会 议 手 册

重庆市.南岸区

2025 年 6 月 7 日-6 月 8 日


重慶工商大學
Chongqing Technology and Business University

目 录

01 会议介绍

中国现场统计研究会多元分析应用专业委员会 2025 年学术年会暨翠湖论坛定于 2025 年 6 月 6 日-6 月 8 日召开，此会由重庆工商大学承办。重庆这座被长江与嘉陵江环抱的“立体山城”，以其魔幻地形、火辣美食与赛博夜景闻名。重庆不仅是一个现代化的都市，还保留了丰富的历史文化遗产和自然景观，是一个集现代与古典、自然与人文于一体的旅游胜地。

会议将邀请国内外世界著名专家学者、业界人士，针对多元分析的科学研究、技术创新、应用推广、面临的机遇和挑战、统计学专业和统计学学科建设、统计学和数学应用领域发展等方面作学术报告和讲座。本次大会旨在推动我国多元分析研究的理论和应用发展，搭建多元分析应用的交流平台，启迪思想，相互促进，合作研究，共同提高，从而推动相关研究领域的发展。

02 会议指南

会议时间及地点

●报到时间：2025 年 6 月 6 日

●报到地点：华商国际会议中心大堂

●会议地点：重庆工商大学图书馆剧场（大礼堂）

●会议时间：2025 年 6 月 7 日-6 月 8 日

会议安排

●6 月 6 日全天：报到

●6 月 7 日上午：开幕式及大会报告

●6 月 7 日下午：大会报告

●6 月 8 日上午：大会报告及分组报告

交通信息

路线 1：重庆江北国际机场 T2 航站楼

从机场乘坐轨道交通 3 号线至重庆工商大学站下车，下车后步行至华商国际会议中心、学豪酒店，全程约 60 分钟；出租车，全程约 30 分钟。

路线 2：重庆江北国际机场 T3 航站楼

乘坐轨道交通 10 号线至重庆北站南广场站，站内换乘轨道交通 3 号线（鱼洞方向）至重庆工商大学站，出站步行至华商国际会议中心、学豪酒店，全程约 70 分钟；出租车，全程约 25 分钟。

路线 3：重庆西站

乘坐轨道交通环线至四公里站，站内换乘轨道交通 3 号线至重庆工商大学站，出站步行至华商国际会议中心、学豪酒店，全程约 51 分钟；出租车，全程约 30 分钟。

路线 4：重庆北站

乘坐轨道交通 3 号线至重庆工商大学站，出站步行至华商国际会议中心、学豪酒店，全程约 40 分钟；出租车，全程约 25 分钟。

路线 5：重庆沙坪坝站

乘坐轨道交通 1 号线至两路口站，站内换乘轨道交通 3 号线至重庆工商大学站，出站步行至华商国际会议中心、学豪酒店，全程约 40 分钟；出租车，全程约 30 分钟。

会务组联系方式

会务组联系邮箱：liuyl@ctbu.edu.cn

会务组联系人：

杨炜明	15215022816	刘亚莉	13678429485
赵娟娟	13677617921	张明平	15683101381
赵培信	15703029859	刘　鎏	18883297010

03 主要参会人员

大会报告嘉宾

洪 永 淼	研究员	中国科学院数学与系统科学研究院
王 兆 军	教 授	南开大学
Jinchi Lv	教 授	南加州大学
周 晓 华	教 授	北京大学
胡 飞 芳	教 授	乔治华盛顿大学
蒋 建 成	教 授	北卡罗来纳大学夏洛特分校
姚 方	教 授	北京大学
周 望	教 授	新加坡国立大学
朱 宏 图	教 授	北卡罗来纳大学教堂山分校
Annie Qu	教 授	加州大学尔湾分校

联合主席

荆 炳 义　陈 义 安

成　　员

蒋 学 军　杨 炜 明　龙 宪 军　赵 培 信　胡 雪 梅

刘 丽 萍　杨 宜 平　陈 明 镜　程 素 丽　丁 黄 艳

秦 湘 斌　任　雪　于 学 霆　刘 婉 琪　易 之 瑶

04 会议日程安排

<table>
<tr><td colspan="4">2025 年 6 月 6 日（星期五）</td></tr>
<tr><td>10:00-22:00</td><td colspan="3">会议报到（华商国际会议中心大堂）</td></tr>
<tr><td colspan="4">2025 年 6 月 7 日（星期六）</td></tr>
<tr><td>时间</td><td>会议内容</td><td>主持人</td><td>地点</td></tr>
<tr><td>8:00-8:30</td><td colspan="2">会议报到（图书馆剧场）</td><td rowspan="10">重庆工商大学图书馆剧场（大礼堂）</td></tr>
<tr><td>8:30-9:00</td><td>开幕式-领导致辞
1. 校领导致辞
2. 张忠占理事长致辞
3. 张日权理事长致辞
4. 陈　敏理事长致辞
5. 邹长亮理事长致辞
6. 荆炳义理事长致辞</td><td>陈义安 教授
（ 重庆工商大学 ）</td></tr>
<tr><td>9:00 - 9:10</td><td colspan="2">合 影</td></tr>
<tr><td>9:10-10:00</td><td>洪永淼研究员报告
（中国科学院）</td><td>荆炳义 教授
（南方科技大学）</td></tr>
<tr><td>10:00-10:45</td><td>王兆军教授报告
（南开大学）</td><td>王维国 教授
（东北财经大学）</td></tr>
<tr><td>10:45-11:00</td><td colspan="2">茶　歇</td></tr>
<tr><td>11:00-11:45</td><td>Jinchi Lv 教授报告
（南加州大学）</td><td>朱利平 教授
（中国人民大学）</td></tr>
<tr><td>11:45-13:30</td><td colspan="3">午 餐（重庆工商大学兰园食堂二楼自助）</td></tr>
</table>

2025 年 6 月 7 日（星期六）			
时间	会议内容	主持人	地点
14:00-14:45	周晓华教授报告 （北京大学）	朱建平 教授 （厦门大学）	重庆工商大学图书馆剧场（大礼堂）
14:45-15:30	胡飞芳教授报告 （乔治华盛顿大学）	孔新兵 教授 （南京审计大学）	
15:30-16:15	蒋建成教授报告 （北卡罗来纳大学夏洛特分校）	张宝学 教授 （首都经济贸易大学）	
16:15-16:25	茶 歇		
16:25-17:10	姚方教授报告 （北京大学）	崔恒建 教授 （首都师范大学）	
17:10-17:55	周望教授报告 （新加坡国立大学）	蒋学军 研究员 （南方科技大学）	
17:55-18:00	宣读常务理事退出和新增名单	蒋学军 研究员 （南方科技大学）	
18:00-20:30	晚 宴（华商国际会议中心三楼宴会厅）		

2025 年 6 月 8 日（星期日）			
时间	会议内容	主持人	地点
8:30 - 9:15	朱宏图教授报告 （北卡罗来纳大学教堂山分校）	荆炳义 教授 （南方科技大学）	图书馆视听阅览室
9:15-10:00	Annie Qu 教授报告 （加州大学尔湾分校）		
10:00-10:15	茶 歇		
10:15-10:40	韩睿渐 （香港理工大学）	主持人： 田波平 教授 （哈尔滨工业大学） 召集人： 韩睿渐 （香港理工大学）	
10:40-11:05	周 帆 （上海财经大学）		
11:05-11:30	Yancheng Yuan （香港理工大学）		
11:30-11:55	Guohao Shen （香港理工大学）		
12:00-13:30	午 餐（重庆工商大学兰园食堂二楼自助）		

2025 年 6 月 8 日（星期日）				
时间	会议内容	主题	主持人	地点
6 月 8 日 8:30-12:30	大会报告 （线上）	大会报告 统计计算专场	荆炳义 教授 （南方科技大学） 田波平 教授 （哈尔滨工业大学）	图书馆视听阅览室
	分组报告 会场一	高维数据分析一	郭旭 教授 （北京师范大学） 解俊山 教授 （河南大学）	图书馆 2 楼多功能厅
	分组报告 会场二	非参数统计	蒋学军 研究员 （南方科技大学） 唐炎林 教授 （华东师范大学）	图书馆会议中心一会议室
	分组报告 会场三	函数型数据分析	邱志平 教授 （福建师范大学） 朱学虎 教授 （西安交通大学）	图书馆会议中心二会议室
	分组报告 会场四	复杂网络分析	刘婧媛 教授 （厦门大学） 刘广应 教授 （南京审计大学）	图书馆会议中心三会议室
	分组报告 会场五	流数据	严晓东 教授 （西安交通大学） 朱复康 教授 （吉林大学）	图书馆会议中心五会议室
	分组报告 会场六	统计推断一	黎德元 教授 （复旦大学） 杨玥含 教授 （中央财经大学）	厚德楼 2017
12:00-13:30	午　餐（重庆工商大学兰园食堂二楼自助）			

2025 年 6 月 8 日（星期日）				
时间	会议内容	主题	主持人	地点
6 月 8 日 8:30-12:30	分组报告 会场七	统计推断二	王德辉 教授 （辽宁大学） 鲜思东 教授 （重庆邮电大学）	厚德楼 3019
	分组报告 会场八	统计推断三	郭精军 教授 （兰州财经大学） 宋晓军 副教授 （北京大学）	厚德楼 3020
	分组报告 会场九	社会经济统计	陈立双 教授 （湖北经济学院） 杨炜明 教授 （重庆工商大学）	明德楼 1019
	分组报告 会场十	贝叶斯与模型平均	张志民 教授 （重庆大学） 李新民 教授 （青岛大学）	博智楼 70201
	分组报告 会场十一	机器学习一	王江峰 教授 （浙江工商大学） 冯峥晖 副教授 （哈尔滨工业大学）	博智楼 70202
	分组报告 会场十二	机器学习二	赵培信 教授 （重庆工商大学） 张军舰 教授 （广西师范大学）	博智楼 70203
	分组报告 会场十三	统计计算 与优化	张飞鹏 教授 （西安交通大学） 郝程程 副教授 （上海对外经贸大学）	博智楼 70204
	分组报告 会场十四	复杂数据建模	胡雪梅 教授 （重庆工商大学） 艾晓辉 副教授 （东北林业大学）	博智楼 70205
	分组报告 会场十五	高维数据与随机矩阵	高照省 教授 （电子科技大学） 胡　江 教授 （东北师范大学）	博智楼 70206
12:00-13:30	午　餐（重庆工商大学兰园食堂二楼自助）			

05 会场分布

会场	地点	会场	地点
大会报告（线上）	图书馆视听阅览室	分组报告会场八	厚德楼 3020
分组报告会场一	图书馆多功能厅	分组报告会场九	明德楼 1019
分组报告会场二	图书馆一会议室	分组报告会场十	博智楼 70201
分组报告会场三	图书馆二会议室	分组报告会场十一	博智楼 70202
分组报告会场四	图书馆五会议室	分组报告会场十二	博智楼 70203
分组报告会场五	图书馆三会议室	分组报告会场十三	博智楼 70204
分组报告会场六	厚德楼 2017	分组报告会场十四	博智楼 70205
分组报告会场七	厚德楼 3019	分组报告会场十五	博智楼 70206

06 分会场具体安排

分组报告会场一

（时间：2025 年 6 月 8 日；地点：图书馆 2 楼多功能厅）

时间	报告人	报告题目
主持人：郭旭、解俊山 Session 召集人：郭旭、解俊山		
8:30-8:55	郭旭 （北京师范大学）	Estimation and inference in quantile regression for high-dimensional partially linear models via convolution smoothing
8:55-9:20	解俊山 （河南大学）	Power Enhancement Test on High-dimensional Regression Coefficients of Partially Linear Models
9:20-9:45	郭朝会 （重庆师范大学）	Robust L_{2,0}-Penalized Rank Regression for High-dimensional Group Selection+
9:45-10:10	HU JIANHUA （上海财经大学）	Random projection-based response best-subset selector for ultra-high dimensional multivariate data
10:10-10:20	茶　歇	
10:20-10:45	林海明 （广东财经大学）	主成分的显著性检验与应用
10:45-11:05	刘梦雅 （华中师范大学）	An automatic multi-scale test for serial correlation of high-dimensional time series
11:05-11:30	孙高明 （河南大学）	Robust Transfer Learning for High-dimensional Sparse and Low-rank Matrix Regression
11:30-11:55	周泽人 （首都经济贸易大学）	Testing High-dimensional White Noise Based on Modified Portmanteaus Tests

分组报告会场二

（时间：2025 年 6 月 8 日；地点：图书馆一会议室）

时间	报告人	报告题目
主持人：蒋学军、唐炎林 Session 召集人：唐炎林、蒋学军		
8:30-8:55	唐炎林 （华东师范大学）	Conformal Prediction with Missing Data
8:55-9:20	刘耀午 （西南财经大学）	A simple and powerful method for composite null testing with application to mediation analysis
9:20-9:45	赵彦勇 （南京审计大学）	A New Single-index Model in Alzheimer's Disease Studies
9:45-10:10	付光辉 （昆明理工大学）	Imbalanced regression and its sparse feature selection
10:10-10:20	茶　歇	
10:20-10:45	蒋学军 （南方科技大学）	Penalized weighted GEEs for high-dimensional longitudinal data with informative cluster size
10:45-11:10	任好洁 （上海交通大学）	Augmented Localized Conformal Testing with Application to Novelty Detection
11:10-11:35	梁宇莉 （广西师范大学）	Testing hierarchical covariance structures under quadratic subspace
11:35-12:00	包亚杰 （南开大学）	Conformal Prediction with Cellwise Outliers: A Detect-then-Impute Approach
12:00 -12:25	王柔琳 （华东师范大学）	Unveiling Compositional Data with the Symmetric Chi-Square Kernel

分组报告会场三

（时间：2025 年 6 月 8 日；地点：图书馆二会议室）

时间	报告人	报告题目
主持人：邱志平、朱学虎 Session 召集人：张荣茂、郭旭		
8:30-8:55	杨莹 （复旦大学）	Online estimation for functional data
8:55-9:20	邱志平 （福建师范大学）	Global tests for detecting change in mean vector functions of multivariate functional data with repeated observations
9:20-9:45	朱学虎 （西安交通大学）	Adaptive functional change point detection
9:45-10:10	朱海斌 （暨南大学）	Estimation of volatility functionals with time-varying price staleness
10:10-10:20	茶　歇	
10:20-10:45	赖廷煜 （广西师范大学）	Testing and measuring the conditional mean (in)dependence for functional data by Martingale difference-angle divergence
10:45-11:10	彭选华 （西南政法大学）	函数型自回归条件异方差积分模型估计及应用
11:10-11:35	王佩淋 （西南政法大学）	函数型 AR-ARCHint-X 模型在比特币市场风险度量中的应用研究
11:35-11:50	俞易安 （南方科技大学）	Analyzing Functional Data with a Mixture of Covariance Structures Using a Curve-Based Sampling Scheme

分组报告会场四

（时间：2025 年 6 月 8 日；地点：图书馆三会议室）

时间	报告人	报告题目
主持人：刘婧媛、刘广应 Session 召集人：刘婧媛、蒋学军		
8:30-8:55	刘婧媛 （厦门大学）	BASIC: Bipartite Assisted Spectral-clustering for Identifying Communities in Large-scale Networks
8:55-9:20	韩　潇 （中国科学技术大学）	Individual-centered partial information in social networks
9:20-9:45	刘秉辉 （东北师范大学）	Community detection in weighted networks via the profile-pseudo likelihood method
9:45-10:10	潘　蕊 （中央财经大学）	Academic Literature Recommendation in Large-scale Citation Networks Enhanced by Large Language Models
10:10-10:20	茶　歇	
10:20-10:45	刘广应 （南京审计大学）	Heterogeneous Autoregressive Model for Symmetric Matrix-valued Time Series
10:45-11:10	王文胜 （杭州电子科技大学）	Fractal Nature of the Solution to Stochastic Heat Equation Driven by Fractional-Colored Noise
11:10-11:35	赵　毅 （哈尔滨工业大学）	时间序列与复杂网络相互表征及应用

分组报告会场五

（时间：2025 年 6 月 8 日；地点：图书馆五会议室）

时间	报告人	报告题目
主持人：严晓东、朱复康 Session 召集人：严晓东、朱复康		
8:30-8:55	严晓东 （西安交通大学）	TestAgent: Strategic A/B test via two-armed agent
8:55-9:20	朱复康 （吉林大学）	Mean-preserving rounding integer-valued ARMA models
9:20-9:45	杨凯 （长春工业大学）	On bivariate self-exciting hysteretic integer-valued autoregressive processes
9:45-10:10	李聪 （吉林大学）	几类时间序列模型在空气质量监测中的应用研究
10:10-10:20	茶　歇	
10:20-10:45	谢锦瀚 （云南大学）	Statistical inference for smoothed quantile regression with streaming data
10:45-11:10	马耀兰 （北方民族大学）	Extreme conditional expectile estimation for heavy-tailed ARMA-GARCH models
11:10-11:35	晏振 （广西师范大学）	Quantile regression model with uniform subsampling for streaming data
11:35-12:00	陈花萍 （河南大学）	A Trinomial Difference Autoregressive Process for the Bounded Z-valued Time Series
12:00-12:25	裴健 （北京建筑大学）	Mixed causal-noncausal count process

分组报告会场六

（时间：2025 年 6 月 8 日；地点：厚德楼 2017）

时间	报告人	报告题目
主持人：黎德元、杨玥含 Session 召集人：杨玥含、周永道		
8:30-8:55	黎德元 （复旦大学）	CTE2: Conditional Tail Expectation Treatment Effect at Extreme Levels
8:55-9:20	邱世芳 （重庆理工大学）	Rank-Based Nonparametric Methods for Evaluating Treatment Equivalence in AB/BA Crossover Trial Designs
9:20-9:45	杨玥含 （中央财经大学）	Unified inference framework and regression adjustment for stratified randomized experiments with ordinal outcomes
9:45-10:10	刘林 （上海交通大学）	Methods-of-moment inference for GLM and observational studies in proportional asymptotics
10:10-10:20	茶　歇	
10:20-10:45	陈建斌 （北京理工大学）	Design and modeling for order-of-addition factorial experiments with interaction effects
10:45-11:10	杨柳青 （中南大学）	Construction of (nearly) orthogonal symmetric Latin hypercube designs
11:10-11:35	胡雪梅 （重庆工商大学）	Robust matrix auto-regressions for matrix-valued time series with heavy-tailed distributions and digital finance analysis in Yangtze River Economic Belt
11:35-11:55	唐一苇 （复旦大学）	High-Dimensional Extreme Quantile Regression
11:55-12:15	王昭文 （复旦大学）	CoVaR under Asymptotic Independence

分组报告会场七

（时间：2025 年 6 月 8 日；地点：厚德楼 3019）

时间	报告人	报告题目
主持人：王德辉、鲜思东 Session 召集人：刘成、张志民		
8:30-8:55	黄丹阳 （中国人民大学）	Pseudo-Likelihood Ratio Screening based on Network Data with Applications
8:55-9:20	鲜思东 （重庆邮电大学）	Granular ball T-spherical fuzzy set for multi-criteria group decision-making in breast cancer treatment
9:20-9:45	夏强 （华南农业大学）	Determining The Number of Factors in Two-Way Factor Model of High-Dimensional Matrix-Variate Time Series: A White-Noise based Method for Serial Correlation Models
9:45-10:10	黄磊 （西南交通大学）	Weighted Estimation Method of High Dimension Portfolio Allocation for Time-Varying Stock Market
10:10-10:20	茶　歇	
10:20-10:45	许晓菲 （武汉大学）	Multi-Period Portfolio Allocation: A One-Shot Stochastic Optimization Approach
10:45-11:10	何平安 （北师香港浸会大学）	Statistical inference of Meta-Elliptical distributions - Based on MSE statistical representation points(MSE-RPs)
11:10-11:35	王可 （长春工业大学）	Estimation for single-index varying-coefficient spatial autoregressive model with index covariate measurement errors
11:35-12:00	霍雨阳 （南开大学）	False Discovery Rate Control over Data-driven Subgroups: A Unified Selective Approach

分组报告会场八

（时间：2025 年 6 月 8 日；地点：厚德楼 3020）

时间	报告人	报告题目
主持人：宋晓军、郭精军 Session 召集人：宋晓军、李周平		
8:30-8:55	宋晓军 （北京大学）	A consistent specification test for Lp-quantile regression
8:55-9:20	赵学靖 （兰州大学）	DyVAESurv: A VAE-enhanced model for dynamic survival analysis
9:20-9:45	郭精军 （兰州财经大学）	Some new results in option pricing using artificial intelligent algorithms
9:45-10:10	谭发龙 （湖南大学）	Goodness-of-Fit Tests for High-Dimensional Regression Models via Projections
10:10-10:20	茶　歇	
10:20-10:45	王春雨 （浙江大学）	Model selection with uncertainty in estimating optimal dynamic treatment regimes
10:45-11:10	张永霞 （北方工业大学）	Statistical Inference for High-dimensional Matrix-variate Factor Models with Missing Observations
11:10-11:35	罗姗姗 （北京工商大学）	Identification and estimation of causal peer effects using instrumental variables
11:35-12:00	王素敏 （河北工业大学）	A New Option Pricing Method Based on Generative Adversarial Networks and Space-Filling Design

分组报告会场九

（时间：2025 年 6 月 8 日；地点：明德楼 1019）

时间	报告人	报告题目
主持人：陈立双、杨炜明 Session 召集人：平卫英、刘成		
8:30-8:55	王 勇 （东北财经大学）	应对气候变化的统计行动：国际标准与国家经验
8:55-9:20	陈立双 （湖北经济学院）	互联网“免费”商品消费及居民真实生活成本指数估计
9:20-9:45	冯凌秉 （江西财经大学）	行政壁垒破除与资本跨区流动——基于企业异地投资的视角
9:45-10:10	朱 贺 （浙江工商大学）	基于数据质量效用的公共数据定价方法研究
10:10-10:20	茶 歇	
10:20-10:45	代金辉 （山东工商学院）	中国数字乡村建设水平：测算、区域差异及影响因素研究
10:45-11:10	沈寒蕾 （中南财经政法大学）	基于因果发现的产业链知识图谱构建与应用
11:10-11:35	陈鑫鹏 （河南财经政法大学）	数据要素共享与城市经济韧性 ——基于双重机器学习的因果推断
11:35-11:50	何义伟 （安徽财经大学）	中国陪诊服务市场调查—基于五省实证数据
11:50-12:05	李静雅 （安徽财经大学）	数字普惠金融赋能乡村振兴—基于县域数据的实证检验
12:05-12:20	卢玄龙 （重庆工商大学）	数字经济的生产力跃迁效应—回归分析与空间异质性

分组报告会场十

（时间：2025 年 6 月 8 日；地点：博智楼 70201）

时间	报告人	报告题目
主持人：李新民、张志民 Session 召集人：张志民、张军舰		
8:30-8:55	李新民 （青岛大学）	Frequentist model averaging under a linear exponential loss
8:55-9:20	吕晶 （西南大学）	Robust Model Averaging Prediction of Longitudinal Response with Ultrahigh-dimensional Covariates
9:20-9:45	夏小超 （重庆大学）	Relative Error Model Averaging for Multiplicative Models
9:45-10:10	王中雷 （厦门大学）	Balanced Active Learning
10:10-10:20	茶　歇	
10:20-10:45	周正 （北京工业大学）	An efficient Bayesian design method for estimating Shapley values.
10:45-11:10	胡爽 （重庆邮电大学）	Estimating the variogram matrix of the multivariate Hüsler-Reiss Pareto model using a truncated approach
11:10-11:25	蔡洪幸 （辽宁大学）	Optimal Design of the Multivariate Bayesian Generalized Likelihood Ratio Schemes for Monitoring the Mean Vector
11:25-11:40	廉爽 （辽宁大学）	Some Elaborate Comparisons of Several Distribution-Free Schemes in Short Production Runs and Some New Results
11:40-11:55	刘海芸 （辽宁大学）	Bayesian Elastic Net variable selection and application for spatial quantile panel autoregressive model

