10@163.com
377	曾杏元	湖南师范大学	xyzeng@hunnu.edu.cn
378	潘皓	湖南师范大学	13174223082@163.com
379	余霞	湖南师范大学	yyyxxx1126@163.com
380	李情情	湖南师范大学	lqq10086110@163.com
381	冯沐怡	湖南师范大学	645766293@qq.com