分组报告会场十一

（时间：2025 年 6 月 8 日；地点：博智楼 70202）

时间	报告人	报告题目
主持人：王江峰、冯峥晖 Session 召集人：毛晓军、李树威		
8:30-8:55	王江峰 （浙江工商大学）	Linear extremile regression and its semi-supervised learning
8:55-9:20	冯峥晖 （哈尔滨工业大学）	Exponential Power Mixture of Experts Model: Estimation, Clustering, and Variable Selection
9:20-9:45	衡佳妮 （北京工商大学）	融合机器学习和统计建模的风电场短期风能预测方法研究
9:45-10:10	沈亮 （青岛理工大学）	Spatial prediction of species invasion using a hybrid of machine learning and geostatistical method
10:10-10:20	茶　歇	
10:20-10:45	宋允全 （青岛理工大学）	Transfer Learning for higher-order partially linear spatial autoregressive model with differential privacy
10:45-11:10	夏思薇 （成都理工大学）	Transfer learning with invariance structures across target and auxiliary sources.
11:10-11:35	周灿 （南京审计大学）	Learning Semi-parametric Tree Models from Mixed Data
11:35-11:50	黄靖翔 （辽宁大学）	Tensor Transfer Learning under Convolutional Smooth Quantile and Its Application
11:50-12:05	张子晗 （辽宁大学）	Double/Debiased Machine Learning for Semiparametric Synthetic Difference-in-Differences Models
12:05-12:20	刘持金 （西安交通大学）	Deep Learning Based Doubly Robust Test for Granger Causality

分组报告会场十二

（时间：2025 年 6 月 8 日；地点：博智楼 70203）

时间	报告人	报告题目
主持人：赵培信、张军舰 Session 召集人：谢峰、王洪		
8:30-8:55	贺百花 （中国科学技术大学）	Optimal Multi-Machine Learning Assisted Semi-Supervised Inference
8:55-9:20	贺莘 （上海财经大学）	Kernel method with latent inputs and its applications to factor-based nonparametric regression
9:20-9:45	李树威 （广州大学）	Interpretable Deep Regression Models with Interval-Censored Failure Time Data
9:45-10:10	任明旸 （上海交通大学）	Model- and similarity-free transfer learning for classification with label noise
10:10-10:20	茶　歇	
10:20-10:45	王小舟 （华东师范大学）	Some progress in transfer learning
10:45-11:10	李灿辉 （河南大学）	A Doubly Robust Matched Learning Approach for Optimal Multi-Category Treatment Regimes in Observational Studies
11:10-11:35	程学伟 （湖南师范大学）	Generalization error bounds of gradient descent for learning a three-layer vision Transformer
11:35-12:00	章晓菲 （中南财经政法大学）	Robust personalized federated learning with sparse penalization

分组报告会场十三

（时间：2025 年 6 月 8 日；地点：博智楼 70204）

时间	报告人	报告题目
主持人：张飞鹏、郝程程 Session 召集人：赵普映、郝程程		
8:30-8:55	张飞鹏 （西安交通大学）	Consistent specification testing of linear censored quantile regression
8:55-9:20	郝程程 （上海对外经贸大学）	Influence Diagnostics for Tensor Regression Models
9:20-9:45	陆婧 （东北师范大学）	A Recursive Stochastic Algorithm for Real-Time Online Parameter Estimation in Item Response Theory: Enhancing Computational Efficiency for Dynamic Educational Assessment
9:45-10:10	马学俊 （苏州大学）	An algorithm for distributed parameter estimation in modal regression models
10:10-10:20	茶　歇	
10:20-10:45	王淳林 （厦门大学）	Semi-parametric inference on inequality measures with non-ignorable non-response using callback data
10:45-11:10	李艳鹏 （哈尔滨工业大学）	Sequential estimation of high-dimensional signal plus noise models under general elliptical frameworks
11:10-11:35	张榕 （云南大学）	A case study on the share holder network effect of stock market data: An SARMA approach
11:35-12:00	王天真 （华东师范大学）	隐私约束下的分布式保形预测

分组报告会场十四

（时间：2025 年 6 月 8 日；地点：博智楼 70205）

时间	报告人	报告题目
主持人：胡雪梅、艾晓辉 Session 召集人：田波平、王德辉		
8:30-8:55	张宇谦 （中国人民大学）	Data integration using covariate summaries from external sources
8:55-9:20	艾晓辉 （东北林业大学）	Exponential synchronization for the complex networks with intermittent event-triggered control under the second-order process
9:20-9:45	王月 （中国科学技术大学）	Estimation and Inference for Nonparametric Expected Shortfall Regression over RKHS
9:45-10:10	刘永欣 （南京审计大学）	Matrix-quantile factor prediction for generalized matrix-variate regression
10:10-10:20	茶　歇	
10:20-10:45	蔡雄 （南京审计大学）	Matrix-Factor-Augmented Regression
10:45-11:10	黄子健 （哈尔滨工业大学）	Stochastic Approximation MM Algorithms for Multiple Responses Mixed-effects Model and its Application
11:10-11:35	李晓妍 （重庆大学）	Efficient distributed estimation for expectile regression in increasing dimensions
11:35-11:50	孙沣扬 （辽宁大学）	A Nonparametric Control Chart for Monitoring Location based on Type II Censoring data
11:50-12:05	顾瑞芹 （辽宁大学）	Iterative Generalised Method of Moments Estimation of Synthetic Difference-in-Differences

分组报告会场十五

（时间：2025 年 6 月 8 日；地点：博智楼 70205）

时间	报告人	报告题目
主持人：高照省、胡江 Session 召集人：冯兴东、胡江		
8:30-8:55	高照省 （电子科技大学）	Sparse Asymptotic PCA: Identifying Sparse Latent Factors Across Time Horizon in High-Dimensional Time Series
8:55-9:20	王籼入 （西南财经大学）	Fast Online Change Point Detection for High-dimensional Regression Models
9:20-9:45	虞龙 （上海财经大学）	Multiplier bootstrap meets high-dimensional PCA
9:45-10:10	吴彬 （中国科学技术大学）	Nonstationary Binary Factor Models: Maximum Likelihood Estimation
10:10-10:20	茶　歇	
10:20-10:45	杨丽 （西安交通大学）	Asymptotic properties of a multicolored random reinforced urn model with an application to multi-armed bandits
10:45-11:10	李玉玲 （长春工业大学）	Application of the KOO Method in High-Dimensional Discriminant Analysis
11:10-11:35	刘芷君 （东北大学）	Asymptotic distributions of four linear hypotheses test statistics under generalized spiked model
11:35-11:50	崔健 （东北师范大学）	On the rate of convergence in the CLT for LSS of large-dimensional sample covariance matrices

07 大会邀请报告人简介及报告摘要

大会邀请学术报告一

报告时间：6 月 7 日上午　报告地点：图书馆千人报告厅

洪永淼，现任中国科学院数学与系统科学研究院冠名首席研究员，中国科学院大学经济与管理学院院长、发展中国家科学院院士、计量经济学会会士，并担任中国教育部高等学校经济学类专业教学指导委员会副主任委员。2010 至 2020 年间，洪教授曾担任康奈尔大学 Ernest S. Liu 经济学与国际研究讲席教授，2009 至 2010 年担任北美华人经济学家学会会长。研究领域涵盖计量经济学理论、时间序列计量经济学、金融计量经济学及统计学。其学术成果发表于 Annals of Statistics, Biometrika, Econometric Theory, Econometrica, International Economic Review, Journal of American Statistical Association, Journal of Business and Economic Statistics, Journal of Econometrics, Journal of Political Economy, Journal of Royal Statistical Society (Series B),等国际知名经济、金融与统计期刊。其最新英文专著《现代计量经济学基础：统一框架》广受学界关注。2014 至 2024 年，洪教授连续 11 年入选爱思唯尔"中国高被引学者（经济学/统计学）"榜单，并于 2022 年荣获国家级教学成果奖（高等教育本科）一等奖。

报告题目：High-dimensional Vector Autoregressions: A Ridge Dynamic Mode Decomposition Approach

报告摘要：This paper presents a novel methodology for estimating high-dimensional vector autoregressive (VAR) models. We use the dynamic mode decomposition (DMD), which is employed to analyze fluid dynamics in fluid mechanics, to recover the latent low-dimensional structure, moving beyond principal component analysis. Our method leverages information from the eigenspace of the transition matrix, enhancing the ability to capture the intrinsic dynamic characteristics of the VAR system and effectively capturing both temporal and cross-sectional dependencies. We introduce a truncated-ridge regularized DMD technique for estimating the high-dimensional transition matrix and establish its favorable large-sample properties. In an empirical application, we forecast inflation rates across 25 countries from 1995 to 2023 using our approach. The results demonstrate improved predictive accuracy and greater interpretability, presenting a robust alternative and providing new insights into the transmission mechanisms of international business cycle fluctuations.

大会邀请学术报告二

报告时间：6 月 7 日上午　报告地点：图书馆千人报告厅

王兆军，南开大学统计与数据科学学院执行院长/教授，教育部长江学者特聘教授，国务院学位委员会统计学科评议组成员，全国统计教材编审委员会委员；中国工业与应用数学学会副理事长，中国统计教育学会副会长，中国工业统计教学研究会副会长，中国概率统计学会副理事长。曾任国家统计专家咨询委员会委员、国家自然科学基金委天元基金领导小组成员、中国现场统计研究会副理事长、天津市现场统计研究会理事长，天津工业与应用数学学会理事长，曾获国务院政府特贴、全国百篇优博指导教师、教育部自然科学二等奖及天津市自然科学一等奖。

报告题目：Semi-supervised distribution learning

报告摘要：This study addresses the challenge of distribution estimation and inference in a semi-supervised setting. In contrast to prior research focusing on parameter inference, this work explores the complexities of semi-supervised distribution estimation, particularly the uniformity problem inherent in functional processes. To tackle this issue, we introduce a versatile framework designed to extract valuable information from unlabeled data by approximating a conditional distribution on covariates. The proposed estimator is derived from the K-fold cross-fitting strategy, exhibiting both consistency and asymptotic Gaussian process properties. Under mild conditions, the proposed estimator outperforms the empirical cumulative distribution function in terms of asymptotic efficiency. Several applications of the methodology are given, including parameter inference and goodness-of-fit tests.

大会邀请学术报告三

报告时间：6 月 7 日上午　报告地点：图书馆千人报告厅

Jinchi Lv，现任美国南加州大学马歇尔商学院数据科学与运筹学系主任、Kenneth King Stonier 讲席教授，并兼任该校数学系教授。他于 2007 年获普林斯顿大学数学博士学位。2016 至 2019 年期间，他曾任南加州大学 McAlister 副教授。其研究领域涵盖统计学、数据科学、人工智能、机器学习及其商业应用，并延伸至区块链与大语言模型方向。学术成果发表于统计学、经济学、商学、计算机科学、信息论、神经科学及生物学等多个学科领域的权威期刊。

报告题目： ARK: Robust Knockoffs Inference with Coupling

报告摘要： We investigate the robustness of the model-X knockoffs framework with respect to the misspecified or estimated feature distribution. We achieve such a goal by theoretically studying the feature selection performance of a practically implemented knockoffs algorithm, which we name as the approximate knockoffs (ARK) procedure, under the measures of the false discovery rate (FDR) and k-familywise error rate (k-FWER). The approximate knockoffs procedure differs from the model-X knockoffs procedure only in that the former uses the misspecified or estimated feature distribution. A key technique in our theoretical analyses is to couple the approximate knockoffs procedure with the model-X knockoffs procedure so that random variables in these two procedures can be close in realizations. We prove that if such coupled model-X knockoffs procedure exists, the approximate knockoffs procedure can achieve the asymptotic FDR or k-FWER control at the target level. We showcase three specific constructions of such coupled model-X knockoff variables, verifying their existence and justifying the robustness of the model-X knockoffs framework. Additionally, we formally connect our concept of knockoff variable coupling to a type of Wasserstein distance. This is a joint work with Yingying Fan and Lan Gao.

大会邀请学术报告四

报告时间：6 月 7 日上午　报告地点：图书馆千人报告厅

蒋建成，北卡大学夏洛特分校数学统计系和数据科学学院的双聘教授，曾兼任南开大学讲座教授。他的研究领域包括非参数建模，金融时间序列分析，生存分析，统计和机器学习，大数据分析等。目前已在包括统计年刊(AOS)，美国统计学会会刊（JASA），英国皇家统计学会会刊(JRSSB)，生物识别杂志(Biometrika)，计量经济学杂志(Journal of Financial Econometrics, Econometrics)等刊物上发表科研论文 70 余篇。自 2018 至 2024 年担任北卡大学数学统计系统计学项目主管。目前兼任该校可信人工智能中心联合研究员。

报告题目： Targeted Inference for High-Dimensional Quantile Regression Models

报告摘要： Quantile regression is a powerful tool for uncovering relationships between predictors and responses, particularly in high-dimensional data where varied effects can be detected. Focusing on key predictors at a specific quantile, while treating others as nuisance variables, presents a unique hypothesis testing challenge. This research introduces an innovative inference framework that employs dimension-reduced convolution-smoothed quantile regression, while avoiding estimating the inverse of high-dimensional covariance matrix of the predictors. By calibrating the regularization parameter, we develop a data-driven test that can be shown to be an oracle test with probability tending to one. To mitigate the selective bias induced by dimension reduction and ensure valid inference, we implement a cross-fitting strategy by dividing the dataset into two parts: one for model selection and the other for parameter estimation. This process yields a fused estimator, derived from an informative weighting method that combines estimators from both dataset partitions. The optimal fused estimator aids in constructing confidence intervals and performing Wald-type tests for targeted parameters. We establish the Bahadur's representation of this estimator and obtain limiting distributions of the test statistics under both null and alternative hypotheses, with the number of parameters diverging to infinity. Advantages of our tests are further highlighted by theoretical power comparisons to some competitive tests. Empirical studies confirm effectiveness of the proposed tests across various linear parameter hypotheses. Additionally, we illustrate the use of the proposed methodology through two real-world data analyses.

大会邀请学术报告五

报告时间：6 月 7 日下午　报告地点：图书馆千人报告厅

胡飞芳，美国乔治华盛顿大学统计系教授。2004 年美国自然科学基金会杰出青年基金的得主。2009 年当选为美国统计协会和国际数理统计协会双料 fellow,是自适应设计与方法的世界领先研究专家。一直致力于统计理论及相关应用的研究。研究内容涉及自适应方法，生物统计，个性化医疗，临床实验设计，大数据分析，线上 A/B 检测的实验设计与分析方法等。2000 年以来国际顶级统计杂志发表学术论文 100 余篇。出版自适应设计与方法的英文专著 1 部。2007 年受美国 FDA 邀请撰写白皮书论文 2 篇。胡飞芳教授曾担任 JASA 和 Annals of Statistics 等国际顶级统计杂志的副主编。胡教授主持六项美国国家自然科学基金研究项目，以及多项香港新加坡研究项目。近些年，他在自适应设计与方法的研究成果引起了同行的高度关注，多次受邀在国际学术会议上做主题报告。组织多次国际学术会议并担任会议共同主席。是多个临床试验的数据与安全监察委员会（DSMB）成员。同时受邀为多家世界五百强企业提供统计咨询。

报告题目： Why Many of Our Key Findings Are Often Misleading: Unobserved Confounders and Their Impacts

报告摘要： In observational studies, it is well known that unobserved cofounders (covariates) could mislead our key findings. Randomized controlled trials (RCTs) are widely regarded as the gold standard for identifying causal effects, largely due to their ability to mitigate confounding. However, even in RCTs, unobserved covariates can threaten the validity of conclusions. In this talk, we explore the role and impact of unobserved covariates in randomized studies by addressing two key questions:

1.Does balancing observed covariates necessarily lead to better balance in unobserved covariates?

2.Can we always trust the conclusions drawn from randomized controlled trials?

For the first question, we demonstrate that when observed covariates are discrete, balancing them often improves the balance of unobserved covariates as well. For the second, we show that if unobserved covariates interact with treatment, the resulting treatment effects may become unidentifiable. This can lead to inconsistent estimates, regardless of the randomization scheme used.

大会邀请学术报告六

报告时间：6 月 7 日下午　报告地点：图书馆千人报告厅

周晓华，北京大学讲席教授，生物统计系创系主任。美国科学促进会（AAAS）、美国统计学会（ASA）和数理统计学会（IMS）会士，并担任国际生物统计学会中国区域分会主席。他在顶尖统计学期刊 JRSSB、JASA、Biometrics、Biometrika、Annals of Statistics 上发表了逾 290 篇统计学方法论和医学研究论文。荣获美国联邦政府退伍军人事务部颁发的“Research Career Scientist 奖”、中国国家自然科学基金“Distinguished Overseas Young Scientist 奖”以及国际贝叶斯分析学会 Mitchell 奖。周教授在诊断医学方法学发展方面做出了许多重要贡献，特别是在解决验证偏倚和不完善金标准偏倚问题上；他在因果推断领域也有重要方法论贡献，尤其体现在处理随机化破坏的实验或复杂临床试验情形的统计方法上；此外，他在医疗成本分析与预测方面贡献卓著，其研究解决了导致医疗成本分布偏离正态性的四大特征问题，学术成就卓著。

报告题目： Nonparametric Receiver Operating Characteristic Curve Analysis with an Imperfect Gold Standard

报告摘要： In this talk, I address the challenge of estimating receiver operating characteristic (ROC) curves and the areas under these curves (AUC) in the context of an imperfect gold standard, a common issue in diagnostic accuracy studies. I discuss the nonparametric identification and estimation of ROC curves and AUCs when the reference standard for disease status is prone to error. When the known or estimable accuracy of this imperfect reference standard is available, under the conditional independent assumption, we show the identifiability of ROC curves and propose a nonparametric estimation method. In cases where the accuracy of the imperfect reference standard remains unknown, we establish that while ROC curves are unidentifiable, the sign of the difference between two AUCs is identifiable. This insight leads us to develop a hypothesis testing method for assessing the relative superiority of AUCs. Compared to the existing methods, the proposed methods are nonparametric so that they do not rely on the parametric model assumptions. In addition, they are applicable to both the ROC/AUC analysis of continuous biomarkers and the AUC analysis of ordinal biomarkers. Our theoretical results and simulation studies validate the proposed methods, which we further illustrate through application in two real-world diagnostic studies. This is a joint work with a PhD student at Peking University, Jiarui Sun.

大会邀请学术报告七

报告时间：6 月 7 日下午　报告地点：图书馆千人报告厅

姚方，北京大学讲席教授、入选国家高层次人才计划，北大统计科学中心主任、概率统计系主任。国际数理统计学会（IMS）Fellow，美国统计学会（ASA）Fellow，获 2024 年第六届科学探索奖（数学物理学领域）。2000 年本科毕业于中国科技大学统计专业，2003 获得加利福尼亚大学戴维斯分校统计学博士学位，曾任职于多伦多大学统计科学系长聘正教授。至今担任 9 个国际统计学核心期刊的主编或编委，包括《加拿大统计学期刊》主编，顶级期刊《Journal of the American Statistical Association》和《Annals of Statistics》编委等。

报告题目：Semiparametric M-estimation with Overparameterized Neural Networks

报告摘要：We focus on semiparametric regression that has played a central role in statistics, and exploit the powerful learning ability of deep neural networks (DNNs) while enabling statistical inference on parameters of interest that offers interpretability. Despite the success of classical semiparametric method/theory, establishing the root-n consistency and asymptotic normality of the finite-dimensional parameter estimator in this context remains challenging, mainly due to nonlinearity and potential tangent space degeneration in DNNs. In this work, we introduce a foundational framework for semiparametric M-estimation, leveraging the approximation ability of overparameterized neural networks that circumvent tangent degeneration and align better with training practice nowadays. The optimization properties of general loss functions are analyzed, and the global convergence is guaranteed. Instead of studying the “ideal” solution to minimization of an objective function in most literature, we analyze the statistical properties of algorithmic estimators, and establish nonparametric convergence and parametric asymptotic normality for a broad class of loss functions. These results hold without assuming the boundedness of the network output and even when the true function lies outside the specified function space. To illustrate the applicability of the framework, we also provide examples from regression and classification, and the numerical experiments provide empirical support to the theoretical findings.

大会邀请学术报告八

报告时间：6 月 7 日下午　报告地点：图书馆千人报告厅

周望，2004 年 7 月起在新加坡国立大学统计系任教，并于 2009 年 1 月获终身教授。现为新加坡国立大学教授，国际著名期刊 Random Matrices-Theory and Applications 的主编。主要研究方向为: High dimensional statistics，Random matrices, SLE,等。近年来发表有高水平论文九十多篇。 其中在概率统计学方面的国际公认的顶尖杂志 Annals of Statistics, Journal of American Statistical Association, Biometrika, Annals of Probability, Probability Theory and Related Fields, Annals of Applied Probability 上发表论文二十余篇。2005 年起主持新加坡政府基金项目十余项。2012 获国际统计学会当选成员（Elected Member of International Statistical Institute）；2021 年获国际数理统计学会(IMS)Fellow。

报告题目：Necessary and sufficient condition for CLT of linear spectral statistics of sample correlation matrices

报告摘要：In this talk, I will discuss the central limit theorem (CLT) for the linear spectral statistics (LSS) of sample correlation matrix, constructed from a pxn data matrix X with independent and identically distributed entries having mean zero, variance one, and infinite fourth moments in the high-dimensional regime n/p tending to c in (0,infinity)\ (1). Especially a necessary and sufficient condition for the CLT is presented.

大会邀请学术报告九

报告时间：6 月 8 日上午　报告地点：图书馆视听阅览室

朱宏图，现任美国北卡罗来纳大学教堂山分校生物统计学系、统计学系、放射学系、计算机科学系及遗传学系凯南杰出教授。2018 至 2020 年间，他曾任滴滴出行首席统计科学家并获授滴滴学者荣誉；2016 至 2018 年担任 MD 安德森癌症中心 Bao-Shan Jing 诊断影像冠名讲席教授。作为国际公认的权威专家，其研究领域涵盖统计学习、医学影像分析、精准医学、生物统计学、人工智能及大数据分析。他曾于 2016 年获美国得克萨斯州癌症预防研究所杰出研究者奖，2019 年荣获国际运筹学与管理科学协会（INFORMS）丹尼尔·H·瓦格纳运筹学实践卓越奖，并获 2025 年国际统计学界最高荣誉之一的 COPSS 斯内德克奖。朱教授在 Nature, Science, Cell, Nature Genetics, Nature Communication, PNAS, AOS, JASA, Biometrika, and JRSSB 等顶级期刊发表论文 340 余篇，并在 NeurIPS、ICLR、ICML、AAAI、KDD 等顶尖学术会议发表 58 篇会议论文。他是国际数理统计学会会士、美国统计学会会士及国际电气与电子工程师协会会士，现任 JASA 的 coordinating editor 和 JASA ACS.的 editor。

报告题目：Conditional Distributional Learning with Non-crossing Quantile Network and applications.

报告摘要：We introduce the Non-Crossing Quantile (NQ) Network, a novel approach for conditional distribution learning. By incorporating non-negative activation functions, the NQ network ensures monotonicity in learned distributions, effectively eliminating the issue of quantile crossing. The NQ network offers a highly adaptable deep distributional learning framework, applicable to a wide range of tasks, from non-parametric quantile regression to causal effect estimation and distributional reinforcement learning (RL). We further establish a comprehensive theoretical foundation for the deep NQ estimator and its application indistributional RL, providing rigorous analysis to support its effectiveness. Extensive experiments demonstrate the robustness and versatility of the NQ network across various domains, including clinical trials, e-commerce, games, and healthcare, highlighting its potential for real-world applications. This is based on a series of joint works with Drs. Shen, Luo, and Shi and Mr. Huang.

大会邀请学术报告十

报告时间：6 月 8 日上午　报告地点：图书馆视听阅览室

Annie Qu，美国加州大学尔湾分校统计系校聘教授。曾获伊利诺伊大学厄巴纳-香槟分校文理学院 Brad and Karen Smith 冠名学者教授荣誉，于 2004 至 2009 年间荣获美国国家科学基金会杰出青年学者奖。Qu 教授是国际数理统计学会会士、美国统计学会会士及美国科学促进会会士，2024 年获颁国际数理统计学会奖章并担任学会特邀讲座嘉宾。2023 至 2025 年，她担任 Journal of the American Statistical Association Theory and Methods 的 Co-Editor；2021 至 2027 年任国际数理统计学会项目秘书长；2025 年将出任美国统计学会理事会分会主席团主席。2025 年，荣获国际数理统计学会最高荣誉卡尔弗奖章。

报告题目： Optimal Transport for Latent Integration with An Application to Heterogeneous Neuronal Activity Data

报告摘要： Detecting dynamic patterns of task-specific responses shared across heterogeneous datasets is an essential and challenging problem in many scientific applications in medical science and neuroscience. In our motivating example of rodent electrophysiological data, identifying the dynamical patterns in neuronal activity associated with ongoing cognitive demands and behavior is key to uncovering the neural mechanisms of memory. One of the greatest challenges in investigating a cross-subject biological process is that the systematic heterogeneity across individuals could significantly undermine the power of existing machine learning methods to identify the underlying biological dynamics. In addition, many technically challenging neurobiological experiments are conducted on only a handful of subjects where rich longitudinal data are available for each subject. The low sample sizes of such experiments could further reduce the power to detect common dynamic patterns among subjects. In this talk, we propose a novel heterogeneous data integration framework based on optimal transport to extract shared patterns in complex biological processes. The key advantages of the proposed method are that it can increase discriminating power in identifying common patterns by reducing heterogeneity unrelated to the signal by aligning the extracted latent spatiotemporal information across subjects. Our approach is effective even with a small number of subjects and does not require auxiliary matching information for the alignment. In particular, our method can align longitudinal data across heterogeneous subjects in a common latent space to capture the dynamics of shared patterns while utilizing temporal dependency within subjects. Our numerical studies on both simulation settings and neuronal activity data indicate that the proposed data integration approach improves prediction accuracy compared to existing machine learning methods.

08 特邀学术报告人简介及报告摘要

统计计算专场报告 1

报告时间：6 月 8 日上午　报告地点：图书馆视听阅览室

报告题目：Statistical ranking with dynamic covariates

报告人：韩睿渐

摘要：We introduce a general covariate-assisted statistical ranking model within the Plackett--Luce framework. Unlike previous studies focusing on individual effects with fixed covariates, our model allows covariates to vary across comparisons. This added flexibility enhances model fitting yet brings significant challenges in analysis. This paper addresses these challenges in the context of maximum likelihood estimation (MLE). We first provide sufficient and necessary conditions for both model identifiability and the unique existence of the MLE. Then, we develop an efficient alternating maximization algorithm to compute the MLE. Under suitable assumptions on the design of comparison graphs and covariates, we establish a uniform consistency result for the MLE, with convergence rates determined by the asymptotic graph connectivity. We also construct random designs where the proposed assumptions hold almost surely. Numerical studies are conducted to support our findings and demonstrate the model's application to real-world datasets, including horse racing and tennis competitions.

报告人简介：韩睿渐，香港理工大学助理教授。2020 年在香港科技大学获得博士学位，之后在香港中文大学从事研究型助理教授工作，并在 2022 年加入香港理工大学。他的研究兴趣主要包括：排序数据分析，高维统计，在线推断，统计机器学习。其科研成果发表在 JASA, Biometrika, AAP, JBES 等期刊上。

统计计算专场报告 2

报告时间：6 月 8 日上午　报告地点：图书馆视听阅览室

报 告 题 目 ： Distributional Off-Policy Evaluation with Deep Quantile Process Regression

报告人：周帆

摘 要 ： This paper investigates the off-policy evaluation (OPE) problem from a distributional perspective, with the aim of modeling the entire distribution of total returns, rather than focusing solely on estimating the expectation (value function), as most existing OPE methods do. Specifically, we introduce a quantile-based approach for OPE using deep quantile process regression, presenting a novel algorithm called Deep Quantile Process regression-based Off-Policy Evaluation (DQPOPE). We provide new theoretical insights into the deep quantile process regression technique, extending existing approaches that estimate discrete quantiles to estimate a continuous quantile function. A key contribution of our work is the rigorous sample complexity analysis for distributional OPE with deep neural networks, bridging theoretical analysis with practical algorithmic implementations. We show that DQPOPE achieves statistical efficiency by estimating the full return distribution using the same sample size required to estimate a single policy value using conventional methods. Furthermore, our empirical studies illustrate that DQPOPE provides significantly more precise and robust policy value estimates than standard methods, thereby enhancing the practical applicability and effectiveness of distributional reinforcement learning approaches.

报告人简介：周帆，上海财经大学统计与管理学院副教授，博士毕业于美国北卡罗莱纳大学教堂山分校。主要研究方向包括强化学习，深度学习，因果推断。在 Journal of American Statistical Association，Journal of Machine Learning Research, Biometrics 等统计学机器学习期刊以及 NeurIPS, ICML, KDD 等国际人工智能顶会接收发表一作通讯文章数十篇，曾获得国际泛华统计协会新研究者奖，北卡教堂山分校 Barry H. Margolin Award，并入选上海市人才计划（青年）。

统计计算专场报告 3

报告时间：6 月 8 日上午　报告地点：图书馆视听阅览室

报告题目：Unified Parameter-Efficient Unlearning for LLMs

报告人：Yancheng Yuan

摘要：The advent of Large Language Models (LLMs) has revolutionized natural language processing, enabling advanced understanding and reasoning capabilities across a variety of tasks. Fine-tuning these models for specific domains, particularly through Parameter-Efficient Fine-Tuning (PEFT) strategies like LoRA, has become a prevalent practice due to its efficiency. However, this raises significant privacy and security concerns, as models may inadvertently retain and disseminate sensitive or undesirable information. To address these issues, we introduce a novel instance-wise unlearning framework, LLMEraser, which systematically categorizes unlearning tasks and applies precise parameter adjustments using influence functions. Unlike traditional unlearning techniques that are often limited in scope and require extensive retraining, LLMEraser is designed to handle a broad spectrum of unlearning tasks without compromising model performance. Extensive experiments on benchmark datasets demonstrate that LLMEraser excels in efficiently managing various unlearning scenarios while maintaining the overall integrity and efficacy of the models.

报告人简介：Dr. Yancheng Yuan is currently an assistant professor in the Department of Applied Mathematics at the Hong Kong Polytechnic University. His primary research interests lie in the theoretical and applied aspects of continuous optimization and machine learning. He has published papers in prestigious journals such as the "SIAM Journal on Optimization," "Journal of Machine Learning Research," and "IEEE Transactions on Neural Networks and Learning Systems", as well as presented at top academic conferences in the field of machine learning such as ICML, NeurIPS, and WWW. He has received the Best Paper Award Finalist (WWW 2021).

统计计算专场报告 4

报告时间：6 月 8 日上午　报告地点：图书馆视听阅览室

报告题目：Deep nonparametric estimation for volatility functions via iterative algorithms.

报告人：Guohao Shen

摘要：This paper introduces a novel approach to estimate nonparametric GARCH models using deep neural networks. We propose an efficient iterative algorithm for training these deep estimators, characterized by ease of implementation and adaptability to various model settings and loss functions. We establish learning guarantees for the proposed method, including non-asymptotic upper bounds on prediction error under mild assumptions. Notably, we demonstrate that our deep neural network estimator can adapt to the true lag dimension of the volatility model even when the input dimension is overspecified. This crucial property ensures optimal performance even with suboptimal input choices. We validate the effectiveness of our approach through extensive simulations, showcasing its superiority over competing methods, particularly in high-dimensional, nonlinear, and complex volatility scenarios.

报告人简介：Guohao Shen，香港理工大学应用数学系助理教授，2022 年获香港中文大学统计学博士学位。主要研究方向为机器学习和非参数统计，重点关注深度学习的统计基础理论。在统计学顶刊 AoS、Biometrika，机器学习和计算机顶刊 JMLR，计量经济顶刊 JoE，以及 NeurIPS 和 ICML 等顶会发表多篇学术论文。2024 年荣获约翰·艾奇森统计学奖（John Aitchison Prize in Statistics）。

高维数据分析专场报告 1

报告时间：6 月 8 日上午　报告地点：图书馆多功能厅

报告题目： Estimation and inference in quantile regression for high-dimensional partially linear models via convolution smoothing

报告人： 郭旭

摘要： This paper aims to develop new estimation and inference procedures for high-dimensional partially linear quantile regression (QR) models. Compared with least squares methods, QR presents unique challenges due to the non-smoothness of its loss function and the non-additivity of conditional quantile. To address the challenges, we apply convolution-smoothing technique to handle the non-smoothness and weighted projection technique to deal with the non-additivity. Specifically, the estimation procedure approximates the non-parametric function by B-spline and employs an L1 regularization for linear coefficients. Theoretically, we establish a new non-asymptotic smoothness-adjusted second-order effect property which holds for a wide range of non-parametric regression methods. Furthermore, we propose a debiased Lasso estimator using a newly proposed projection strategy. The strategy involves estimating the conditional density function of random errors, which introduces an uncontrollable error. We adopt the double smoothing technique to address the issue and establish asymptotic normality for debiased estimator. The proposed methods are evaluated through numerical simulations and an analysis of the relationship between maternal age and infant birth weight.

报告人简介： 郭旭，现为北京师范大学统计学院教授，博士生导师。郭老师一直从事回归分析中复杂假设检验的理论方法及应用研究，近年来旨在对高维数据发展适当有效的检验方法。部分成果发表在 JRSSB, JASA，Biometrika 和 JOE。现主持国家自然科学基金青年科学基金项目 B 类（原优秀青年科学基金）。曾荣获北师大第十一届“最受本科生欢迎的十佳教师”，北师大第十八届青教赛一等奖和北京市第十三届青教赛三等奖。

高维数据分析专场报告 2

报告时间：6 月 8 日上午　报告地点：图书馆多功能厅

报告题目：Power Enhancement Test on High-dimensional Regression Coefficients of Partially Linear Models

报告人：解俊山

摘要：This paper is concerned with the testing on regression coefficients of high-dimensional partially linear regression models. A power enhancement test method is proposed based on the U-statistics test via random integration technique, and the asymptotic distribution of the proposed test statistic is investigated when both the dimension and the sample size go to infinity.　It is shown that the proposed test method outperforms the existing method in terms of power under some situations by the asymptotic relative efficiency. The finite sample performance of the proposed test method is evaluated through numerical studies, and the empirical results demonstrate that the proposed test method is more powerful than the existing tests in a wide range of alternative hypothesis settings.

报告人简介：解俊山，河南大学教授，博士生导师。美国明尼苏达大学统计系、香港浸会大学数学系访问学者。主要研究方向为高维统计推断和随机矩阵理论。以第一作者（或通讯作者）在 J.Theor. Probab.、J.Multivariate Anal.、Sci.China:Math. 等发表 SCI 学术论文 30 余篇。

高维数据分析专场报告 3

报告时间：6 月 8 日上午　报告地点：图书馆多功能厅

报告题目：Robust L_{2,0}-Penalized Rank Regression for High-dimensional Group Selection

报告人：郭朝会

报告摘要：Sparse group selection is the process of selecting a small part of nonoverlapping groups to achieve the good interpretability and prediction on the response, and it has recently seen increasing applications in machine learning, image processing and bio-medical fields. However, developing robust and efficient algorithms for group selection remains a challenging research topic due to the computational complexity and potential outliers in high-dimensional settings. Motivated by the superior performance of rank-based methodology, we design a fast and efficient algorithm based on the $\ell_{2,0}$ penalty to achieve the goal of robust group selection for a given size of active groups s. This new algorithm can iteratively detect the active groups and exclude the irrelevant ones. When is not less than (the true size of active groups), we theoretically prove that the proposed algorithm covers the true subset of active groups with high probability and the estimation error of the solution sequence generated by our algorithm decays to the optimal error bound in a few iterations. Moreover, coupled with the group Bayesian information criterion, an adaptive algorithm is further introduced to determine the optimal. Theoretically, without any prior knowledge of , the proposed adaptive algorithm is able to exactly identify the true subset of active groups with probability approaching to one. Finally, extensive simulation examples show that our method outperforms existing competitors, resulting in significant improvements in terms of efficiency and accuracy of group selection and parametric estimation. The Bardet-Biedl syndrome gene expression data set is also analyzed to illustrate the application of our proposed method.

报告人简介：郭朝会，重庆师范大学教授，硕士生导师。2016 年毕业于重庆大学统计学专业获得博士学位，2018 年 10 月至 2019 年 1 月在新加坡国立大学统计与应用概率系访学。主持国家社会科学基金青年 1 项、国家自然科学基金青年 1 项、重庆市自然科学基金 5 项。主要研究方向为因子模型、稳健估计、纵向数据，已在《Journal of Business & Economic Statistics》、《Computational Statistics and Data Analysis》、《SCIENCE CHINA Mathematics》等 SCI 期刊发表多篇学术研究成果。

高维数据分析专场报告 4

报告时间：6 月 8 日上午　报告地点：图书馆多功能厅

报告题目： Random projection-based response best-subset selector for ultra-high dumensional multivariate data

报告人： HU JIANHUA

摘要： In this talk, we propose a random projection-based response best-subset selector to perform response variable selection in ultra-high dimensional multivariate data, where both the dimensions of response and predictor variables are substantially greater than the sample size.

This method is developed by integrating the response best-subset selector and random projection technique which is applied to reduce dimensionality of predictors. Under a multivariate tail eigenvalue condition, such a random projection-based dimensionality reduction of predictors only leads to an ignorable error between the original and dimension-reduced models. A computational procedure is presented. The proposed method exhibits model consistency under some certain conditions. The efficiency and merit of the proposed method are strongly supported by extensive finite-sample simulation studies. A real breast cancer dataset spanning 22 chromosomes are analyzed to demonstrate the proposed method.

报告人简介： HU JIANHUA，上海财经大学统计与数据科学学院研究员，国际权威统计杂志《Journal of Multivariate Analysis》副主编、中国环境统计学会大数据科学分会常务理事。研究方向为高维数据/大数据、空间面板数据、多元分析等。

高维数据分析专场报告 5

报告时间：6 月 8 日上午　报告地点：图书馆多功能厅

报告题目：主成分的显著性检验与应用

报告人：林海明

摘要：主成分分析的优良性准则和检验，是美国科学院院士 Anderson 等教授待解决的问题。本文按照检验要素及其构架，提出了 7 个待解决的问题，并逐一进行了探讨。首先，基于主成分模型及其解，构建了包含公共成分、特殊成分和误差的扩展模型；其二，建立了主成分分析的优良性准则；其三，提出了主成分的如下 5 个检验：变量与主成分相关阵的简单结构检验，特殊成分检验，公共成分检验，共性方差检验，主成分命名检验以及修正措施，解决了 7 个待解决的问题；最后，结合实例详细阐述了主成分显著性检验的应用步骤。

报告人简介：林海明，广东财经大学教授，广东省现场统计学会常务理事，广东省应用统计学会副理事长。研究方向为多元统计模型与应用。

高维数据分析专场报告 6

报告时间：6 月 8 日上午　报告地点：图书馆多功能厅

报告题目：An automatic multi-scale test for serial correlation of high-dimensional time series

报告人：刘梦雅

摘要：This article proposes an automatic multi-scale test for detecting serial correlation of high-dimensional time series (HDTS) from the perspective of time-frequency analysis. Three theoretical tools fuel the construction and implementation of the test, including the L_2-norm, the maximum overlap discrete wavelet transform (MODWT) and a Bayesian Information Criterion (BIC)-like penalty term. The three accomplish, in turn, data dimensionality reduction, scale-by-scale correlation time-frequency analysis and data-driven selection of the optimal number of scales, thus completing the implementation of our testing principle. Under some mild conditions, the limiting null distribution of the proposed test is proved to be chi-square with degrees of freedom 1, and the testing power of our test is analyzed in theory under a general alternative hypothesis.

报告人简介：刘梦雅，华中师范大学数学与统计学院副教授，主要研究方向是时间序列分析，研究成果发表在 JBES、JTSA 等期刊，主持国自然青年项目和湖北省青年项目各一项，入选湖北省“楚天学子”人才计划。

高维数据分析专场报告 7

报告时间：6 月 8 日上午　报告地点：图书馆多功能厅

报告题目： Robust Transfer Learning for High-dimensional Sparse and Low-rank Matrix Regression

报告人： 孙高明

摘要： In this paper, we study the robust transfer learning problem for sparse and low-rank matrix regression models, aiming to enhance the learning performance on the target data by leveraging information from transferable source data. We propose a sparse and low-rank Huber regression method under heavy-tailed noise in high-dimensional settings, based on a nonconvex optimization problem with rank and sparsity constraints. When the set of transferable sources is known, we develop a transfer learning algorithm on the sparse and low-rank Huber regression, and establish non-asymptotic estimation error bounds under both Frobenius and nuclear norms. When the set of transferable sources is unknown, we further develop a data-driven transferable source detection approach to identify informative sources, and establish the consistency of the detection under the sparse low-rank Huber regression transfer learning setting. Numerical simulations and real data analysis demonstrate the effectiveness of the proposed algorithms.

报告人简介： 孙高明，博士，河南大学数学与统计学院讲师，硕士生导师, 华东师范大学博士，研究方向为统计机器学习， 曾在 Statistics、Metrika、Stat.等期刊发表论文多篇。

高维数据分析专场报告 8

报告时间：6 月 8 日上午　报告地点：图书馆多功能厅

报告题目：Testing High-dimensional White Noise Based on Modified Portmanteaus Tests

报告人：周泽人

摘要：For high-dimensional time series, we propose a new test to detect white noise that is not necessarily assumed to be independent and identically distributed. The test can be viewed as a modified portmanteau test in high dimensions, and the critical value of the test statistic is approximated by a multiplier bootstrap method. We provide asymptotic properties of our test under the null hypothesis. The usefulness of our tests is demonstrated by simulations and one real example, particularly for detecting dense alternatives.

报告人简介：周泽人，首都经济贸易大学硕士生导师。主要研究方向：时间序列分析，假设检验。

高维数据分析专场报告 9

报告时间：6 月 8 日上午　报告地点：图书馆多功能厅

报告题目：On the rate of convergence in the CLT for LSS of large-dimensional sample covariance matrices

报告人：崔健

摘要：The convergence rate is a nature probability problem after the establishment of CLT. A direct issue when applying the CLT is controlling the empirical size given sample size, which requires convergence rates to quantify the discrepancy between the true and limiting distributions is a matrix with independent and identically distributed real or complex entries, nonrandom Hermitian nonnegative definite matrix with spectral norm uniformly bounded in. In this paper, based on Stein's method, we establish that if as , the rate of the normalized linear spectral statistics of the sample covariance matrices converging in distribution to the standard normal distribution is for any fixed.

报告人简介：崔健，东北师范大学。

非参数统计专场报告 1

报告时间：6 月 8 日上午　报告地点：图书馆一会议室

报告题目：Conformal Prediction with Missing Data

报告人：唐炎林

摘要：Conformal prediction is a distribution-free method for uncertainty quantification that ensures finite sample guarantee. However, its validity relies on the assumption of data exchangeability. In this talk, I will introduce conformal prediction approaches tailored for missing responses and nonignorable missing data. To provide an asymptotic conditional coverage guarantee for a given subject, we propose constructing prediction regions by establishing the highest posterior density region of the target. This method is more accurate under complex error distributions, such as asymmetric and multimodal distributions, making it beneficial for personalized and heterogeneous scenarios. I will present some numerical results to illustrate their effectiveness.

报告人简介：唐炎林，华东师范大学统计学院教授，博士生导师，统计学系主任；入选国家“万人计划”青年拔尖人才、上海市浦江人才计划。2012 年 1 月博士毕业于复旦大学统计系，同年 5 月加入同济大学，2019 年 1 月加入华东师范大学。主要研究方向为分位数回归、共形预测、高维异质性数据统计推断，主持多项国家自然科学基金、上海市自然科学基金，担任 SCI 期刊 Statistica Sinica、Journal of the Korean Statistical Society 的编委。在 Biometrika、JRSSB、PNAS、Biometrics 等发表论文 40 余篇。

非参数统计专场报告 2

报告时间：6 月 8 日上午　报告地点：图书馆一会议室

报告题目：A simple and powerful method for compoite null testing with application to mediation analysis

报告人：刘耀午

摘要：Large-scale mediation analysis has received increasing interest in recent years, especially in genome-wide epigenetic studies. The statistical problem in large-scale mediation analysis concerns testing composite null hypotheses in the context of large-scale multiple testing. The classical Sobel’s and joint significance tests are overly conservative and therefore are underpowered in practice. In this work, we propose a testing method for large-scale composite null hypothesis testing to properly control the type I error and hence improve the testing power. Our method is simple and essentially only requires counting the number of observed test statistics in a certain region. Non-asymptotic theories are established under weak assumptions and indicate that the proposed method controls the type I error well and is powerful. Extensive simulation studies confirm our non-asymptotic theories and show that the proposed method controls the type I error in all settings and has strong power. A data analysis on DNA methylation is also presented to illustrate our method.

报告人简介：刘耀午，西南财经大学统计学院教授。他的研究兴趣包括统计遗传学，大规模假设检验，全基因组关联性分析等。他的多项研究成果发表于 JASA，JRSSB, American Journal of Human G 和 enetics 等统计学和遗传学知名期刊。

非参数统计专场报告 3

报告时间：6 月 8 日上午　报告地点：图书馆一会议室

报告题目：A New Single-index Model in Alzheimer’s Disease Studies

报告人：赵彦勇

摘要：Alzheimer’s disease (AD) is the major cause of dementia in the elderly, and investigations on the impact of risk factors on neurocognitive performance are crucial in preventative treatment. While existing statistical regression models, such as single-index models, have proven effective tools for uncovering the relationship between the neurocognitive scores and covariates of interest such as demographic information, clinical variables, and neuroimaging features, limited research has explored scenarios where jump discontinuities exist in the regression patterns and the covariates are unobservable but measured with errors, which are common in real applications. To address these challenges, we propose a single-index measurement error jump regression model (SMEJRM) that can handle both jump discontinuities and measurement errors in image covariates introduced by different image processing software. This development is motivated by data from 168 patients in the Alzheimer’s Disease Neuroimaging Initiative. We establish both the estimation procedure and the corresponding asymptotic results. Simulation studies are conducted to evaluate the finite sample performance of our SMEJRM and the estimation procedure. The real application reveals that jump discontinuities do exist in the relationship between neurocognitive scores and some covariates of interest in this study.

报告人简介：赵彦勇，南京审计大学统计与数据科学学院教授，博士生导师；入选江苏省“333 高层次人才培养工程”。2016 年 6 月博士毕业于东南大学数学系，同年 7 月加入南京审计大学。主要研究方向为半参数模型、高维统计推断、分布式推断、复杂数据统计建模，主持 2 项国家自然科学基金、4 项省部级基金（含 1 项重点项目）。在中国科学:数学、统计研究、JMLR、JMVA 等发表论文 40 余篇，相关成果获全国商业科技进步奖二等奖、江苏省统计科研优秀成果奖二等奖等，担任中国现场统计研究会生存分析分会常务理事，中国现场统计研究会理事等。

非参数统计专场报告 4

报告时间：6 月 8 日上午　报告地点：图书馆一会议室

报告题目：Imbalanced regression and its sparse feature selection

报告人：付光辉

摘要：Imbalanced regression has emerged as a novel research direction in machine learning, distinct from classical imbalanced classification. Imbalanced classification addresses the issue of uneven distributions of discrete class labels; while imbalanced regression focuses on predicting continuous target variables that exhibit severe skewness or long-tailed distributions. Such issues are particularly prevalent in scientific fields like medical diagnosis (e.g., predicting indicators of rare diseases) and meteorological forecasting (e.g., extreme weather events), where accurate prediction of rare values holds higher significance. Feature selection is a crucial step in imbalanced regression in high-dimensional settings. However, conventional (sparse) feature selection exhibits inherent limitations when handling imbalanced regression problems: the dominance of common samples often leads to neglect of features critical for predicting rare values. To address this challenge, we propose the SerEnet method which innovatively combines Squared Error-Relevance with respect to a cutoff t (SER_t) with sparse regularization. Through a dual optimization, SerEnet achieves two goals: it optimizes prediction errors for rare instances by SER_t; and it strengthens the selection of highly discriminative features through L1 sparse regularization constraints. Experimental results demonstrate that compared to traditional methods, SerEnet not only more accurately identifies key features but also enhances prediction accuracy for rare values across multiple real-world datasets, while maintaining overall prediction stability.

报告人简介：付光辉，男，现为昆明理工大学教授，博士生导师，云南省“兴滇英才支持计划”青年人才，曼彻斯特大学访问学者。现主要从事生物医学大数据（如高原心血管疾病和茶区头颈部肿瘤数据）的统计机器学习与不平衡学习研究。主持国家自然科学基金 3 项，主持云南省基础研究计划重点项目和面上项目各 1 项。以第一或通讯作者发表相关学术论文近 30 篇，获云南省自然科学三等奖 1 次。

非参数统计专场报告 5

报告时间：6 月 8 日上午　报告地点：图书馆一会议室

报告题目：Penalized weighted GEEs for high-dimensional longitudinal data with informative cluster size

报告人：蒋学军

摘要：High-dimensional longitudinal data has become increasingly common in recent studies. Penalized generalized estimating equations (GEEs) are frequently employed to model such data. However, the desirable properties of the GEE method break down when the outcome of interest is related to the cluster size, a phenomenon known as informative cluster size. In this article, we address this issue by formulating the effect of informative cluster size and proposing a novel weighted GEE approach to mitigate its adverse impact. We demonstrate that the penalized weighted GEE approach achieves consistency in both model selection and estimation. Theoretically, we establish that the proposed penalized weighted GEE estimator is asymptotically equivalent to the oracle estimator, when we know the true model in advance. This finding indicates that the penalized weighted GEE approach retains the excellent properties of the GEE method and is robust to informative cluster sizes, thereby extending the applicability of the GEE method to more complex situations. Simulations and a real data application demonstrate that the penalized weighted GEE have outstanding performance over existing alternative methods.

报告人简介：蒋学军，南方科技大学统计与数据科学系教学副主任，研究员、博士生导师，于 2009 年博士毕业于香港中文大学统计系，2009-2010 在港中文从事博士后研究，2010-2013 任中南财经政法大学副教授，2013 年 07 月加入南方科技大学，入选深圳市海外高层次人才孔雀计划（2016），曾获南方科技大学杰出教学奖（2018），深圳市优秀教师（2018），主持和完成国家（广东省）自然科学基金、深圳市基础研究面上项目等 10 余项。其主要研究方向包括分位数回归、变量选择、假设检验、高维统计推断，金融统计与计量，迁移学习等，已在 Biometrika, Bernoulli, Statistica Sinica, Econometrics Journal，Science China-Mathematics 等国际一流统计学及计量经济学期刊上发表 SCI&SSCI 论文近 60 篇，授权专利 2 项及出版教材一部。目前担任 Statistic and Its Interface 副主编，国内学会任职主要有中国现场统计研究会-教育统计与管理分会副理事长、多元分析应用专业委员会秘书长等。

非参数统计专场报告 6

报告时间：6 月 8 日上午　报告地点：图书馆一会议室

报告题目：Augmented Localized Conformal Testing with Application to Novelty Detection

报告人：任好洁

摘要：The enhancement of test performance by emphasizing the effect of localized regions around the test sample has attracted significant attention. However, such localization inevitably reduces the effective sample size for inference. From the perspective of conformal inference, this paper proposes Augmented Localized Conformal p-values (ALCP) by recalibrating an augmented conditional distribution estimator. This estimator efficiently captures local information by fully integrating both calibration and test data into its kernel estimation. We examine that these p-values with augmented estimators are valid in finite sample and could improve testing performance. We then introduce a novel algorithm that applies these augmented p-values to novelty detection and incorporates a conditional calibration technique. Additionally, a data-driven bandwidth selection method is provided that ensures finite-sample false discovery rate control while improving testing power. Both simulated and real data experiments demonstrate the advantages of our proposed ALCP.

报告人简介：任好洁，上海交通大学数学科学学院长聘教轨副教授，18 年博士毕业于南开大学，随后在宾州州立大学从事博士后研究。她的研究方向包括预测推断、统计异常探查、在线学习与监控、选择推断等。在 JASA，Biometrika 等杂志和机器学习顶会 ICML，NeurIPS 上发表学术论文 20 余篇。

非参数统计专场报告 7

报告时间：6 月 8 日上午　报告地点：图书馆一会议室

报告题目：Testing hierarchical covariance structures under quadratic subspace

报告人：梁宇莉

摘要：In this paper the Rao score and likelihood ratio tests for hypothesis related to the covariance structure of multivariate models are studied. For the hypothesis related to hierarchical dependence such that the intercalss and intraclass correlations belong to the space of commutative quadratic structures the Rao score and likelihood ratio test statistics are derived as well as the exact distributions of the likelihood ratio test is determined. Simulation studies are conducted and presented results are applied to real data example.

报告人简介：梁宇莉，女，博士，广西师范大学数学与统计学院，曾任瑞典林奈大学经济学院经济与统计系助理教授，硕士生导师，瑞典厄勒布鲁大学商学院统计系教职和瑞典国家统计局方法论部门资深统计师(senior methodologist)，于乌普萨拉大学获统计学硕士学位，后于斯德哥尔摩大学获统计学博士学位。主要研究兴趣为含协方差结构的多元模型和高维数据分析，已在包括 TEST，Biometrical Journal, Statistical Paper 在内的国际 SCI 期刊上发表论文数篇。

非参数统计专场报告 8

报告时间：6 月 8 日上午　报告地点：图书馆一会议室

报告题目： Conformal Prediction with Cellwise Outliers: A Detect-then-Impute Approach

报告人： 包亚杰

摘要： Conformal prediction is a powerful tool for constructing prediction intervals for black-box models, providing a finite sample coverage guarantee for exchangeable data. However, this exchangeability is compromised when some entries of the test feature are contaminated, such as in the case of cellwise outliers. To address this issue, this paper introduces a novel framework called detect-then-impute conformal prediction. This framework first employs an outlier detection procedure on the test feature and then utilizes an imputation method to fill in those cells identified as outliers. To quantify the uncertainty in the processed test feature, we adaptively apply the detection and imputation procedures to the calibration set, thereby constructing exchangeable features for the conformal prediction interval of the test label. We develop two practical algorithms, PDI-CP and JDI-CP, and provide a distribution-free coverage analysis under some commonly used detection and imputation procedures. Notably, JDI-CP achieves a finite sample 1-2α coverage guarantee. Numerical experiments on both synthetic and real datasets demonstrate that our proposed algorithms exhibit robust coverage properties and comparable efficiency to the oracle baseline.

报告人简介： 包亚杰，南开大学助理研究员，博士毕业于上海交通大学，研究兴趣包括共形推断、联邦学习等，主要成果发表在 Biometrika、CSDA、EJS 等统计学期刊以及 ICML、ICLR、NuerIPS 等机器学习会议上。

非参数统计专场报告 9

报告时间：6 月 8 日上午　报告地点：图书馆一会议室

报告题目：Unveiling Compositional Data with the Symmetric Chi-Square Kernel

报告人：王柔琳

摘 要： Traditional compositional data analysis methods, such as log-ratio transformations and Aitchison geometry, have significant limitations: they require that all components be non-zero and that the number of components remain fixed. This paper introduces a novel nonparametric technique based on the symmetric Chi-Square Kernel, which allows for components to be zero and accommodates an infinite number of components. As an initial application, we employ this technique to address the two-sample testing problem for compositional data. We provide the statistical properties of the proposed nonparametric test, including test power, consistency of the test statistic, and its asymptotic distribution. Simulation studies demonstrate that the proposed test outperforms traditional methods in terms of sensitivity and computational efficiency. Applications to real-world data underscore its practical relevance, particularly in detecting subtle differences between distributions in high-dimensional settings.

报告人简介：王柔琳，华东师范大学博士后，博士毕业于中国科学技术大学管理学院统计学。研究方向聚焦非参数统计推断与假设检验理论等领域。在 Annals of Statistics，Communications in Mathematics and Statistics，Canadian Journal of Statistics 等期刊上发表多篇论文。

函数型数据分析专场报告 1

报告时间：6 月 8 日上午　报告地点：图书馆二会议室

报告题目：Online estimation for functional data

报告人：杨莹

摘要：Functional data analysis has attracted considerable interest and is facing new challenges, one of which is the increasingly available data in a streaming manner. In this article we develop an online nonparametric method to dynamically update the estimates of mean and covariance functions for functional data. The kernel-type estimates can be decomposed into two sufficient statistics depending on the data-driven band- widths. We propose to approximate the future optimal bandwidths by a sequence of dynamically changing candidates and combine the corresponding statistics across blocks to form the updated estimation. The proposed online method is easy to compute based on the stored sufficient statistics and the current data block. We derive the asymptotic normality and, more importantly, the relative efficiency lower bounds of the online estimates of mean and covariance functions. This provides insight into the relationship between estimation accuracy and computational cost driven by the length of candidate bandwidth sequence.

报告人简介：杨莹，复旦大学应用数学中心青年研究员，相辉青年学者，2022 年于北京大学统计科学中心取得博士学位，2024 年在中国科学院数学与系统科学研究院完成博士后工作。研究方向为复杂数据的实时动态算法和基于因果推断的政策评估问题。相关成果发表于 JASA、JRSSB 等期刊。

函数型数据分析专场报告 2

报告时间：6 月 8 日上午　报告地点：图书馆二会议室

报告题目：Global tests for detecting change in mean vector functions of multivariate functional data with repeated observations

报告人：邱志平

摘要：In many scientific and technological fields, multivariate functional data are often repeatedly observed under varying conditions over time. A fundamental question is whether the mean vector function remains consistently equal throughout the entire period. This paper introduces two novel global testing statistics that leverage integration technique to address this issue. The asymptotic distributions of the proposed test statistics under the null hypothesis are derived, and their root-n consistency is established. Simulation studies are conducted to evaluate the numerical performance of the proposed tests, which are further illustrated through an analysis of publicly available EEG motion data.

报告人简介：邱志平，福建师范大学教授。2014 年 6 月毕业于上海财经大学统计学专业，获经济学博士学位。2016 年 11 月至 2018 年 6 月于美国埃默里（Emory）大学生物统计与生物信息系从博士后研究工作。主要从事生存数据分析、函数型数据分析、半参数与非参数模型、统计机器学习等方面研究工作。在国际统计学知名杂志《Statistica Sinica》、《Scandinavian Journal of Statistics》、《Journal of Multivariate Analysis》等期刊上发表 20 多篇学术论文。主持国家自然科学基金面上项目 1 项，教育部人文社科一般项目 1 项和福建省自然科学基金面上项目多项。曾获上海财经大学研究生优秀博士学位论文、首届全国高校数学微课程教学设计竞赛华东赛区一等奖等奖项

函数型数据分析专场报告 3

报告时间：6 月 8 日上午　报告地点：图书馆二会议室

报告题目：Adaptive functional change point detection

汇报人：朱学虎

摘要:：This paper develops an adaptive dimension reduction methodology specifically tailored for functional change point detection, with a particular focus on preserving critical information. The core of our methodology is to reduce infinite-dimensional functional data into finite-dimensional vector or scalar representations. This innovative strategy effectively addresses the prevalent issue of information loss commonly associated with traditional dimension reduction techniques, such as Functional Principal Component Analysis (FPCA). Building upon this foundation, we propose the MPULSE criterion to establish a comprehensive functional change point detection framework. This framework seamlessly integrates dimension reduction with precise change point identification, creating a unified analytical pipeline. The robustness and efficacy of our method are validated through extensive simulation studies and real-world data applications, consistently demonstrating superior performance compared to existing methods.

报告人简介：朱学虎，博士，西安交通大学教授，博士生导师，主要从事统计学习、高维数据分析及应用统计等领域的研究，截至目前，已发表学术论文 30 余篇，包括 Journal of the American Statistical Association、Journal of Business & Economic Statistics、Science China Mathematics 等期刊以及应用领域权威期刊 IEEE Transactions on Geoscience and Remote Sensing、计算机顶级会议 NeurIPS 等。先后主持科技部重点研发计划子课题、国家自然科学基金面上项目、国家自然科学基金重点项目课题等，入选陕西省高校青年杰出人才支持计划和仲英青年学者等荣誉。

函数型数据分析专场报告 4

报告时间：6 月 8 日上午　报告地点：图书馆二会议室

报告题目：Estimation of volatility functionals with time-varying price staleness

报告人：朱海斌

摘要：It is widely observed that zero returns frequently occur in the real financial market, especially for high-frequency data. Such a phenomenon is known as price staleness. In this paper, we are interested in investigating the effects of price staleness on the estimation of volatility functionals. Some specific quantities include integrated volatility, covariance, beta, correlation, and many others, all of which play critical roles in financial econometrics. We first estimate the spot covariance by the empirical realized estimator, and our analyses show that price staleness brings in bias. After correcting the bias, we propose a global estimator of volatility functionals by applying the plug-in technique. Under suitable conditions, consistency and central limit theorem results are established for the estimator. Moreover, the estimation under the further presence of jumps and microstructure noise is also discussed. We conduct extensive simulation studies to assess the finite sample performance of our estimator and verify our theoretical results. As an empirical application, by using some real high-frequency data, we demonstrate that our staleness-corrected estimator of beta is more suitable for evaluating intraday ETF tracking performance than the traditional one without considering the price staleness.

报告人简介：朱海斌，中国澳门大学数学博士，现任暨南大学统计与数据科学系助理教授、第三届常任轨教师。主要研究领域为高频金融计量、金融机器学习、生物信息学等。先后在 Journal of Business & Economic Statistics、Journal of Empirical Finance 等期刊上发表论文。

函数型数据分析专场报告 5

报告时间：6 月 8 日上午　报告地点：图书馆二会议室

报告题目：Testing and measuring the conditional mean (in)dependence for functional data by Martingale difference-angle divergence

报告人：赖廷煜

摘要：We proposed a new nonparametric method to test and measure conditional mean (in)dependence for functional data. This new metric has some appealing properties: it is nonnegative and equals to zero if and only if the conditional mean independence holds; it is invariant under linear transformations of the predictor; it doesn’t require the moment condition for the predictor variable. Based on this measure, two test procedures for conditional mean independence are proposed for functional data. One uses a wild bootstrap while the other uses the limiting standard normal distribution. The tests are consistent and perform well in finite sample simulations. We further propose some requirements for a reasonable conditional mean dependence measure and demonstrate that our metric has those properties. A real data example is introduced to illustrate the application of the proposed method.

报告人简介：赖廷煜，男，博士，广西师范大学数学与统计学院讲师，研究方向为函数型数据分析。在《Journal of Multivariate Analysis》，《Statistica Sinica》，《Computational Statistics & Data Analysis》等期刊上发表多篇文章。

函数型数据分析专场报告 6

报告时间：6 月 8 日上午　报告地点：图书馆二会议室

报告题目：函数型自回归条件异方差积分模型估计及应用

报告人：彭选华

摘要：日内新息显著影响投资决策，并可能引发价格异常波动的叠加现象。现有的函数型自回归条件异方差模型（FARCH）主要依赖历史数据预测日内波动，但未能充分考虑实时新息的冲击。本文引入日内新息，构建函数型自回归条件异方差积分模型（FARCHint），分析了模型平稳解的存在性，并提出基于基函数展开的最小二乘估计方法，进而用于预测价格日内波动的马太特征。仿真算例表明，FARCHint 在预测性能上优于传统的 FARCH 模型。该模型用于预测 ChatGPT 概念股和比特币的日内收益率、波动率以及风险价值，验证了应用的有效性和科学合理性，实现日内风险价值的动态测度，从而为金融风险管理提供了新模型。

报告人简介：彭选华，男，重庆人，理学硕士，管理学博士，西南政法大学经济学院副教授，硕士生导师； 全国金融量化分析与计算专业委员会会员，中国灾害防御协会风险分析专业委员会会员，中国量化投资学会会员，重庆市工业与应用数学学会理事，全国大学生数学建模竞赛（重庆赛区）委员会专家；主要从事金融大数据科学、金融量化分析、加密资产风险管理等经济科学前沿基础与应用研究。多次应邀在中国数量经济技术经济年会、经济学年会、金融学年会、中国现场统计研究会高维数据分会、重庆工业与应用数学年会等做专场学术报告，已独立出版 1 部学术专著《金融风险价值量化分析》，已在《数量经济技术经济研究》、《系统工程理论与实践》、《数理统计与管理》、《工业管理最优化学报》等刊物发表 20 余篇专业技术论文；长期为系统科学与复杂性学报(Journal of Systems Science & Complexity)、应用经济学报（Applied Economics）、经济学模型(Economic modeling)等英文刊物审稿；主持 5 项重庆市自然科学基金面上项目、教委科技项目、教委人文社科项目和重庆社科规划项目，参研多项国家科学基金，有承担基础与前沿应用研究的较强业务技能。

函数型数据分析专场报告 7

报告时间：6 月 8 日上午　报告地点：图书馆二会议室

报告题目：函数型 AR-ARCHint-X 模型在比特币市场风险度量中的应用研究

报告人：王佩淋

摘要：在加密资产高频交易背景下，比特币市场因独特的价格形成机制与多重影响因素的非线性交互作用，对传统金融模型构成重大的挑战。针对现有研究在刻画高频数据动态特征与多因子协同效应方面的不足，本文创新性地构建函数型 AR-ARCHint-X 模型，该模型结合了函数型自回归和条件异方差模型，综合考虑了成交量对价格均值和条件异方差的影响，以及日内动态因素对市场波动的作用。随后，本文对协变量的弱平稳性进行研究，并采用最小二乘法进行参数估计，通过优化联合估计方法提升模型精度。在实证分析中，本文以比特币市场的高频数据为对象，采用函数型典型相关性分析（FCCA）方法探讨比特币市场中的价量关系，结果表明，成交量函数与收益率波动在时频域存在显著耦合效应。最后，对构建的函数型 AR-ARCHint-X 模型参数估计的结果进行分析，解释市场收益率、波动率和成交量之间的动态传导机制。结果显示，该模型在拟合能力上表现良好，能够有效地捕捉市场的动态特征，本研究为加密资产市场的风险管理提供了兼具理论创新与实践价值的分析框架。

报告人简介：王佩淋，西南政法大学，硕士研究生。

函数型数据分析专场报告 8

报告时间：6 月 8 日上午　报告地点：图书馆二会议室

报告题目：Analyzing Functional Data with a Mixture of Covariance Structures Using a Curve-Based Sampling Scheme

报告人：俞易安

摘要：Motivated by distinct walking patterns in real-world free-living gait data, this study proposes an innovative curve-based sampling scheme for the analysis of functional data characterized by a mixture of covariance structures. Traditional approaches often fail to adequately capture inherent complexities arising from heterogeneous covariance patterns across distinct subsets of the data. We introduce a unified Bayesian framework that integrates a nonlinear regression function with a continuous-time hidden Markov model, enabling the identification and utilization of varying covariance structures. One of the key contributions is the development of a computationally efficient curve-based sampling scheme for hidden state estimation, addressing the sampling complexities associated with high-dimensional, conditionally dependent data. This study details the Bayesian inference procedure, examines the asymptotic properties to ensure the structural consistency of the model, and demonstrates its effectiveness through simulated and real-world examples.

报告人简介：俞易安，南方科技大学，研究生。

复杂网络分析专场报告 1

报告时间：6 月 8 日上午　报告地点：图书馆三会议室

报告题目：BASIC: Bipartite Assisted Spectral-clustering for Identifying Communities in Large-scale Networks

报告人：刘婧媛

摘要：Community detection, which focuses on recovering the group structure within networks, is a crucial and fundamental task in network analysis. However, the detection process can be quite challenging and unstable when community signals are weak. Motivated by a newly collected large-scale academic network dataset from the Web of Science, which includes multi-layer network information, we propose a Bipartite Assisted Spectral-clustering approach for Identifying Communities (BASIC), which incorporates the bipartite network information into the community structure learning of the primary network. The accuracy and stability enhancement of BASIC is validated theoretically on the basis of the degree-corrected stochastic block model framework, as well as numerically through extensive simulation studies. We rigorously study the convergence rate of BASIC even under weak signal scenarios and prove that BASIC yields a tighter upper error bound than that based on the primary network information alone. We utilize the proposed BASIC method to analyze the newly collected large-scale academic network dataset from statistical papers. During the author collaboration network structure learning, we incorporate the bipartite network information from author-paper, author-institution, and author-region relationships. From both statistical and interpretative perspectives, these bipartite networks greatly aid in identifying communities within the primary collaboration network.

报告人简介：刘婧媛，厦门大学经济学院统计学与数据科学系教授、博士生导师，国家级青年人才计划入选者，厦门大学南强卓越教学名师，南强青年拔尖人才 A 类。主要从事复杂数据的统计方法研究、网络数据建模与推断、统计基因学等领域的工作，在国际权威学术期刊发表论文 30 余篇，担任 AOAS 等权威期刊编委，入选福建省杰出青年科研人才计划。担任全国现场统计研究会教育统计与管理分会副理事长、统计交叉科学研究分会副理事长等。

复杂网络分析专场报告 2

报告时间：6 月 8 日上午　报告地点：图书馆三会议室

报告题目：Individual-centered partial information in social networks

报告人：韩潇

摘要：In statistical network analysis, we often assume either the full network is available or multiple subgraphs can be sampled to estimate various global properties of the network. However, in a real social network, people frequently make decisions based on their local view of the network alone. Here, we consider a partial information framework that characterizes the local network centered at a given individual by path length L and gives rise to a partial adjacency matrix. Under L = 2, we focus on the problem of (global) community detection using the popular stochastic block model (SBM) and its degree-corrected variant (DCSBM). We derive general properties of the eigenvalues and eigenvectors from the signal term of the partial adjacency matrix and propose new spectral-based community detection algorithms

that achieve consistency under appropriate conditions. Our analysis also allows us to propose a new centrality measure that assesses the importance of an individual’s partial information in determining global community structure. Using simulated and real networks, we demonstrate the performance of our algorithms and compare our centrality measure with other popular alternatives to show it captures unique nodal information. Our results illustrate that the partial information framework enables us to compare the viewpoints of different individuals regarding the global structure.

报告人简介：韩潇，中国科学技术大学管理学院特任教授，研究方向为大维随机矩阵；高维统计推断，入选国家创新人才计划青年项目，主持青年基金项目与面上基金项目各一项。

复杂网络分析专场报告 3

报告时间：6 月 8 日上午　报告地点：图书馆三会议室

报告题目：Community detection in weighted networks via the profile-pseudo likelihood method

报告人：刘秉辉

摘要：In this paper, we consider the issue of community detection in weighted networks. Most methods addressing this issue, particularly those statistical approaches based on likelihood optimization, suffer from a notable drawback: the necessity to specify in advance the particular form of the distribution of edge weights conditional on the community labels. This requirement dictates that algorithms based on likelihood optimization must be custom-tailored exclusively to the specific form of distribution, which exhibits significant limitations in practical applications where the form of distribution is unknown. To address this limitation, we propose two novel methods based on the expectation profile-pseudo likelihood maximization, for community detection in both undirected and directed weighted networks, which are applicable to various types of weighted networks and independent of the specific form of the conditional distribution of the edge weights. In theory, we establish weak and strong consistency, respectively, of the resulting community label estimations within the sub-exponential family, and then establish the convergence of the proposed algorithms. In simulation studies, we demonstrate significant advantages of the proposed methods across a wide range of conditional distributions and parameter settings, both in terms of community detection accuracy and computational efficiency. In practical applications, we showcase the applicability of the proposed methods on three real-world weighted networks.

报告人简介：刘秉辉，东北师范大学数学与统计学院，教授、统计系主任；主要研究方向为统计机器学习和网络数据分析；在统计学及相关领域发表论文三十余篇，部分成果发表在 JASA、AOS、AOAS、AIJ、JMLR、JOE、JBES 等上；主持国家自然科学基金多项，入选国家级青年人才计划，担任中国现场统计研究会因果推断分会副理事长、中国现场统计研究会统计交叉科学研究分会副理事长等。

复杂网络分析专场报告 4

报告时间：6 月 8 日上午　报告地点：图书馆三会议室

报告题目：Academic Literature Recommendation in Large-scale Citation Networks Enhanced by Large Language Models

报告人：潘蕊

摘要：Literature recommendation is essential for researchers to find relevant articles in an ever-growing academic field. However, traditional methods often struggle due to data limitations and methodological challenges. In this work, we construct a large citation network and propose a hybrid recommendation framework for scientific article recommendation. Specifically, the citation network contains 190,381 articles from 70 journals, covering statistics, econometrics, and computer science, spanning from 1981 to 2022. The recommendation mechanism integrates network-based citation patterns with content-based semantic similarities. To enhance content-based recommendations, we employ text-embedding-3-small model of OpenAI to generate an embedding vector for the abstract of each article. The model has two key advantages: computational efficiency and embedding stability during incremental updates, which is crucial for handling dynamic academic databases. Additionally, the recommendation mechanism is designed to allow users to adjust weights according to their preferences, providing flexibility and personalization. Extensive experiments have been conducted to verify the effectiveness of our approach. In summary, our work not only provides a complete data system for building and analyzing citation networks, but also introduces a practical recommendation method that helps researchers navigate the growing volume of academic literature, making it easier to find the most relevant and influential articles in the era of information overload.

报告人简介：潘蕊，中央财经大学统计与数学学院教授、博士生导师，中央财经大学龙马学者青年学者。主要研究领域为网络结构数据的统计建模、时空数据的统计分析等。在 Annals of Statistics、Journal of the American Statistical Association、Journal of Business & Economic Statistics 等期刊发表论文 30 余篇。著有中文专著《数据思维实践》、《网络结构数据分析与应用》。主持国家自然科学基金项目、全国统计科学研究项目。

复杂网络分析报告 5

报告时间：6 月 8 日上午　报告地点：图书馆三会议室

题目：Heterogeneous Autoregressive Model for Symmetric Matrix-valued Time Series

报告人：刘广应

摘要：In the fields of economics and finance, there is a vast amount of matrix-valued time series data,some of which are symmetric matrices, such as the realized covariance matrix obtained from financial high-frequency data. In this paper, we construct the Matrix Heterogeneous AutoRegressive (MHAR) model for symmetric matrix-valued time series to describe their dynamics. The coefficient matrices are estimated using the Alternating Direction Method of Multipliers (ADMM), and the asymptotic properties of these estimators are established. Simulation results confirm these asymptotic properties. We then use our MHAR model to predict future realized covariance matrices and apply these predicted matrix values in portfolio management. Compared to the other models considered for predicting realized covariance, both simulation studies and real-data analysis demonstrate that our proposed model offers optimal predictive accuracy and investment performance.

报告人简介：刘广应，南京审计大学教授、博士生导师，校润泽学者。研究方向：金融高频数据、应用统计、金融风险管理、深度学习，2011 年复旦大学统计学系概率论与数理统计专业理学博士。2019.7-2022.6 主持一项国家社会科学基金一般项目基于深度学习的金融高频数据波动率预测及其应用研究，2023.12-2028.12 国家社科重大项目课题，中国式现代化的统计监测评价问题研究子课题负责人。2016.1-2018.12 主持国家自然科学基金青年项目，含微观噪音半鞅的预平均统计量渐近理论及其在金融高频数据的应用。

复杂网络分析专场报告 6

报告时间：6 月 8 日上午　报告地点：图书馆三会议室

报告题目：Fractal Nature of the Solution to Stochastic Heat Equation Driven by Fractional-Colored Noise

报告人：王文胜

摘要：In this paper, the solution to a stochastic heat equation driven by fractional-colored noise is studied. This solution is a random function of time and space. For a fixed point in space, the resulting random function of time has exact, dimension dependent, global continuity moduli and laws of the iterated logarithm (LILs). It is obtained that the set of fast points at which LILs fail of this process occur infinitely often is a random fractal, whose size is evaluated by its Hausdorff dimension. These points of this process are everywhere dense with power of the continuum almost surely, and their hitting probabilities are determined by the packing dimension dimP (E) of the target set E.

报告人简介：王文胜，杭州电子科技大学经济学院教授、博导。主要研究方向是概率极限理论、数理金融。长期以来主要从事概率统计和数理金融等领域的教学和科研工作。兼任中国工程概率统计学会常务理事、浙江省现场统计研究会副理事长、浙江省学位委员会经济类专业学位研究生教育指导委员会成员。先后主持和参与国家自然科学基金项目多项、主持和参与教育部、上海市、浙江省等省部级科研项目多项，在国内外学术期刊发表学术论文 110 余篇，出版教材 1 部。获浙江省高校优秀科研成果三等奖 1 项，浙江省自然科学优秀论文二等奖 3 项。入选教育部新世纪优秀人才支持计划和浙江省中青年学科带头人培养计划。

复杂网络分析专场报告 7

报告时间：6 月 8 日上午　报告地点：图书馆三会议室

报告题目：时间序列与复杂网络相互表征及应用

报告人：赵毅

摘要：本报告围绕数据科学两类典型范式（时间序列和复杂网络）相互转换、表征开展研究。现实系统的复杂性决定了仅凭单一范式我们只能获得对象的特定属性。为解决这一问题，人们更多关注它们的内在关系。受先前工作的启发，我们提出了利用拟等距同构映射的方法，建立时间序列和复杂网络等价转换的理论，并且将该转换方法扩展到多元时间序列与多层网络等价转换的情况。在此基础上，将多元时间序列传递熵的方法拓展到多层网络范式提出层间熵的概念，证明二者是严格等价的，从而给出一般意义下多元时间序列与多层复杂网络相互表征提供理论依据。进一步，量化了两者之间的对应关系，以期从这二者的双重视角全面理解复杂系统。

报告人简介：赵毅, 哈尔滨工业大学（深圳）理学院, 教授，博士生导师，深圳市鹏城学者特聘教授、英国数学与应用学会会士。主要研究方向包括应用动力系统、非线性时间序列分析、复杂系统和数据科学。他担任广东省复杂动力系统理论与智能计算创新团队负责人和广东省大学数学公共课程教学团队负责人。近五年主持完成国家自然科学基金面上项目 1 项，广东省自然科学基金面上项目 2 项，深圳市孔雀计划技术创新项目 1 项，参与主持国家重点研发计划云计算和大数据重点专项子课题 1 项。目前主持国家自然科学基金面上项目 1 项。他获得深圳市第四届教育教学科研优秀成果特等奖和第五届教学成果二等奖，并荣获深圳市十佳师德标兵和哈工大（深圳）教学名师奖。

流数据专场报告 1

报告时间：6 月 8 日上午　报告地点：图书馆五会议室

报告题目：TestAgent: Strategic A/B test via two-armed agent

报告人：严晓东

摘要：Detecting a minor average treatment effect is a major challenge in large-scale applications, where even minimal improvements can have a significant economic impact. Traditional methods, reliant on normal distribution-based or expanded statistics, often fail to identify such minor effects because of their inability to handle small discrepancies with sufficient sensitivity. This work leverages a counterfactual outcome framework and proposes a maximum probability-driven two-armed bandit process by weighting the mean volatility statistic, which controls Type I error. The implementation of permutation methods further enhances the robustness and efficacy. The established strategic central limit theorem (SCLT) demonstrates that our approach yields a more concentrated distribution under the null hypothesis and a less concentrated one under the alternative hypothesis, greatly improving statistical power. The experimental results indicate a significant improvement in the A/B testing, highlighting the potential to reduce experimental costs while maintaining high statistical power.

报告人简介：严晓东，西安交通大学数学与统计学院教授，博士生导师，入选国家级青年人才项目和校内青拔 A 类支持计划，滴滴盖亚学者，研究成果发表在统计学著名期刊 JRSSB、AOS、JASA 和计量经济著名期刊 JOE 以及人工智能顶级会议 AAAI、AISTAT 等。

流数据专场报告 2

报告时间：6 月 8 日上午　报告地点：图书馆五会议室

报告题目：Mean-preserving rounding integer-valued ARMA models

报告人：朱复康

摘要： In the past four decades, research on count time series has made significant progress, but research on Z-valued time series is relatively rare. Existing Z-valued models are mainly of autoregressive structure, where the use of the rounding operator is very natural. Because of the discontinuity of the rounding operator, the formulation of the corresponding model identifiability conditions and the computation of parameter estimators need special attention. It is also difficult to derive closed-form formulae for crucial stochastic properties. We rediscover a stochastic rounding operator, referred to as mean-preserving rounding, which overcomes the above drawbacks. Then, a novel class of Z-valued ARMA models based on the new operator is proposed, and the existence of stationary solutions of the models is established. Stochastic properties including closed-form formulae for (conditional) moments, autocorrelation function, and conditional distributions are obtained. The advantages of our novel model class compared to existing ones are demonstrated. In particular, our model construction avoids identifiability issues such that maximum likelihood estimation is possible. A simulation study is provided, and the appealing performance of the new models is shown by several real-world data sets.

报告人简介：朱复康，吉林大学数学学院教授、博士生导师，吉林国家应用数学中心副主任、院长助理、概率统计与数据科学系主任。2008 年博士毕业，2013 年破格晋升教授，2021 年任唐敖庆领军教授。主要从事时间序列分析和金融统计的研究，已经在 Annals of Applied Statistics、Journal of Business & Economic Statistics、Statistica Sinica、Scandinavian Journal of Statistics、Journal of Time Series Analysis、中国科学-数学等期刊上发表论文多篇，主持国家自然科学基金面上项目 3 项和青年基金 1 项。曾获教育部自然科学奖二等奖、吉林省科学技术奖二等奖、吉林省享受省政府津贴专家、长春市有突出贡献专家等奖励，连续两年(2023-2024)入选美国斯坦福大学发布的全球前 2%顶尖科学家榜单。现任中国现场统计研究会、全国工业统计学教学研究会、中国数学会概率统计分会等学会的理事或常务理事。现任 SCI 期刊 Statistical Papers、Journal of Statistical Computation and Simulation 的 Associate Editor，是 JASA、JRSSB、JBES、AoAS 等 80 余个 SCI 期刊的匿名审稿人。指导的研究生 1 人获得吉林省优秀博士学位论文、3 人获得吉林省优秀硕士学位论文。

流数据专场报告 3

报告时间：6 月 8 日上午　报告地点：图书馆五会议室

报告题目：On bivariate self-exciting hysteretic integer-valued autoregressive processes

报告人：杨凯

摘要：This paper introduces a bivariate hysteretic integer-valued autoregressive (INAR) process driven by a bivariate Poisson innovation. It deals well with the buffered or hysteretic characteristics of the data. Model properties such as stationarity and ergodicity are studied in detail. Parameter estimation problem is also well address via methods of two-step conditional least squares (CLS) and conditional maximum likelihood (CML). The boundary parameters are estimated via triangular grid searching algorithm. The estimation effect is verified through simulations based on three scenarios. Finally, the new model is applied to the offence counts in New South Wales (NSW), Australia.

报告人简介：杨凯，副教授，博士生导师，现任长春工业大学数学与统计学院副院长，吉林省高层次人才，曾赴日本岛根大学学术访问，兼任全国工业统计教学研究会理事，中国现场统计研究会大数据统计分会理事，中国现场统计研究会多元分析应用专业委员会理事，中国现场统计研究会经济与金融统计分会理事，吉林省工业与应用数学学会理事等。主要研究领域包括时间序列分析、高维数据分析、区间数据分析、矩阵数据分析等。主持国家自然科学基金面上项目、青年基金项目各 1 项，吉林省自然科学基金面上项目 1 项，吉林省博士后基金择优资助项目 1 项，吉林省产业关键核心技术攻关项目 1 项（子课题负责人），吉林省教育厅科学研究项目 1 项，横向科研项目 2 项，以主要参加人身份参与国家级、省部级科研项目 5 项。以第一作者、通讯作者身份在 Applied Mathematical Modelling, CSDA，TEST 等杂志发表 SCI/SSCI 论文 30 余篇，其中二区以上论文 6 篇，ESI 高被引论文 2 篇，ESI 热点论文 1 篇，发表 CSSCI 论文 2 篇。主持省级研究生精品课建设项目 1 项，先后荣获 2023 和 2024 年吉林省优秀硕士学位论文指导教师，第四届全国高校数学微课程教学设计竞赛全国一等奖，全国应用统计专业学位研究生教育教学成果二等奖，指导学生参加学科竞赛获得国家级奖项 10 余项。

流数据专场报告 4

报告时间：6 月 8 日上午　报告地点：图书馆五会议室

报告题目：几类时间序列模型在空气质量监测中的应用研究

报告人：李聪

摘要：随着城市化的加速，空气质量监测对于公共健康和环境管理变得越来越重要。本文探讨了用整数值时间序列模型捕捉空气质量指标之间的复杂相依关系的方法，并根据模型性质提出新的控制图，以监控过程均值和方差的变化。本研究还利用我国多个地区某一时期的 AQI 和 CO 等时间序列数据，展示了这些方法在空气质量评估方面的应用，该研究可为决策制定和政策规划提供支持。

报告人简介：李聪，吉林大学副教授，博士生导师。一直从事时间序列建模、预测与监控等方面的研究，主持国家自然科学基金和国家社会科学基金等项目，参与“长春市市长公开电话”等政校合作项目，现任中国现场统计研究会资源与环境统计分会理事、大数据统计分会理事。

流数据专场报告 5

报告时间：6 月 8 日上午　报告地点：图书馆五会议室

报告题目：Statistical inference for smoothed quantile regression with streaming data

报告人：谢锦瀚

摘要： In this talk, we introduce a fully online updating method for statistical inference in smoothed quantile regression with streaming data. For low-dimensional data, we present an incremental updating algorithm to obtain the smoothed quantile regression estimator with the streaming data set. The proposed estimator allows us to construct asymptotically exact statistical inference procedures. Within the realm of high-dimensional data, we develop an online debiased lasso procedure to accommodate the special sparse structure of streaming data. The proposed online debiased approach is updated with only the current data and summary statistics of historical data and corrects an approximation error term from online updating with streaming data. Furthermore, theoretical results such as estimation consistency and asymptotic normality are established to justify its validity in both settings. Our findings are supported by simulation studies and applications.

报告人简介：谢锦瀚，云南大学数学与统计学院副教授，先后在香港中文大学统计系、加拿大阿尔伯塔大学数学与统计科学系、美国北卡罗来纳大学教堂山分校生物统计系从事博士后研究工作。其研究成果已在多个国际顶刊上发表，涵盖统计学、人工智能和计量经济学领域，包括 JASA、JMLR, JoE 和 JBES 等。

流数据专场报告 6

报告时间：6 月 8 日上午　报告地点：图书馆五会议室

报告题目：Extreme conditional expectile estimation for heavy-tailed ARMA-GARCH models

报告人：马耀兰

摘要：Expectiles have recently received considerable attention due to their coherence as a measure of tail risk. Estimating conditional expectiles, particularly at both intermediate and extreme levels, is crucial in quantitative risk management. This article proposes an ARMA-GARCH model that accommodates fewer finite moments and assumes innovations follow a Pareto-type tail distribution. We apply the two-step self-weighted procedure of He, Peng, Zhang and Zhao (2022) to forecast extreme conditional expectiles. Employing extreme value theory, we estimate the extremal conditional expectile and develop a unified asymptotic theory for the conditional expectile estimator, which incorporates both intermediate and extreme scenarios. Our Monte Carlo simulations demonstrate that the proposed approach significantly improves coverage probabilities compared to other competing methods across various contexts, especially in extreme scenarios. Finally, an empirical application to the daily negative log-returns of 14 financial asset indices shows that our method consistently outperforms forecasts from Hoga (2022) and the Peaks Over Threshold (POT) approach during both normal and crisis periods.

报告人简介：马耀兰，北方民族大学数学与信息科学学院副教授、硕士生导师。2018 年博士毕业于浙江大学数学科学学院。主要从事时间序列分析和非参数统计的研究，在统计学和计量经济学期刊 Econometric Theory、Statistica Sinica、Journal of Time Series Analysis、Journal of Financial Econometrics 等期刊上发表论文多篇，先后主持国家自然科学基金项目 1 项、宁夏自然科学基金 3 项。

流数据专场报告 7

报告时间：6 月 8 日上午　报告地点：图书馆五会议室

报告题目：Quantile regression model with uniform subsampling for streaming data

报告人：晏振

摘 要： In this paper, we propose an online updating algorithm that combines subsampling to solve the quantile regression model's coefficients, addressing two fundamental characteristics of current streaming data: large volume and the inability to store all data simultaneously. To verify the effectiveness of the proposed algorithm, we compare its absolute performance gap and computational cost with those of the Oracle quantile regression model and the updatable quantile regression model under four different error scenarios. The numerical simulation results demonstrate that this algorithm is well-suited to the characteristics of streaming data, does not require accessing historical data, has low memory consumption, high efficiency, and shows no significant increase in error compared to traditional quantile regression coefficient estimation methods. In addition to its practical performance, we also theoretically prove the Oracle property of the algorithm.

报告人简介：晏振，男，博士，广西师范大学副教授，研究领域主要包括函数型数据、分位回归、经验似然等方面。先后主持国家自然科学基金青年项目 1 项、广西自然科学基金项目 2 项，参与国家级科研课题 4 项。在《数理统计与管理》和《Communication in Statistics-Theory and Method》等期刊发表学术论文 10 余篇。曾获北京市第十二届统计科研优秀成果评比优秀课题论文一等奖。兼任中国商业统计学会理事、全国工业统计学教学研究会民族统计与数据科学分会常务理事、全国工业统计学教学研究会青年统计学家协会理事、广西统计学会理事、广西数学会理事等职务。

流数据专场报告 8

报告时间：6 月 8 日上午　报告地点：图书馆五会议室

报告题目：A Trinomial Difference Autoregressive Process for the Bounded Z-valued Time Series

报告人：陈花萍

摘要：This article tackles the modeling challenge of bounded Z-valued time series by proposing a novel trinomial difference autoregressive process. This process not only maintains the autocorrelation structure presenting in the classical binomial GARCH model, but also facilitates the analysis of bounded Z-valued time series with negative or positive correlation. We verify the stationarity and ergodicity of the couple process (comprising both the observed process and its conditional mean process) while also presenting several stochastic properties. We further discuss the conditional maximum likelihood estimation and establish their asymptotic properties. The effectiveness of these estimators is assessed through simulation studies, followed by the application of the proposed models to two real datasets.

报告人简介：陈花萍博士，河南大学数学与统计学院讲师，硕士生导师, 研究方向:时间序列分析，主持河南省博士后启动基金、河南省自然科学基金-青年项目，参与国家自然科学基金-面上项目. 先后在 Journal of Time Series Analysis、Journal of Statistical Planning and Inference、AStA Advances in Statistical Analysis、Metrika、Journal of Statistical Computation and Simulation、Stat 等国际权威期刊发表论文 18 篇，其中发表在 Stat 上的论文被评为年度高被引论文，发表在 JTSA 上的论文获得开封市第二十届自然科学优秀学术成果奖一等奖。

流数据专场报告 9

报告时间：6 月 8 日上午　报告地点：图书馆五会议室

报告题目：Mixed causal-noncausal count process

报告人：裴健

摘要： Recently, Gouriéroux and Lu (2021) introduced a class of (Markov) noncausal count processes. These processes are obtained by time-reverting a standard count process (such as INAR(1)), but have quite different dynamic properties. In particular, they can feature bubble-type phenomena, which are epochs of steady increase, followed by sharp decreases. This is in contrast to usual INAR and INGARCH type models, which only feature “reverse bubbles”, that are epochs of sharp increase followed by steady decreases. In practice, however, in many datasets, sudden jumps and crashes are rare, while it is more frequent to observe epochs of steady increase or decrease. This paper introduces the mixed causal-noncausal integer-valued autoregressive (m-INAR(1,1)) process, obtained by superposing a causal and a noncausal INAR(1) process sharing the same sequence of error terms. We show that this process inherits some key properties from the noncausal INAR(1), such as the bi-modality of the predictive distribution and the irreversibility of the dynamics, while at the same time allowing different accumulation and burst speeds for the bubble. We propose a GMM estimation method, investigate its finite sample performance, develop testing procedures, and apply the methodology to stock transaction data.

报告人简介：裴健，北京建筑大学理学院，讲师，博士，2015 年 9 月-2024 年 6 月期间，本硕博连读于吉林大学，于 2019 年 6 月取得学士学位，2024 年 6 月取得博士学位。研究方向为时间序列分析，一直从事时间序列模型的推广、推断和应用，目前共发表学术论文 6 篇，其中以第一作者在《Test》、《Statistical Papers》、《Journal of Applied Statistics》等 SCI 期刊发表论文 4 篇，主持项目 1 项。

统计推断专场一报告 1

报告时间：6 月 8 日上午　报告地点：厚德楼 2017

报告题目：CTE2: Conditional Tail Expectation Treatment Effect at Extreme Levels

报告人：黎德元

摘要：The subgroup average treatment effect, which evaluates effects for predefined sub populations, is often more interpretable than the pointwise conditional average treatment effect. In this paper, we focus on the extreme subpopulations and propose the conditional tail expectation treatment effect (CTE2). By combining inverse probability weighting technique with extreme value theory, we develop estimators for CTE2 at extreme levels. We establish the asymptotic normality of our estimators through a novel multivariate causal tail empirical process framework. The finite-sample performance of the estimator is demonstrated through simulation studies. We further illustrate the utility of our approach by applying it to estimate the CTE2 of college education on wages.

报告人简介：黎德元，复旦大学管理学院教授，中国现场统计研究会教育统计与管理分会副理事长、常务理事。主要研究极值统计、风险管理、分布式统计推断、分位数回归。

统计推断专场一报告 2

报告时间：6 月 8 日上午　报告地点：厚德楼 2017

报告题目：Rank-Based Nonparametric Methods for Evaluating Treatment Equivalence in AB/BA Crossover Trial Designs

报告人：邱世芳

摘要：The primary objective of a randomized controlled trial is to evaluate the equivalence between two treatments. Treatment effects are quantified by the probability that a subject receiving one treatment achieves a better outcome (or a 'win') compared to a subject receiving the other treatment in this article. Without requiring distributional assumptions, nonparametric tests based on confidence intervals for win probability have been developed to analyze the effects of two treatments in standard two-period AB/BA crossover trials. Although crossover designs help control inter-subject variability, unmeasured carryover effects can compromise the validity of the results. Therefore, the win probability of the carryover effect is first employed to examine the absence of sequence order effects between treatments. Five simultaneous confidence interval methods-normal-distribution-based, Student's t-distribution-based, adjusted Student's t-distribution-based, logit-transformation-based, and arsinh-transformation-based-have been developed to evaluate the two win probabilities and assess the effectiveness of the two treatments. Simulation results indicate that all confidence interval methods perform satisfactorily in terms of empirical coverage probability, empirical interval width, and left and right non-coverage probabilities. To illustrate the proposed methods, an AB/BA crossover trial involving patients with chronic obstructive pulmonary disease (COPD) is utilized.

报告人简介：邱世芳，教授，硕士生导师。重庆市数学学会副理事长，中国现场统计研究会多元分析分会常务理事、资源与环境分会理事，全国工业统计学教学研究会理事。重庆市巴南区“菁英计划”人才。1989.09-1993.07，获得重庆师范学院数学教育专业学士学位；2005.09-2008.07，获得云南大学概率论与数理统计专业硕士学位（中国科学院数学与系统科学研究院联合培养）；2008.09-2011.01，获得云南大学概率论与数理统计专业博士学位（中国科学院数学与系统科学研究院联合培养）。2011.01-至今，重庆理工大学理学院教师。主持了国家自然科学基金项目、重庆市自然科学基金创新发展联合基金项目等项目的研究，发表学术论文 40 余篇，被 SCI 收录近 30 篇。

统计推断专场一报告 3

报告时间：6 月 8 日上午　报告地点：厚德楼 2017

报告题目： Unified inference framework and regression adjustment for stratified randomized experiments with ordinal outcomes

报告人： 杨玥含

摘要： Assessing the causal effects of interventions on ordinal outcomes is essential in many biological experiments and behavioral studies. Using the framework of potential outcomes, causal effects are defined by comparing outcomes under treatment and control conditions. However, interpreting average causal effects can be challenging for ordinal outcomes, which are ranked or ordered. To address this, we develop a unified inference framework for stratified randomized experiments, focusing on the relative treatment effect—the difference between the probabilities of the treatment being beneficial versus harmful. We derive sharp bounds for this effect based on the marginal distribution of potential outcomes, accommodating both regression adjusted and unadjusted effects. Covariate adjustment is leveraged to refine these bounds and improve estimation precision. The proposed estimators are validated through numerical simulations and empirical analyses.

报告人简介： 杨玥含，中央财经大学统计与数学学院教授，2014 年博士毕业于北京大学数学科学学院，主要研究方向包括因果推断、迁移学习、复杂数据建模等。文章发表于 JASA、Biometrika、JBES、Pattern Recognition 等期刊。

统计推断专场一报告 4

报告时间：6 月 8 日上午　报告地点：厚德楼 2017

报告题目： Methods-of-moment inference for GLM and observational studies in proportional asymptotics

报告人： 刘林

摘要： In this talk, we discuss how to conduct inference properly for GLMs and causal estimands in observational studies when the data dimension is proportional to the sample size. Recent works have been exclusively about performing a debias surgery on regularized MLE, such as the Lasso, ridge and Firth regularizations. However, we discovered that under the commonly used assumptions in the literature, such as Gaussian designs, the high-dimensional GLM can be succintly represented by a system of two nonlinear equations. By directly solving this nonlinear system, we can obtain sharp theoretical guarantees over the linear and quadratic forms of the regression coefficients. We also discuss the implications for estimating causal estimands that arise from observational studies. This is based on joint work with Xingyu Chen and Rajarshi Mukherjee.

报告人简介： 刘林，目前于上海交通大学自然科学研究院、数学科学学院与上海交大-耶鲁生物统计与数据科学联合中心任职副教授，研究兴趣为因果推断、半参数理论、生物信息学、以及因果推断方法在生物医学中的应用。

统计推断专场一报告 5

报告时间：6 月 8 日上午　报告地点：厚德楼 2017

报告题目： Design and modeling for order-of-addition factorial experiments with interaction effects

报告人： 陈建斌

摘要： In scientific research and industrial engineering, the response of an experiment is often determined by both the levels of factors and their order of addition. Such experiments are referred to as order-of-addition factorial (OofAF) experiments. The interactions between factor levels and their order of addition frequently have a substantial impact on the response. For instance, in automotive manufacturing, painting components prior to welding may result in heat-induced damage due to welding temperatures, thereby increasing the defect rate. To address these interactions in OofAF experiments, we introduce a class of conditional main effects (CMEs) to capture interaction effects by measuring the conditional effect of a pairwise ordering factor at a fixed level of another factor. Leveraging these CMEs, we propose a statistically interpretable model, called the conditional main effects pairwise ordering model, for analyzing OofAF experiments. Under this model, we develop a systematic construction method to generate a D-optimal fractional OofAF design, which reduces experimental costs while enabling efficient estimation of main effects and CMEs. Comparative case studies are presented to illustrate that this proposed model outperforms existing approaches in predictive accuracy when analyzing OofAF experiments with significant interaction effects.

报告人简介： 陈建斌 ，北京理工大学数学与统计学院副研究员，美国普渡大学博士后，2020 年毕业于南开大学数据与统计科学学院，获博士学位。曾先后前往宾州州立大学统计系，佐治亚大学进行学术访问。目前研究方向包括：试验设计，添加次序试验，统计学习理论，可靠性分析，机器学习等。先后曾在 Tecnometrics, European Journal of Operational Research 等期刊上发表 20 余篇重要期刊论文。

统计推断专场一报告 6

报告时间：6 月 8 日上午　报告地点：厚德楼 2017

报告题目：Construction of (nearly) orthogonal symmetric Latin hypercube designs

报告人：杨柳青

摘要：Latin hypercube designs (LHDs) have been widely used as computer experimental designs when a linear model is _tted. The symmetric LHD (SLHD), as a special kind of LHDs, can guarantee that the estimates of second-order e_ects and main e_ects are uncorrelated. In this paper, we propose two methods to construct orthogonal SLHDs (OSLHDs) and nearly orthogonal SLHDs (NOSLHDs). The _rst method can generate new designs based on existing OSLHDs. Some new NOSLHDs with exible run sizes and good nearly orthogonality can be constructed. Moreover, the resulting OSLHDs of the second construction method have better strati_cation properties than existing OSLHDs.

报告人简介：杨柳青，博士，现为中南大学数学与统计学院讲师，主要从事统计试验设计与数据分析的研究，现已在 JRSSB、JASA、Sinica 等统计学重要期刊上正式发表 7 篇论文；主持国家自然科学基金青年项目 1 项、湖南省自然科学基金青年项目 1 项。

统计推断专场一报告 7

报告时间：6 月 8 日上午　报告地点：厚德楼 2017

报告题目：Robust matrix auto-regressions for matrix-valued time series with heavy-tailed distributions and digital finance analysis in Yangtze River Economic Belt

报告人：Hu Xuemei（胡雪梅）

摘要：In this paper we firstly replace an ordinary least square loss (LSL) by a general robust loss including Huber loss (HL) and exponential squared Huber loss (ESHL, a hybrid of squared loss for relatively small errors and exponential squared loss for relatively large ones), and construct a fire-new robust matrix auto-regressive model(RMAR) to efficiently handle matrix-valued time series with heavy-tailed distributions such as multivariate t distribution and multivariate exponential distribution. Secondly, we review an alternating iterative least squares estimation procedure for MAR under LSL(ILSE-LSL), propose a fire-new alternating iterative estimation (IE) method including IE-HL and IE-ESHL for RMAR under HL and ESHL, respectively. Monte Carlo simulation results show that except for ILSE-LSL, the proposed IE-HL and IE-ESHL can efficiently deal with heavy-tailed distributions, and further verify that IE-ESHL performs best, IE-HL performs second, ILSE-LSL performs third, and LSE-LSL(least squares estimation under least squares loss) for VAR (vector autoregression) performs worse. Finally, we apply the relatively optimal IE-ESHL method to analyze digital finance situations from 2011 to 2021 in the Yangtze River Economic Belt, and predict the digital finance indexes from 2022 to 2026.

报告人简介：胡雪梅，重庆工商大学数学与统计学院教授，博士生导师，伦敦政治经济学院国家公派访问学者,重庆经开区经济运行局挂职副局长,中南大学理学博士，中科院数学与系统科学研究院博士后， 第四批重庆市学术技术带头人(统计学)，第五批重庆市高等学校优秀人才支持计划人选，重庆市“统计学”研究生导师团队负责人，《随机过程》市级一流线下课程负责人。目前已对半变系数模型的统计推断、半参数模型的经验似然、随机扩散方程的稳健推断、高维数据模型的统计学习等展开了系统研究，在 IEEE Transaction on Information Theory、Journal of Multivariate Analysis、Expert Systems with Applications、Statistical Papers、North American Journal of Economics and Finance、Journal of Forecasting 等学术期刊上发表论文 50 多篇，其中 SCI/SSCI 收录 30 篇，主持完成国家自然科学基金和教育部人文社科等省部级以上科研项目 10 余项， 获得重庆市科学技术奖二等奖，主编出版学术专著 2 部《高维统计模型的估计理论与模型识别》和《高维数据模型的统计学习方法与预测精度评估》，参编英文专著 1 部。

统计推断专场一报告 8

报告时间：6 月 8 日上午　报告地点：厚德楼 2017

题目：High-Dimensional Extreme Quantile Regression

报告人：唐一苇

摘要：The estimation of conditional quantiles at extreme tails is of great interest in numerous applications. Various methods that integrate regression analysis with an extrapolation strategy derived from extreme value theory have been proposed to estimate extreme conditional quantiles in scenarios with a fixed number of covariates. However, these methods become less effective in high-dimensional settings, where the number of covariates grows with the sample size. In this article, we develop new estimation methods tailored for extreme conditional quantiles with high-dimensional covariates. We establish the asymptotic properties of the proposed estimators and demonstrate their superior performance through simulation studies, particularly in scenarios of growing dimension and high dimension where existing methods may fail. Furthermore, the analysis of auto insurance data validates the efficacy of our methods in estimating extreme conditional insurance claims and selecting important variables.

报告人简介：唐一苇，复旦大学，博士研究生。

统计推断专场一报告 9

报告时间：6 月 8 日上午　报告地点：厚德楼 2017

题目：CoVaR under Asymptotic Independence

报告人：王昭文

摘要：There has been increasing interest in measuring systemic risk since the global financial crisis of 2007-2009. Accurate assessment of systemic risk would enable regulators to introduce suitable policies to mitigate the risk as well as allow individual institutions to monitor their vulnerability to market movements. One popular measure of systemic risk is the conditional value-at-risk (CoVaR), proposed by Adrian and Brunnermeier (2016). While Nolde et al. (2022) discusses the estimation of CoVaR under asymptotic dependence, it does not address the case of asymptotic independence, which is receiv- ing increasing attention in quantitative risk management. In this work, we develop a semi-parametric methodology to estimate CoVaR for asymptotically independent pairs within the framework of bivariate extreme value theory. We use parametric modelling of the bivariate extremal structure to address data sparsity in the joint tail regions and prove consistency and asymptotic normality of the proposed estimator. The performance of the estimator is illustrated via simulation studies and a real data example.

报告人简介：王昭文，复旦大学，博士研究生。

统计推断专场二报告 1

报告时间：6 月 8 日上午　报告地点：厚德楼 3019

报 告 题 目 ： Pseudo-Likelihood Ratio Screening based on Network Data with Applications

报告人：黄丹阳

摘要：Social network platforms today generate vast amounts of data, including network structures and a large number of user-defined tags, which reflect users' interests. The dimensionality of these personalized tags can be ultra-high, posing challenges for model analysis in targeted preference analysis. Traditional categorical feature screening methods overlook the network structure, which can lead to incorrect feature set and suboptimal prediction accuracy. This study focuses on feature screening for network-involved preference analysis based on ultra-high-dimensional categorical tags. We introduce the concepts of self-related features and network-related features, defined as those directly related to the response and those related to the network structure, respectively. We then propose a pseudo-likelihood ratio feature screening procedure that identifies both types of features. Theoretical properties of this procedure under different scenarios are thoroughly investigated. Extensive simulations and real data analysis on Sina Weibo validate our findings.

报告人简介：黄丹阳，中国人民大学教授，博士生导师，北京大数据协会副会长。专注研究大规模数据的高效计算与统计理论，包括高维数据降维分析，复杂网络建模，分布式计算。并关注统计理论在中小企业数字化中的应用。主持国家自然科学基金面上项目 2 项，国家自然科学基金青年项目 1 项。参与国家社会科学基金重大项目 1 项，国家重点研发计划“社会治理与智慧社会科技支撑”重点专项 1 项。北京市科协青年人才托举工程，北京市优秀人才培养资助。

统计推断专场二报告 2

报告时间：6 月 8 日上午　报告地点：厚德楼 3019

报告题目：Granular ball T-spherical fuzzy set for multi-criteria group decision-making in breast cancer treatmen

报告人：鲜思东

摘要：随着 DeepSeek 的强势入场，我们进入了国产大模型带来的颠覆性创新时代。面对新一轮的机遇与挑战，如何进行统计学习、知识创新以及智能决策？本报告在介绍基于数据驱动的知识发现与和基于 x -语言模型的智能决策研究工作的基础上。重点引入了一种新颖的框架，称为粒球 T-球形模糊集（GBT-SFS），以应对癌症治疗方案选择中专家面对复杂不确定信息及歧义性的挑战，给出复杂认知信息环境下，统计学习与复杂系统的融合框架及其智能决策过程等。

报告人简介：鲜思东，男，中共党员，重庆邮电大学二级教授，硕士生导师，重庆市高校“双带头人”教师党支部书记工作室书记，重庆市人工智能+智能决策学科群执行负责人，统计学学科负责人，复杂系统智能分析与决策重庆市高校重点实验副主任，兼职中国商业统计学会常务理事，CSIAM 大数据与人工智能专业委员会委员，中国现场统计学会大数据统计分会常务理事，中国商业统计学会数据科学与商业智能分会常务理事，重庆市特种设备安全标准委员会副主任委员，重庆软件产业专家顾问委员会委员等。全国大学生数学建模竞赛优秀指导教师，中国研究生数学建模竞赛全国先进个人等。研究兴趣为：大数据统计、模式识别、智能决策、模糊时间序列分析等。主持省部级以上项目 10 余项（含重点 5 项），横向研究课题 10 余项；主持重庆市一流本科课程 2 门。在 IEEE TFS、INS、ESWA、AIIM、JOTA、IJIS、ASOC、EAAI 等重要期刊上发表 SCI 论文 60 余篇，SCI 一二区 30 余篇。出版专著 1 部，主编教材 3 部。

统计推断专场二报告 3

报告时间：6 月 8 日上午　报告地点：厚德楼 3019

报告题目：Determining The Number of Factors in Two-Way Factor Model of High-Dimensional Matrix-Variate Time Series: A White-Noise based Method for Serial Correlation Models

报告人：夏强

摘要：In this paper, we study a new two-way factor model for high-dimensional matrix-variate time series. To estimate the number of factors in this two-way factor model, we decompose the series into two parts: one being a non-weakly correlated series and the other being a weakly correlated noise. By comparing the difference between two series, we can construct white-noise based signal statistics to determine the number of factors in row loading matrix (column loading matrix).

Furthermore, to mitigate the negative impact on the accuracy of the estimation, which is caused by the interaction between the row loading matrix and the column loading matrix, we propose a transformation so that the transformed model only contains the row loading matrix (column loading matrix). We define sequences of ratios of two test statistics as signal statistics to determine the number of factors and derive the consistence of the estimation.

We implement the numerical studies to examine the performance of the new methods.

报告人简介：夏强，男，教授，博士，毕业于中国人民大学统计学院，接受和发表的论文 20 多篇，其中 SCI（或 SSCI）论文 14 篇，科学出版社出版专著 1 部，国际斯伯林格出版专著的一个章节，人民邮电出版社出版教材 1 部，主持国家和省部级基金 6 项。

统计推断专场二报告 4

报告时间：6 月 8 日上午　报告地点：厚德楼 3019

报告题目：Weighted Estimation Method of High Dimension Portfolio Allocation for Time-Varying Stock Market

报告人：黄磊

摘要：The primary goal of portfolio allocation is to achieve higher returns while controlling risk. In the background of large portfolios, the classical Mean-Variance Portfolio (MVP) model faces high-dimensional challenges, which severely reduces the accuracy of the estimation. Meanwhile, policy changes and market structure shifts may contribute to time-varying stock market to some extent. To address the poor performance of existing methods based on the classical MVP model in the time-varying stock market, this article proposes a Time-Varying Weighted Large Portfolio (TVWLP) method. It is grounded in the factor model and introduces various weight functions to adjust the estimated factor loading matrix, and timely adjust the mean vector and covariance matrix. Furthermore, we establish the asymptotic properties of the timely adjusted method under some regular conditions. We validate the effectiveness of TVWLP through some simulations. Finally, in empirical studies, we demonstrate the superiority of TVWLP with both high and low time-varying effects.

报告人简介：黄磊，2015 年博士毕业于新加坡国立大学，现任职于西南交通大学数学学院统计系，副教授，研究生导师，科研兴趣包括半参数回归模型、时间序列分析、生物统计。在国内外学术期刊 The Annals of Statistics, Journal of Business & Economic Statistics, Energy, Statistics in Medicine, Journal of Nonparametric Statistics, Computational Statistics & Data Analysis 等发表论文 20 余篇。主持国家自然科学基金青年项目 1 项、省级项目 3 项、校级项目 6 项。2020 年获四川省数学会首届应用数学奖一等奖。2021 年入选中国现场统计研究会理事。

统计推断专场二报告 5

报告时间：6 月 8 日上午　报告地点：厚德楼 3019

报告题目： Multi-Period Portfolio Allocation: A One-Shot Stochastic Optimization Approach

报告人： 许晓菲

摘要： This paper presents a multi-period factor investing framework utilizing a one-shot stochastic optimization approach, which efficiently addresses the complexities of multi-period decision-making under uncertainty. Traditional portfolio optimization models, based on single-period analyses and static beta estimations, often fail to capture the evolving dynamics of financial markets. By extending the classical mean-variance optimization to a multi-period setting and employing a one-shot stochastic optimization methodology, we optimize portfolio weights across multiple periods simultaneously. This approach leverages Gaussian Processes and the long-term utility function to model the dynamic and stochastic nature of asset returns, ensuring theoretical consistency and favorable sample complexity properties. Our empirical analysis, which includes a comparison of dynamic versus static beta estimations, demonstrates that the proposed one-shot optimization delivers superior performance metrics such as Compound Annual Growth Rate and Sharpe Ratio. Additionally, portfolios optimized using machine learning-driven forecasts exhibit superior risk-return trade-offs and resilience across diverse market conditions compared to conventional benchmark strategies like Buy and Hold and Minimum Variance portfolios. These findings provide valuable insights for portfolio managers aiming to optimize risk-adjusted returns over multiple investment horizons, thereby contributing an efficient and theoretically grounded methodology to the finance literature and advancing the practice of long-term factor-based investing.

报告人简介： 许晓菲，武汉大学数学与统计学院特聘副研究员。博士毕业于新加坡国立大学。主要研究方向为函数型时间序列分析、整值时间序列分析和预测等。在 JBES，AOAS，Statistica Sinica 等国际学术期刊上发表论文数篇，主持/参与国家自然科学基金青年项目和重点项目等。

统计推断专场二报告 6

报告时间：6 月 8 日上午　报告地点：厚德楼 3019

报告题目： Statistical inference of Meta-Elliptical distributions - Based on MSE statistical representation points(MSE-RPs)

报告人： 何平安

摘要： Meta-distributions, constructed based on copulas, effectively capture the underlying correlations between variables, even when their marginal distributions vary significantly. This approach enables more accurate and robust multivariate modeling, particularly in scenarios involving extreme values or heavy-tailed distributions. In this talk, we will use the Mean Squared Error (MSE) criterion to generate MSE representative points (MSE-RPs) for the Meta-Elliptical distribution and compare them with traditional Monte Carlo (MC) and Quasi-Monte Carlo (QMC) methods. Our findings show that MSE-RPs outperform these methods in estimating both the mean and covariance matrix of the Meta-Elliptical distribution, demonstrating greater efficiency. Some new properties of Gaussian copula and meta Gaussian distributions are also investigated.

报告人简介： 何平安，北师香港浸会大学，博士研究生。

统计推断专场二报告 7

报告时间：6 月 8 日上午　报告地点：厚德楼 3019

报告题目：Estimation for single-index varying-coefficient spatial autoregressive model with index covariate measurement errors

报告人：王可

摘要：This paper proposes a single-index varying coefficient spatial autoregressive model which has measurement errors in the index covariates. We combine a local-linear smoother based simulation-extrapolation (SIMEX) algorithm, the estimation equation and the profile maximum likelihood method to estimate our model. Under some regular conditions, we derive the asymptotic properties of the unknown estimators and the link function. Some simulation studies indicate our estimation method performs well. Finally, our method is illustrated with the real dataset of Boston Housing Price.

报告人简介：王可，博士毕业于吉林大学，师从王德辉教授。现为长春工业大学师资博士后，合作导师杨凯副教授。主要研究兴趣包括空间统计，非参数统计，测量误差模型等。主要成果发表于 Spatial Statistics 和 Statistical Papers 等杂志。

统计推断专场二报告 8

报告时间：6 月 8 日上午　报告地点：厚德楼 3019

报告题目： False Discovery Rate Control over Data-driven Subgroups: A Unified Selective Approach

报告人： 霍雨阳

摘要： Multiple testing over subgroups of interest is crucial for uncovering the underlying structure of a problem. In cases where prior knowledge of group partition is unavailable, a data-driven subgrouping algorithm is applied to the data. Controlling the false discovery rate (FDR) in such data-driven subgroups is challenging because the subgrouping process distorts the distribution of test statistics, creates complex dependence within groups, and results in a random number of hypotheses. To overcome these issues, we propose a unified selective approach named SCAR (Swapping-CAlibrating-Referencing), which consists of three key components: a swapping re-subgrouping procedure to construct selective p-values, a calibrating procedure to decouple the within group dependence, and a referencing step to adjust for the randomness in subgroup sizes. We demonstrate that SCAR achieves finite-sample FDR control in a distribution-free and algorithm-agnostic regime. Unlike existing methods based on marginal p-values, our approach leverages selective p-values that fully utilize subgrouping information to significantly boost statistical power. Numerical experiments on both synthetic and real datasets confirm that SCAR not only maintains valid FDR control over data-driven subgroups but also outperforms compared benchmark methods across various subgrouping algorithms and data-generating scenarios.

报告人简介： 霍雨阳，南开大学统计与数据科学学院 2022 级博士生，师从邹长亮教授，研究兴趣为预测推断，多重检验和选择后推断。在顶级期刊 JASA 和 Biometrika，机器学习顶级会议 NeurIPS 上发表多篇论文，现主持博士生基金项目。

统计推断专场三报告 1

报告时间：6 月 8 日上午　报告地点：厚德楼 3020

报告题目：A consistent specification test for Lp-quantile regression

报告人：宋晓军

摘要：In this article, we propose a nonparametric test for the correct specification of parametric Lp-quantile regression over a continuum of quantile levels. The test is based on continuous functions of a residual-marked empirical process. We show that the test is consistent and has nontrivial power against a sequence of local alternatives approaching the null at a parametric rate. Since the limiting distribution of the test statistic depends on the underlying data-generating process in a complicated way, we propose a simple multiplier bootstrap procedure to approximate the critical values. A Monte Carlo study shows that the asymptotic results provide good approximations for small sample sizes.

报告人简介：宋晓军，男，北京大学光华管理学院商务统计与经济计量系副教授，博士生导师，西班牙马德里卡洛斯三世大学经济学博士。主要研究兴趣是理论计量经济学，包括非参数/半参数方法，假设检验和自助法，以及计量经济学的应用等。论文发表在 Annals of Applied Statistics，Biometrics，Econometric Theory，Journal of Applied Econometrics，Journal of Business & Economic Statistics 和 Journal of Econometrics 等国际期刊。主持和参加自然科学基金面上项目和重点项目等。自 2020 年 1 月起担任 Economic Modelling 副主编。

统计推断专场三报告 2

报告时间：6 月 8 日上午　报告地点：厚德楼 3020

报告题目：DyVAESurv: A VAE-enhanced model for dynamic survival analysis

报告人：赵学靖

摘要：Survival analysis plays a crucial role in discovering the relationship between covariates and event times, which is widely applied in individual risk prediction. Traditional methods often rely on static features, results in the difficulty of capturing time dependencies, while dynamic survival models have limitations in capturing heterogeneity among individuals. To address these issues, a deep learning model, integrating original features with latent representations generated by a Variational Autoencoder (VAE), is proposed to better characterize individual survival trajectories. The model employs a dynamic feature extractor that selects the appropriate Transformer or LSTM branch based on the sequence length, thus adapting to different time series data. The model retains the information of the original data upon joint representation, it also utilizes the latent variables of the VAE to effectively capture the heterogeneity among individuals. Experimental results demonstrate that the proposed model achieves superior performance in terms of C-index and IBS, providing a novel solution for dynamic prediction based on survival analysis.

报告人简介：赵学靖，兰州大学数学与统计学院教授，副院长，概率论与数理统计博士生导师，应用统计专业学位硕士研究生导师,兰州大学应用统计硕士专业学位授权点负责人。担任中国工业统计学教学研究会第十届理事会常务理事；甘肃省统计学会第九届理事会副秘书长。研究方向为生存分析与寿命检验，统计机器学习，高维统计。在统计专业期刊发表论文 40 余篇，其中 SCI/EI 收录近 30 篇。主持国家自然科学基金面上项目一项，春晖计划合作交流项目一项，留学回国人员科研启动项目一项，兰州大学高校基本科研项目两项。

统计推断专场三报告 3

报告时间：6 月 8 日上午　报告地点：厚德楼 3020

报告题目： Some new results in option pricing using artificial intelligent algorithms

报告人： 郭精军

摘要： In this talk,we will give some new results in option pricing. Firstly, a dynamic ensemble framework is introduced,which integrates parametric and non-parametric pricing models. Within this framework, we propose a time-varying parametric pricing model optimized using artificial intelligence algorithms. Secondly,establishing the loss function between the trading data and modeled value, the implied volatility at different moments solved using the global optimal double annealing algorithm is found to differ from the generalized autoregressive conditional heteroskedasticity (GARCH) volatility and historical volatility.

报告人简介： 郭精军，男，二级教授，博士，博士生导师，担任甘肃经济发展数量分析研究中心主任和统计与数据科学学院副院长。担任全国工业统计学教学研究会理事等。入选国家留学基金委西部地区人才培养特别项目，入选甘肃省省级重点人才项目。入选甘肃省飞天学者人选。省级一流本科课程和省级教学团队负责人。研究兴趣：随机分析、金融统计与风险管理等。主持国家自然科学基金项目 3 项、省部级项目 5 项。在 Resources Policy、Acta Mathematica Scientia 等重要期刊上发表学术论文 40 余篇。出版学术著作 2 部，出版教材 3 部。2 次获甘肃省优秀硕士论文指导教师称号。获得甘肃省研究生教育优秀导师称号。

统计推断专场三报告 4

报告时间：6 月 8 日上午　报告地点：厚德楼 3020

报告题目： Goodness-of-Fit Tests for High-Dimensional Regression Models via Projections

报告人： 谭发龙

摘要： In this talk, we proposed a new method for testing the goodness of fit for high dimensional generalized linear regression models when the number of covariates may be much larger than the sample size. Most existing model checking methods in the literature does not work for high dimension regression models as they suffer from the curse of dimensionality and rely on the asymptotic linearity and normality of the estimator of the parameters. Our method is based on random projections which largely avoid the “curse of dimensionality”. Further, our test only need the convergence rate of the estimators of the high dimensional parameters and does not rely on the asymptotic expansion or the normality of these estimators. The asymptotic properties of the test statistics are investigated under the null and the local and global alternatives when the number of covariates is much larger than the sample sizes. We further proposed a combination method to enhance the power performance of the tests. Detailed simulation studies and a real data analysis are conducted to illustrate the effectiveness of our methodology.

报告人简介： 谭发龙，湖南大学金融与统计学院教授、博士生导师。主要研究兴趣包括人工智能的数学与统计基础、统计机器学习、高维模型检验、高维经验过程、充分性降维、因果推断与精准医疗等。相关的研究工作发表在 Annals of Statisitcs、Biometrika、CVPR、Statistica Sinica 等国际期刊与会议上，主持和参与国家自然科学基金、省部级项目 10 余项。

统计推断专场三报告 5

报告时间：6 月 8 日上午　报告地点：厚德楼 3020

报告题目：Model selection with uncertainty in estimating optimal dynamic treatment regimes

报告人：王春雨

摘要：Optimal dynamic treatment regimes (DTRs), as a key part of precision medicine, have gained more and more attention recently. To inform clinical decision making, interpretable and parsimonious models for contrast functions are preferred, which inevitably to some extent suffer from misspecification. It is therefore important to properly evaluate the performance of the candidate interpretable models and select one with better approximation to the unknown contrast function. Moreover, as a DTR usually involves multiple decision points, an inaccurate approximation at a latter decision point will then affect its estimation at an earlier point when a backward induction algorithm is applied. This paper aims to perform model selection for contrast functions in the context of learning optimal DTRs from observed data. Note that the relative performance of candidate models may heavily depend on the sample size when, for example, the comparison is made between parametric models and tree-based models. Therefore, instead of investigating the limit behavior of each candidate model and developing methods to asymptotically select the correct one, we focus on the finite sample performance of each model and attempt to perform model selection under a given sample size. In other words, the estimate of interest is sample-size-specific. To this end, we adopt the counterfactual cross-validation metric and propose a novel method to estimate the variance of the metric. Supplementing the cross-validation metric with its estimated variance allows us to characterize the uncertainty on model selection under a given sample size and facilitates conducting a hypothesis test over a preferred model structure. Simulation studies are provided to demonstrate (i) the performance of our proposed variance estimator and (ii) the improvement achieved by incorporating model selection for contrast functions in estimating optimal DTRs. We finally apply our method to the analysis of Sequential Treatment Alternatives to Relieve Depression (STAR*D) data.

报告人简介：王春雨，浙江大学数学科学学院助理教授，百人计划研究员，剑桥大学 MRC Biostatistics Unit 博士后。

统计推断专场三报告 6

报告时间：6 月 8 日上午　报告地点：厚德楼 3020

报告题目：Statistical Inference for High-dimensional Marltrix-variate Factor Models with Missing Observations

报告人：张永霞

摘要：This paper develops an inferential theory for high-dimensional matrix-variate factor models with missing observations. We propose an easy-to-use all-purpose method that involves two straightforward steps. First, we perform principal component analysis on two re-weighted covariance matrices to obtain the row and columrloadings. Second, we utilize these loadings along with the matrix-variate data to derive the factors. We develop an inferential theory that establishes the consistency and the rate of convergence under general conditions and missing patterns. The simulation results demonstrate the adequacy of the asymptotic results in approximating the properties of a finite sample. Finally, we illustrate the application of our method using a real numerical dataset.

报告人简介：张永霞，女，博士后，北方工业大学理学院讲师，毕业于人民大学。

统计推断专场三报告 7

报告时间：6 月 8 日上午　报告地点：厚德楼 3020

报告题目：Nonstationary Binary Factor Models: Maximum Likelihood Estimation

报告人：吴彬

摘要：This paper develops an asymptotic theory for general factor models incorporating observable covariates and unobservable factors generated by an integer process. This framework encompasses scenarios where single-indexes are nonstationary and cointegrated. For nonstationary single-indexes, the maximum likelihood estimator (MLE) of the coefficients is consistent under the condition, as both the cross-sectional dimension and the temporal dimension approach infinity. Similarly, the MLE of all nonstationary factors is consistent when is determined by the property of the link function. Notably, the limiting distributions of the factors exhibit a novel phenomenon: they are correlated with time due to the convergence of the Hessian matrix to zero as increase. In the case of cointegrated single-indexes, the MLEs of both factors and coefficients remain consistent, with the rate of convergence improving to. Additionally, a distinct phenomenon arises in the limiting distributions compared to nonstationary single-indexes: the dual rate of convergence of the coefficients increases from, and the class of limiting distributions changes accordingly. Moreover, the limiting distributions of the factors no longer depend on in the cointegrated case. Monte Carlo simulations verify the accuracy of the estimates. In an empirical application, we apply our model to the arrival of jumps in financial markets, where we extract the jump factor and find its excellent asset pricing performance.

报告人简介：吴彬，中国科学技术大学统计与金融系博士后。先后发表论文 A simulation-based method for estimating systemic risk measures , European Journal of Operational Research , 2024、Variance swaps with mean reversion and multi-factor variance , European Journal of Operational Research , 2024 , 315(1): 191-212。

统计推断专场三报告 8

报告时间：6 月 8 日上午　报告地点：厚德楼 3020

报告题目：Identification and estimation of causal peer effects using instrumental variables

报告人：罗姗姗

摘要：In social science researches, causal inference regarding peer effects often faces significant challenges due to homophily bias and contextual confounding. For example, unmeasured health conditions (e.g., influenza) and psychological states (e.g., happiness, loneliness) can spread among closely connected individuals, such as couples or siblings. To characterize such interference problems, we define four causal estimands for dyadic data that capture both direct and spillover effects. We employ dual instrumental variables to achieve nonparametric identification of these causal estimands in the presence of unobserved confounding. We then derive the efficient influence functions for these estimands under the nonparametric model. Additionally, we develop a triply robust and locally efficient estimator that remains consistent even under partial misspecification of the observed data model. The proposed robust estimators can be easily adapted to flexible approaches such as machine learning estimation methods, provided that certain rate conditions are satisfied. Finally, we illustrate our approach through simulations and an empirical application evaluating the causal peer effects of retirement on fluid cognitive perception among couples.

报告人简介：罗姗姗，北京工商大学数学与统计学院讲师，主要从事因果推断及其在生物医学中的应用研究。研究成果发表于 JASA、Biometrics、Statistica Sinica 和 ICML 等国际知名期刊和会议上。目前担任中国现场统计研究会因果推断分会副秘书长，主持国家自然科学基金青年项目。

统计推断专场三报告 9

报告时间：6 月 8 日上午　报告地点：厚德楼 3020

报告题目：A New Option Pricing Method Based on Generative Adversarial Networks and Space-Filling Design

报告人：王素敏

摘要：This paper proposes a novel financial option pricing method, namely the Generative Adversarial Network - space-filling sampling method based pricing (abbreviated as GAN - SFSP). It combines a Generative Adversarial Network (GAN) with a sampling method based on space-filling design. The main aim is to improve the accuracy and efficiency of option pricing by overcoming the limitations of traditional sampling methods in option pricing. In this study, GAN is used to model the process of the underlying asset, with the purpose of providing data inputs that conform to actual financial situations for the sampling method, serving as an alternative to traditional Quasi-Monte Carlo in option pricing.

报告人简介：王素敏，现任河北工业大学理学院讲师。她博士毕业于东北师范大学。其主要研究方向包括试验设计、大规模数据的抽样与分析以及统计机器学习等多个方向。目前已在 Science China Mathematics、Statistical Papers 等期刊发表 6 篇学术论文，并主持一项国家自然科学基金青年基金

社会经济统计专场报告 1

报告时间：6 月 8 日上午　报告地点：明德楼 1019

报告题目：应对气候变化的统计行动：国际标准与国家经验

报告人：王勇

摘要：气候变化正在深刻影响着人类的生存和经济社会的发展，准确及时的气候变化核算能够为人类更好的应对气候变化提供有力支撑。气候变化核算是环境经济核算的重要内容，联合国统计委员会先后发布了三个关于环境经济核算的国际标准：《2012 年环境经济核算体系：中心框架》（SEEA2012）、《2012 年环境经济核算体系：实验性生态系统核算》（SEEA2014）和《环境经济核算体系：生态系统核算》（SEEA2021）。基于这三个国际标准，全球各国开展了气候变化的统计核算工作，为人类更好的应对气候变化提供了较好的统计基础。该文首先梳理了三个国际标准的历史演进，从核算对象、核算范围、类型分类、账户表式四个方面对三个国际标准进行比较分析，重点介绍了最新的 SEEA2021；其次，对各国开展气候变化统计核算的现状以及实践经验进行归纳总结，并对我国当前气候变化核算的现状以及实践经验进行重点分析；最后提出我国开展气候变化核算的进一步完善与工作推进。

报告人简介：王勇，东北财经大学统计学院副院长，教授、博士生导师，国家社科基金重大项目首席专家。辽宁省文化名家暨“四个一批”人才、“兴辽英才计划”青年拔尖人才、中国统计教育学会常务理事兼青年经济统计学者分会副会长。主持 30 余项国家级和省部级等研究课题，在国内外高水平学术期刊发表学术论文 90 余篇。

社会经济统计专场报告 2

报告时间：6 月 8 日上午　报告地点：明德楼 1019

报告题目：互联网“免费”商品消费及居民真实生活成本指数估计

报告人：陈立双

摘要： 互联网“免费”商品创造了大量消费者剩余，显著影响消费者时间分配及生活成本，对 CPI 编制带来重要挑战。本文基于家庭生产及时间分配理论，探讨了纳入互联网“免费”商品后 CPI 的创新编制方法，据此编制了中国城镇地区居民真实生活成本指数 CPIT，获得了以下富有启发性的结论：（1）“免费”商品和传统商品对消费者消费效用水平的贡献度各占约 50%，两者在满足消费者效用水平方面没有显著差异。（2）纳入“免费”商品和考虑居民消费时间机会成本后，样本期间内 CPIT 指数较传统 CPI 指数年平均提高了 1.3 个百分点，“免费”商品实际上并非真正免费，其消费反而可能让居民承担了更高的时间机会成本。（3）CPIT 序列与 CPI 序列间具有较强的一致性与同步性，并保持相对长期的均衡关系。在反映通胀水平变化趋势上，CPIT 也具有较好的代表性。本研究将为统计部门开展“免费”商品消费的统计核算及数字经济时代 CPI 创新编制等提供参考。

报告人简介：陈立双，男，统计学博士，教授，湖北经济学院旅游与酒店管理学院副院长，硕士生导师，湖北经济学院“腾龙学者”，校学术委员会委员。主要研究领域为国民经济核算、统计理论与方法、大数据统计分析等，在《统计研究》《财贸经济》《统计与信息论坛》《统计与决策》等学术期刊发表学术论文 30 多篇，出版学术专著 1 部。主持国家社科基金、省社科基金等课题近 10 项，参与教育部哲学社会科学重大攻关项目、国家社科基金重点项目等多项课题。担任《统计与信息论坛》《西部论坛》学术期刊匿名审稿人，荣获湖北经济学院“科研先进个人奖”（2023）。

社会经济统计专场报告 3

报告时间：6 月 8 日上午　报告地点：明德楼 1019

报告题目：行政壁垒破除与资本跨区流动——基于企业异地投资的视角

报告人：冯凌秉

摘要：在新发展格局下，资本跨区自由流动是优化资源配置、促进区域均衡发展以及推动全国统一市场建设的核心驱动力。本文将 2019 年发布的《关于开展妨碍统一市场和公平竞争的政策措施清理工作的通知》视为准自然实验，以 2016—2021 年中国沪深 A 股上市公司为研究样本，通过双重差分方法分析了该清理工作对企业异地投资的影响。研究发现，清理工作的开展显著促进了企业的异地投资，且这一结论经过一系列稳健性检验后依然成立。机制分析表明，清理工作通过降低企业成本、提高企业效率和提高市场竞争充分性，促进了企业异地投资的增加。使用机器学习中的广义随机森林模型，进一步的异质性分析发现，清理工作在年轻企业、小规模企业以及政商关系较差、法治环境较弱的地区对企业异地投资的促进效应更为显著。本文拓展了关于全国统一市场建设的研究，为进一步深化政府治理、推动跨区资本流动和提升区域资源配置效率提供了政策建议。

报告人简介：冯凌秉，江西财经大学统计与数据科学学院副教授，国势研究院副院长，入选江西省“赣鄱俊才-高校领军人才”。研究方向为国势学、信用风险与机器学习。主持国家自科等省部级课题 6 项，在 Review of Finance, Journal of Hydrology, International Journal of Forecasting 等期刊发表论文 30 余篇。

社会经济统计专场报告 4

报告时间：6 月 8 日上午　报告地点：明德楼 1019

报告题目：基于数据质量效用的公共数据定价方法研究

报告人：朱贺

摘要： 公共数据定价是公共数据市场化运营的前提，也是数据要素市场建设的重要一环。本文提出了一种综合考虑公共数据质量与效用的定价模型。从数据科学视角分析了基于效用理论的公共数据定价机理；在构建公共数据质量测度指标体系的基础上，运用机器学习算法探究公共数据在质量层面的效用，进而建立公共数据授权运营制度下公共数据定价模型，并基于运营单位利润最大化确定公共数据价格求解方法。最后，以浙江省公共数据平台上的民宿数据集为例，计算了基于质量效用的最优价格，并验证了模型的适用性与稳健性，结果表明本文提出的定价方法可以实现具体交易场景下运营单位利润与用户效用最大化。本研究为公共数据定价提供了新的思路和方法参考，对推动我国公共数据市场化运营具有重要意义。

报告人简介：朱贺，浙江工商大学统计与数学学院副教授，硕士生导师，经济统计系主任。研究方向为数字经济统计核算，主持国家社科基金一般项目和青年项目等课题 10 余项，在 SSCI、SCI、CSSCI 等收录期刊发表学术论文 10 余篇，获浙江省科技进步三等奖，浙江省教育教学成果特等奖。

社会经济统计专场报告 5

报告时间：6 月 8 日上午　报告地点：明德楼 1019

报告题目：中国数字乡村建设水平：测算、区域差异及影响因素研究

报告人：代金辉

摘要：本文依据数字乡村内涵及数字化赋能理论，创新性的从数字基础设施建设、智慧农业、乡村数字经济、数字化治理效能、智慧绿色乡村五个维度构建评价体系。首先，采用 CRITIC-EW-改进 TOPSIS 法测算 2018-2022 年中国 30 个省（自治区、直辖市，西藏、港澳台除外）数字乡村建设水平，从全国、区域、省际层面多角度深入分析其发展现状及时空演变趋势。再次，采用 Dagum 基尼系数及分解法与核密度估计考察中国数字乡村建设水平的空间相对差异与绝对差异，揭示差异动态变化并量化差异程度。最后，运用空间杜宾模型探索影响中国数字乡村发展的重要因素。

报告人简介：代金辉，女，博士，副教授，现任山东工商学院统计学院应用统计教研室主任、教工第一党支部书记，硕士生导师。教育部硕士学位论文评审专家。先后获得山东工商学院女职工“建工立业标兵”、山东工商学院创新创业教育“优秀教师”、山东工商学院“巾帼读书榜样”等荣誉称号。主要研究方向为经济社会统计、计量经济分析方法及应用。

社会经济统计专场报告 6

报告时间：6 月 8 日上午　报告地点：明德楼 1019

报告题目：基于因果发现的产业链知识图谱构建与应用

报告人：沈寒蕾

摘要：应用自然语言处理和因果发现算法，基于国民经济行业分类、投入产出表等，从产业、行业到企业，构建一个多层次多维度产业链知识图谱，为监管、投融资和招商等应用领域提供服务。

报告人简介：沈寒蕾，中南财经政法大学，统计与数学学院 讲师，研究方向：因果科学、经济统计。

社会经济统计专场报告 7

报告时间：6 月 8 日上午　报告地点：明德楼 1019

报告题目：数据要素共享与城市经济韧性——基于双重机器学习的因果推断

报告人：陈鑫鹏

摘要：数据要素共享是推进区域经济高质量发展的重要举措，为提高城市应对风险的能力与增强经济韧性带来了新的契机。基于 2002-2022 年 288 个地级市面板数据，以公共数据开放平台构建为准自然实验，采用双重机器学习模型实证检验公共数据开放对城市经济韧性的影响。研究发现，公共数据开放平台建立对城市经济韧性具有显著的带动效应。就影响机制而言，公共数据开放平台建立能够通过提高信息化发展水平、促进劳动要素和资本要素集聚而提升城市经济韧性能力，且该路径具有长期效应。此外，数据共享弱化了地理壁垒，打破了传统数据在空间上的隔阂，降低经济韧性发展区域差距，具有显著的区域协调效应。本文研究结论为我国建立数据要素共享市场机制、提升城市经济韧性、加快区域协调发展提供了理论基础和政策启示。

报告人简介：陈鑫鹏，河南财经政法大学，讲师。

社会经济统计专场报告 8

报告时间：6 月 8 日上午　报告地点：明德楼 1019

报告题目：中国陪诊服务市场调查—基于五省实证数据

报告人：何义伟

摘要："本文以 2012 年和 2017 年投入产出表为基础，深入探究长三角地区产业关联结构的动态演变与空间分异特征。研究发现，长三角地区核心驱动型产业不断强化，化学制品、金属精炼・压延加工业等成为经济增长的关键力量；区域分工持续深化，形成了以上海为服务核心、苏浙为转化高地、安徽为配套基地的“核心 - 边缘”梯度结构；新兴产业与传统产业分化明显，数字经济和高端制造业加速发展，传统劳动密集型产业对经济增长的贡献逐渐减弱；输入输出结构呈现空间分异，各省市在产业发展模式和对外依存度上存在显著差异。基于上述结论，文章提出长三角应强化核心产业协同、优化梯度分工机制、统筹能源与生态安全，同时针对各省市产业发展特点给出了具体建议，旨在推动长三角地区产业协同发展，提升区域整体竞争力。

报告人简介：何义伟，安徽财经大学，硕士研究生。

社会经济统计专场报告 9

报告时间：6 月 8 日上午　报告地点：明德楼 1019

报告题目：数字普惠金融赋能乡村振兴—基于县域数据的实证检验

报告人：李静雅

摘要：乡村振兴是实现社会主义现代化的关键组成部分，也是新时代建设农业强国的重要任务。实践经验表明，数字普惠金融是推进乡村振兴有力工具之一。本文基于数字普惠金融发展指数和熵值法测算了 2014—2021 年 1560 个县域的乡村振兴指数，从影响机制和异质性分析检验了数字普惠金融助推乡村振兴的效应。研究发现：数字普惠金融能够通过激励县域人工智能发展及双创水平提升显著推动乡村振兴，通过工具变量法和多重稳健性检验后，结果依然稳健；并且在胡焕庸线东侧、距离省会越近以及有淘宝村的县域，促进作用越显著；进一步结合中国县域发展的实际情况，充分利用发挥数字普惠金融的优势，为促进乡村振兴提供了政策建议。

报告人简介：李静雅，安徽财经大学，研究生。

社会经济统计专场报告 10

报告时间：6 月 8 日上午　报告地点：明德楼 1019

报告题目：数字经济的生产力跃迁效应—回归分析与空间异质性

报告人：卢玄龙

摘要：为探究数字经济对新质生产力的促进作用以及数字经济的空间异质性，本文首先用熵值法、主成分分析法、因子分析法测度了 2011~2022 年中国 31 个省份数字经济与新质生产力的综合发展水平，其次利用双向固定效应模型进行计量分析。结果发现：第一，数字经济显著促进了新质生产力发展，不同分组的数字经济对新质生产力的影响存在显著差异。第二，作用机制分析表明，数字经济带来的更多是实质性创新而不是策略性创新。第三，利用莫兰指数对中国各省份数字经济的空间分布特征进行进一步分析，结果表明数字经济呈现明显的空间聚集现象。本文的研究为推动数字经济赋能新质生产力发展提供了效应、机制、地区差异和空间集聚的理解。

报告人简介：卢玄龙，重庆工商大学，硕士研究生。

贝叶斯与模型平均专场报告 1

报告时间：6 月 8 日上午　报告地点：博智楼 70201

报告题目：Frequentist model averaging under a linear exponential loss

报告人：李新民

摘要：This paper introduces a new model averaging approach to consider uncertainty in model specification using an asymmetric loss, linear exponential (LINEX) loss function. We are motivated by the existing model-averaging prediction analysis studies being based on symmetric loss functions, which cannot meet practical situations where different weights are needed for over-prediction and under-prediction. The existing approaches cannot be used for the asymmetric loss. The proposed model averaging estimator established via the LINEX model averaging (LMA) criterion is shown to be optimal in achieving the lowest possible LINEX loss. We demonstrate the superiority of the LMA method and its effectiveness in movie forecasting and bitcoin volatility forecasting applications. Compared to other methods, the LMA estimator effectively reduces asymmetric loss and performs reasonably well even in the case of symmetric loss.

报告人简介：李新民，青岛大学数学与统计学院教授，博士生导师。2004 年 7 月于中国科学院数学与系统科学研究院获理学博士学位。主要研究领域：Fiducial 推断、计算机试验、模型平均、大数据统计等。在 Science China Mathematics, Journal of Computational and Graphical Statistics，Statistica Sinica，Statistics and Computing 等国内外重要期刊发表高水平学术论文，先后主持了国家自然科学基金面上项目和国家社会科学基金一般项目 3 项。学术兼职主要有：全国工业统计教学研究会理事，中国现场统计研究会试验设计分会、大数据统计分会、资源与环境统计分会、生存分析分会和多元分析应用专业委员会等分会常务理事，全国工业统计教学研究会理事，中国数学会均匀设计分会理事等。

贝叶斯与模型平均专场报告 2

报告时间：6 月 8 日上午　报告地点：博智楼 70201

报告题目： Robust Model Averaging Prediction of Longitudinal Response with Ultrahigh-dimensional Covariates

报告人： 吕晶

摘要： Model averaging is an attractive ensemble technique to construct fast and accurate prediction. Despite of having been widely practiced in cross-sectional data analysis, its application to longitudinal data is rather limited so far. We consider model averaging for longitudinal response when the number of covariates is ultrahigh. To this end, we propose a novel two-stage procedure in which variable screening is first conducted and then followed by model averaging. In both stages, a robust rank-based estimation function is introduced to cope with potential outliers and heavy-tailed error distributions, while the longitudinal correlation is modeled by a modified Cholesky decomposition method and properly incorporated to achieve efficiency. Asymptotic properties of our proposed methods are rigorously established, including screening consistency and convergence of the model averaging predictor, with uncertainties in the screening step and selected model set both taken into account. Extensive simulation studies demonstrate that our method outperforms existing competitors, resulting in significant improvements in screening and prediction performance. Finally, we apply our proposed framework to analyze a human microbiome dataset, showing the capability of our procedure in resolving robust prediction using massive metabolites.

报告人简介：吕晶，西南大学副教授，硕士生导师。2015 年毕业于重庆大学统计学专业获得理学博士学位。2018 年 2 月至 2019 年 1 月在新加坡国立大学统计与应用概率系从事博士后研究。目前主要从事超高维统计和模型平均研究，部分研究结果发表于《 J. Royal Statist. Soc., Ser. B》、《J. Amer. Statist. Assoc.》、《Statist. Sinica》、《 J. Royal Statist. Soc., Ser. C》、《数学学报》。主持完成国家青年基金项目 1 项、重庆市面上项目 2 项、2020 重庆市留学人员回国创新支持计划 1 项。2024 年入选第四批重庆市学术技术带头人后备人选。

贝叶斯与模型平均专场报告 3

报告时间：6 月 8 日上午　报告地点：博智楼 70201

报告题目：Relative Error Model Averaging for Multiplicative Models

报告人：夏小超

摘要：We propose a relative error model averaging (REMA) approach to predict positive response values under a set of multiplicative error models. To estimate the parameters in each candidate multiplicative model, we utilize a relative error loss as the empirical objective function. Specifically, we consider two commonly used loss functions: the least product relative error (LPRE) and the least absolute relative error (LARE), under which two model averaging estimators, REMA-LPRE and REMA-LARE, are proposed accordingly. The optimal weight vector is chosen by minimizing a jackknife version of the relative error loss. Theoretically, it is shown that under some technical conditions, our proposed model averaging estimators enjoy asymptotic optimality under the two losses, respectively, in the sense that its loss defined by a final prediction error (FPE) is asymptotically identical to that of the best yet infeasible model averaging estimator. Furthermore, we propose a model-based screening approach to deal with the high-dimensional data setting when the number of candidate models are extremely large, and then present an extension to relax the sum-to-one constraint. Extensive simulations and empirical applications are conducted to demonstrate the practical performance of our approach.

报告人简介：夏小超，重庆大学数学与统计学院副教授，硕士生导师。2015 年重庆大学统计学专业博士毕业。目前主要感兴趣的研究方向是海量数据下的分布式估计、子抽样和模型平均。在统计学和计量经济学期刊 Journal of Econometrics、Biometrics、Statistica Sinica、Scandinavian Journal of Statistics、Statistics and Computing、CSDA、JSPI 等期刊上发表论文 20 余篇。主持完成国家自然科学基金青年项目、湖北省自然科学基金 1 项，中央高校基金项目 1 项。

贝叶斯与模型平均专场报告 4

报告时间：6 月 8 日上午　报告地点：博智楼 70201

报告题目：Balanced Active Learning

报告人：王中雷

摘 要： Active learning aims to minimize annotation costs and enhance learning efficiency by intelligently selecting the most informative data points for labeling. In this paper, we propose a balanced active learning algorithm, incorporating the estimated labels in the sampling procedure. Theoretical properties are investigated under regularity conditions. Numerical experiments demonstrate the advantage of the proposed algorithm compared with state-of-art alternatives.

报告人简介：王中雷，厦门大学王亚南经济研究院副教授，研究兴趣包括抽样调查以及人工智能在气象水文等科学的交叉应用。多篇成果发表于 JRSS-B、JASA、Biometrika 以及 Nature Communications 等国际知名期刊。

贝叶斯与模型平均专场报告 5

报告时间：6 月 8 日上午　报告地点：博智楼 70201

报告题目：An efficient Bayesian design method for estimating Shapley values.

报告人：周正

摘要：The Shapley value is a well-known concept in cooperative game theory, which provides a fair way to allocate cooperative gains (or costs) to players. However, computing Shapley values for a game involving d players requires evaluating values of all 2^d player coalitions, which is infeasible for a large d. Recently, the Shapley value has found widespread application in artificial intelligence, data science and many other emerging fields. In problems within these domains, calculating the value of a single coalition can be quite costly, further increasing the difficulty of obtaining Shapley values. In response to the computational issue, this paper proposes an efficient design-based method for estimating Shapley values. This method provides a prior distribution for values of a cooperative game, and then selects the design formed by key coalitions to provide Bayesian estimations of Shapley values by minimizing their posterior variances. The proposed method has been theoretically proven to ensure fairness for every player. Moreover, theoretical results are developed to substantially decrease the complexity in design and estimation processes. Compared with existing methods, this method offers three key advantages. First, it only requires computing the values of at least d^2-d+1 different coalitions to provide precise estimates of Shapley values for d players. Second, it allows for statistical inference of the estimates, so the confidence intervals of Shapley values can be provided. Third, it imposes no restrictions on the types of cooperative games, exhibiting strong robustness. Multiple case studies demonstrate that, our method outperforms existing methods in terms of estimation accuracy and identifying key players, at the same or lower costs.

报告人简介：周正，博士，目前是北京工业大学统计系讲师。他的研究方向是试验设计，也包括其在数据科学和机器学习中的应用。他在 Biometrika, Technometrics, IEEE TKDE 等期刊上发表了共计 5 篇文章。

复杂数据建模专场报告 6

报告时间：6 月 8 日上午　报告地点：博智楼 70201

报告题目：Estimating the variogram matrix of the multivariate Hüsler-Reiss Pareto model using a truncated approach

报告人：胡爽

摘要：In multivariate extreme value analysis, it is often the case that where distinct groups of risk variables may exhibit joint extremes while others do not, rather than all risk variables are always large simultaneously. Under such circumstances, the extremal dependence between some components of the limiting multivariate Pareto distribution may be weak, leading to substantial biases in the empirical estimation of the associated extremal variogram matrix. In this paper, under the assumption of a H\""{u} sler--Reiss Pareto model, we propose two classes of truncated estimators for the variogram matrix based on the method of moments and maximum likelihood estimation. We investigate the asymptotic properties of the proposed estimators and demonstrate through simulation studies that, in certain settings, the truncated estimators can significantly reduce estimation biases compared to conventional approaches.

报告人简介：胡爽，重庆邮电大学，讲师。

贝叶斯与模型平均专场报告 7

报告时间：6 月 8 日上午　报告地点：博智楼 70201

报告题目：Optimal Design of the Multivariate Bayesian Generalized Likelihood Ratio Schemes for Monitoring the Mean Vector

报告人：蔡洪幸

摘要："Bayesian approaches have gained popularity in dealing with operational research and management problems, including process monitoring or product quality assessments.In particular, Bayesian approaches are often used to deal with multivariable complex processes. This paper analyses and evaluates a multivariate Bayesian generalized likelihood ratio, abbreviated as MBGLR, scheme to monitor the mean vector of a multivariate Gaussian process distribution. The Monte Carlo simulation is employed to evaluate the detection performance of the MBGLR scheme, and the optimal window size is provided in terms of the simulation results. The optimal design of the MBGLR scheme is presented based on the relative mean index (RMI). Furthermore, the MBGLR scheme is compared with the multivariate Bayesian cumulative sum, referred to as MBCUSUM,and a combination of two MBCUSUM, denoted as 2MBCUSUM, schemes according to the steady-state average time to signal (SSATS). The simulation results demonstrate that the MBGLR scheme exhibits superior competitiveness in overall detection performance. We consider a case study involving the lumber manufacturing process to illustrate the implementation of the MBGLR scheme. Some concluding remarks are offered."

报告人简介：蔡洪幸，辽宁大学，研究生。

贝叶斯与模型平均专场报告 8

报告时间：6 月 8 日上午　报告地点：博智楼 70201

报告题目：Some Elaborate Comparisons of Several Distribution-Free Schemes in Short Production Runs and Some New Results

报告人：廉爽

摘要：The trend towards small batch production is accelerating due to advancements in manufacturing processes and the evolution of customer demands. Besides, there is often an unknown or non-normal distribution of process observations in real manufacturing scenarios. To address this challenge, we study the WRS scheme for monitoring the location parameter, three distribution-free schemes for jointly monitoring location and scale parameters, and the Cramér-von Mises scheme for general shift detection in short production runs. In this paper, we employ an extensive Monte-Carlo simulation to compare the performance of these SPM schemes under various underlying distributions. This evaluation takes account of factors such as the truncated average run length and the relative mean index. Furthermore, we investigate the impact of various charting parameters on the detection performance. A numerical example is offered. The paper concludes with a summary of our findings.

报告人简介：廉爽，辽宁大学数学与统计学院，研究生。

贝叶斯与模型平均专场报告 9

报告时间：6 月 8 日上午　报告地点：博智楼 70201

报告题目：Bayesian Elastic Net variable selection and application for spatial quantile panel autoregressive model

报告人：刘海芸

摘要：In this paper, we apply Elastic Net penalty method to analyze Bayesian spatial quantile panel autoregressive model and derive full conditional posterior distribution based on asymmetric Laplace distribution. The Markov Chain Monte Carlo (MCMC) algorithm is used for parameter posterior estimation and variable selection. We compare this method with the least absolute shrinkage and selection operator (LASSO) and adaptive LASSO quantile regression method when different degrees of correlation among explanatory variables. Additionally, we explore the advantages of Bayesian Elastic Net quantile regression method under different random disturbance terms and sample sizes. Simulation studies illustrate that this method is especially effective in scenarios where there is correlation in explanatory variables and small samples. In empirical research part, we select Economic Value Added (EVA) of listed companies in four conceptual sectors: Internet finance, Internet industry, Internet insurance and Internet medical care, as response variables, select 15 explanatory variables such as operating profit margin, and construct a Bayesian Elastic Net spatial quantile panel autoregressive model. Empirical result indicates that return on total assets, inventory turnover rate, long-term capital liability ratio, interest earned multiple and total asset growth rate have a greater impact on EVA, the Elastic Net penalty method eliminates remaining 10 variables.

报告人简介：刘海芸，辽宁大学，研究生。

机器学习专场一报告 1

报告时间：6 月 8 日上午　报告地点：博智楼 70202

报告题目：Linear extremile regression and its semi-supervised learning

报告人：王江峰

摘要：The extremile (Daouia et al., 2019) is a novel and coherent measure of risk, determined by weighted expectations rather than tail probabilities. However, existing studies on extremile are difficult to generalize to linear models due to the challenge of obtaining a $\sqrt{n}$ -consistent estimator of unknown parameters in the model. To address this issue, this article proposes a new definition of linear extremile regression and its estimation method. In many practical data analyses, linear regression models may not be accurate. Therefore, we have developed a semi-supervised framework for the proposed linear extremile regression using unlabeled data, which is often present in practical problems. Both simulations and real data analyses have been conducted to illustrate the finite sample performance of the proposed methods.

报告人简介：王江峰，男，江西南昌人，1978 年 11 月生，博士(后)，统计学教授，博士生导师。现任应用概率统计研究所所长，全国工业统计学教学研究会理事，入选校西湖学者拔尖人才入选人员。主要研究领域有：复杂生存数据分析、高维数据分析、分位数回归方法等。主持国家社会科学基金面上项目 2 项，以及教育部人文社科基金等省部项目 5 项；其中代表性课题有：国家社科基金面上项目“删失指标随机缺失下分位数回归模型的研究及其应用”；国家社科基金面上项目“复杂数据下 CQR 模型的研究及其应用”。先后在《中国科学》、《数学学报》(中，英文版)、《数学进展》、《Front. Math. China》、《Statistics》、《J. Statist. Plann. Inference》、《Statist. Papers》、《Adv. Statist. Anal》、《J. Korean. Stat. Soc》、《Statist. Probab. Lett》等期刊上发表特级、一级以及 SCI/SSCI 收录的论文 40 余篇，出版专著 1 部(A 类)。

机器学习专场一报告 2

报告时间：6 月 8 日上午　报告地点：博智楼 70202

报告题目：Exponential Power Mixture of Experts Model: Estimation, Clustering, and Variable Selection

报告人：冯峥晖

摘要："The mixture of experts model is a popular framework for modeling heterogeneity in data for regression, classification, and clustering. For regression and cluster analyses of continuous data, MoE usually uses normal experts following the Gaussian distribution. However, for a set of data containing a group or groups of observations with heavy tails or outliers, the use of normal experts is unsuitable and can unduly affect the fit of the MoE model. Motivated by the flexibility of the exponential power distribution, we propose EPMoE model and provide the corresponding estimation method, taking into account variable selection and the theoretical properties of the model. Numerical simulations demonstrate the effectiveness of our proposed method and algorithm, and finally, we provide a real data analysis to illustrate the application of our"

报告人简介：冯峥晖，女，统计学博士，哈尔滨工业大学（深圳）理学院副教授，中国现场统计研究会多元分析应用专业委员会理事。主要研究方向包括函数型数据分析，密度估计，大数据，充分降维，变量选择。在统计学国际期刊 Journal of the American Statistical Association, Computational Statistics and Data Analysis 等发表论文 20 余篇，主持国家和省级研究项目 4 项，主持教学项目 3 项。曾获得全国经济管理实验教学“联奕奖”优秀教师称号（全国十佳），哈尔滨工业大学（深圳）第二届青年教师教学竞赛一等奖。

机器学习专场一报告 3

报告时间：6 月 8 日上午　报告地点：博智楼 70202

报告题目：融合机器学习和统计建模的风电场短期风能预测方法研究

报告人：衡佳妮

摘要：电力低碳化是目前减少二氧化碳排放、缓解气候变化、实现全球范围内可持续发展的重要要求。随着能源需求的不断增加，为保证这一目标的实现，可再生能源并网的开发利用势在必行。风能作为一种分布广泛、资源丰富且无污染的可再生能源，是新能源开发和建设的重点领域。然而，由于风能具有随机性、间歇性和波动性等特点，使得大规模风电并网始终存在输送和消纳等技术问题，对电力系统的稳定性造成极大的冲击，给电网安全运营和管理来了严峻的挑战。研究表明，精确可靠的短期风功率预测能够有效解决大规模风电并网带来的系统运行风险，并对提升电网精准调度、平衡备用容量和健全电网运营管理机制具有重要的指导作用。本研究通过数据挖掘、机器学习和多种统计建模策略，提出了一系列风能时间序列预测模型，有效提升了风电时间序列的预测精度，丰富了面向清洁能源应用的人工智能建模理论。

报告人简介：衡佳妮，北京工商大学数学与统计学院副教授，硕士生导师，Data Science and Management 期刊青年编委等，主持国自然青年基金项目一项，参与国家社科基金重大项目，国家自然科学基金面上项目等。研究方向包括预测理论与方法、统计大数据分析与机器学习在能源环境领域的应用等。在国际杂志 Applied Energy、Energy 等国际重要 SCI/SSCI 学术期刊发表论文 20 余篇。

机器学习专场一报告 4

报告时间：6 月 8 日上午　报告地点：博智楼 70202

报告题目：Spatial prediction of species invasion using a hybrid of machine learning and geostatistical method

报告人：沈亮

摘要：Modeling ecological patterns and processes often involve large-scale and complex high-dimensional spatial data. Due to the nonlinearity and multicollinearity of ecological data, traditional geostatistical methods have faced great challenges in model accuracy. As machine learning has increased our ability to construct models on big data, the main focus of the study is to propose the use of statistical models that hybridize machine learning and spatial interpolation methods to cope with increasingly large-scale and complex ecological data. Here, two machine learning algorithms, were combined with ordinary kriging to model plant invasions across the eastern United States. The accuracies of the hybrid models and conventional models were evaluated by 10-fold cross-validation. Based on an invasive plants dataset of 15 ecoregions across the eastern United States, the results showed that the hybrid algorithms were significantly better at predicting plant invasion when compared to commonly used algorithms in terms of RMSE and paired-samples *t*-test (with the *p*-value < .0001). Besides, the additional aspect of the combined algorithms is to have the ability to select influential variables associated with the establishment of invasive cover, which cannot be achieved by conventional geostatistics. Higher accuracy in the prediction of large-scale biological invasions improves our understanding of the ecological conditions that lead to the establishment and spread of plants into novel habitats across spatial scales.

报告人简介：沈亮，青岛理工大学商学院副教授，普渡大学访问学者，主持国家统计科学研究项目、教育部人文社科项目、山东省自然科学基金(面上、青年)等项目。主要研究方向为生态统计、机器学习、空间统计。

机器学习专场一报告 5

报告时间：6 月 8 日上午　报告地点：博智楼 70202

报告题目：Transfer Learning for higher-order partially linear spatial autoregressive model with differential privacy

报告人：宋允全

摘要：The effectiveness of transfer learning in leveraging external information to improve the learning performance and predictive accuracy of target domain models is widely recognized. In terms of privacy protection, differential privacy has been shown to provide robust and consistent privacy protection. However, there is no research on transfer learning and differential privacy in the context of higher-order partially linear spatial autoregressive models. In this study, we fill this gap by introducing a differential privacy-based transfer learning approach for this model. We propose different algorithms for the cases where the transferable sources are known and unknown, respectively. Through extensive simulation experiments and practical applications, we validate the effectiveness of the proposed methods.

报告人简介：宋允全，中国石油大学（华东）理学院数据科学与统计系，副教授，博士生导师，主要从事空间网络数据统计建模及应用研究，发表相关论文 60 余篇，授权国家发明专利 4 项。主持数据科学领域科研项目 5 项，包括国家重点研发计划子课题 1 项、教育部人文社科基金 1 项、山东省自然科学基金 3 项。

机器学习专场与报告 6

报告时间：6 月 8 日上午　报告地点：博智楼 70202

报告题目：Transfer learning with invariance structures across target and auxiliary sources.

报告人：夏思薇

摘要：This paper studies the transfer learning problem under multiple source data with invariance structure among all sources, aiming to improve the estimation and prediction on target source by borrowing knowledge from auxiliary sources. To boost transfer learning, we propose a novel estimation method, called Joint Invariant Adaptive Transferred Estimation (JIATE), enabling a more accurate evaluation. JIATE learns the invariance structure by recognizing invariant features, which have similar effects on responses, in all sources, imposing adaptive penalty strengths to different features. It overcomes the challenge caused by the stringent similarity between target and auxiliary sources, and the noisy information from auxiliary samples. The theoretical analysis shows that JIATE guarantees selection and estimation consistency. Extensive simulations and a real-data experiment on cancer data analysis verify the effectiveness and accuracy of our proposed method.

报告人简介：夏思薇，成都理工大学数学科学学院讲师。2021 年博士毕业于重庆大学数学与统计学院。主要研究方向包括模型选择、特征选择、无模型方法和迁移学习。文章发表于 Knowledge-Based Systems、Applied Mathematics and Computation 等期刊

机器学习专场一报告 7

报告时间：6 月 8 日上午　报告地点：博智楼 70202

报告题目：Learning Semi-parametric Tree Models from Mixed Data

报告人：周灿

摘要：Causal discovery and representation involving latent variables and structures have attracted growing interest in the era of artificial intelligence, particularly for their critical role in understanding real-world data. While many existing methods focus exclusively on either purely continuous or purely discrete data, this paper addresses the challenge of learning latent structures from mixed data. We propose a novel semi-parametric tree model capable of handling mixed data， and develop an algorithm for learning the structure of this model using additive information distances. We demonstrate that this algorithm efficiently and accurately recovers the true structure, given the information distances. Additionally, the sample-based version of the structural learning algorithm achieves probabilistic approximate correctness, with a finite sample bound established for exact structural recovery. Both simulated and real data are used to assess the performance of our proposed algorithm, with experimental results showing that our algorithm can effectively discover latent hierarchical structures behind mixed data.

报告人简介：周灿，博士，南京审计大学统计与数据科学学院讲师。主要研究方向是图模型和混合数据分析。主持江苏省自然科学青年基金 1 项。在 Journal of Machine Learning Research 等期刊发表多篇论文。

机器学习专场一报告 8

报告时间：6 月 8 日上午　报告地点：博智楼 70202

报告题目：Tensor Transfer Learning under Convolutional Smooth Quantile and It’s Application

报告人：黄靖翔

摘要："Transfer learning, as one of the most important and effective means to improve the performance of target data by utilizing the information in the source data, has received extensive attention from scholars. In this paper, tensor data is introduced into the transfer learning framework under convolutional smooth quantile regression, and the classical Tucker decomposition is used for dimension reduction while keeping the internal structure of the tensor undamaged. For the oracle condition, we give a two-step estimation algorithm and establish error bounds for the estimates. In the case where the transferable set is unknown, we choose to filter the transferable set using the loss of the model from the nature of transfer. In addition to, we also considered the use of data transfer to recover the missing data. Numerical simulation results show that our screening method can effectively screen out the transferable set and thus avoid the occurrence of negative transfer, and the scale of missing data in the datasets leads to different results of data migration recovery; finally, we demonstrate the effectiveness of our proposed algorithm by using the air quality inspection data from different areas in Beijing."

报告人简介：黄靖翔，辽宁大学，研究生。

机器学习专场一报告 9

报告时间：6 月 8 日上午　报告地点：博智楼 70202

报告题目： Double/Debiased Machine Learning for Semiparametric Synthetic Difference-in-Differences Models

报告人：张子晗

摘要："We propose a new estimator for causal effects when the data do not satisfy the parallel trends assumption in high-dimensional settings. Our estimator combines insights from the semiparametric difference-in-differences and synthetic difference-indifferences methods, and compensates for the lack of parallel trends by re-weighting units. The score/moment function used for estimating effects satisfies Neyman orthogonality, which allows for the application of machine learning in the first-step estimation. Additionally, we establish the consistency and asymptotic normality of the estimator. In simulation experiments, our estimator performs well compared to other estimators. Finally, we apply this estimator to evaluate the effects of green finance on green innovation, and the results show that green finance has a significant positive effect on green innovation."

报告人简介：张子晗，辽宁大学数学与统计学院，研究生。

机器学习专场一报告 10

报告时间：6 月 8 日上午　报告地点：博智楼 70202

报告题目：Deep Learning Based Doubly Robust Test for Granger Causality

报告人：刘持金

报告摘要：We propose a doubly robust test for Granger causality via deep learning. The proposed test asymptotically controls the type I error, while achieving power approaches one. Our methodology makes several innovations. First, we integrate deep learning techniques within the framework of statistical Granger causality testing. Second, we derive a doubly robust test statistic that employs a semiparametric estimation while achieving a parametric convergence rate. Third, our method performs particularly well in handling high-dimensional settings with multiple lag orders. Fourth, a multiplier bootstrap method is employed to obtain the critical value, thereby reducing computational time costs. We validate the effectiveness of our test through numerical studies. Moreover, we apply our method to the price and volume data from three indices, and the results demonstrate the presence of Granger causality across multiple lags.

报告人简介：刘持金，目前就读于西安交通大学。主要研究方向为格兰杰因果关系、时间序列分析等。

机器学习专场二报告 1

报告时间：6 月 8 日上午　报告地点：博智楼 70203

报告题目：Optimal Multi-Machine Learning Assisted Semi-Supervised Inference

报告人：贺百花

摘要： Statistical inference is a fundamental problem with broad applications across fields such as biomedical research, economic analysis, and quality control. Recently, AI or ML-labeled synthetic data has emerged as a powerful tool for reducing the need for human annotations, offering innovative solutions to semi-supervised learning challenges. However, existing methods have yet to fully exploit the predictive power of diverse machine learning models. In this paper, we propose an efficient, robust, and reliable strategy to address the problem of constructing confidence intervals. Specifically, we introduce a new method called Weighted Prediction Power Inference (WPPI), which utilizes an efficiency-oriented-D-optimal criterion to aggregate the capabilities of different models. We demonstrate that the resulting interval sets achieve the smallest interval size with nominal coverage as the sample size increases. This theoretical guarantee holds when at least one machine learning algorithm is either efficient or weighted efficient. Additionally, our framework provides a principled approach to constructing estimation intervals that account for shifts in the covariate distribution between labeled and unlabeled data. We further illustrate the effectiveness of our method through experiments on both synthetic and real-world datasets.

报告人简介：贺百花，博士，中国科学技术大学特任副教授，主要从事模型平均以及多源域学习等领域的研究，相关论文发表在 JASA，JMLR，IJOC 等统计和机器学习期刊，主持国家自然科学基金重大项目子课题以及青年基金，入选省部级人才计划支持。

机器学习专场二报告 2

报告时间：6 月 8 日上午　报告地点：博智楼 70203

报告题目： Kernel method with latent inputs and its applications to factor-based nonparametric regression

汇报人： 贺莘

摘要： Kernel methods, particularly kernel ridge regression (KRR), are time-proven, powerful nonparametric regression techniques known for their rich capacity, analytical simplicity, and computational tractability. The analysis of their predictive performance has received continuous attention for more than two decades. However, in many modern regression problems where the feature inputs used in KRR cannot be directly observed and must instead be inferred from other measurements, the theoretical foundations of KRR remain largely unexplored. In this talk, we introduce a novel approach for analyzing KRR with latent feature inputs, and we also apply our general results to factor-based nonparametric regression models and establish the minimax optimality of KRR when the latent inputs are derived via principal component analysis. Our theoretical findings are further corroborated by simulation studies.

报告人简介： 贺莘，上海财经大学统计与数据科学学院， 副教授，博士生导师。主要研究方向为统计机器学习，尤其是基于再生核希尔伯特空间的非参数方法及其应用、图模型及网络数据分析等。已有二十余篇研究论文发表在国际权威期刊上。主持自然科学青年基金项目，上海市面上项目以及入选了上海市浦江人才计划。

机器学习专场二报告 3

报告时间：6 月 8 日上午　报告地点：博智楼 70203

报告题目： Interpretable Deep Regression Models with Interval-Censored Failure Time Data

报告人： 李树威

摘要： Deep neural networks (DNNs) have become powerful tools for modeling complex data structures through sequentially integrating simple functions in each hidden layer. In survival analysis, recent advances of DNNs primarily focus on enhancing model capabilities, especially in exploring nonlinear covariate effects under right censoring. However, deep learning methods for interval-censored data, where the unobservable failure time is only known to lie in an interval, remain underexplored and limited to specific data type or model. This work proposes a general regression framework for interval-censored data with a broad class of partially linear transformation models, where key covariate effects are modeled parametrically while nonlinear effects of nuisance multi-modal covariates are approximated via DNNs, balancing interpretability and flexibility. We employ sieve maximum likelihood estimation by leveraging monotone splines to approximate the cumulative baseline hazard function. To ensure reliable and tractable estimation, we develop an EM algorithm incorporating stochastic gradient descent. We establish the asymptotic properties of parameter estimators and show that the DNN estimator achieves minimax-optimal convergence. Extensive simulations demonstrate superior estimation and prediction accuracy over state-of-the-art methods. Applying our method to the Alzheimer's Disease Neuroimaging Initiative dataset yields novel insights and improved predictive performance compared to traditional approaches.

报告人简介： 李树威, 统计学博士, 现任广州大学经济与统计学院副教授。2017 年 6 月博士毕业于吉林大学统计系。主要研究方向为生物统计、应用统计，相关研究成果发表在 Biometrika、Biometrics、Statistica Sinica 等期刊上。主持国家自然科学基金面上项目、国家自然科学基金青年项目、广东省自然科学基金面上项目、广州市科技局项目等。现任广东省现场统计学会常务理事、中国现场统计研究会资源与环境分会理事等。

机器学习专场二报告 4

报告时间：6 月 8 日上午　报告地点：博智楼 70203

报告题目：Model- and similarity-free transfer learning for classification with label noise

汇报人：任明旸

摘要：Transfer learning, which seeks to enhance the learning of the target domain by leveraging relevant auxiliary domains, has garnered significant attention in the statistical community. However, most existing methods rely on correct model specification and restrictive similarity assumptions (such as overall similarity or shared subspaces) between the target and auxiliary domains. While these assumptions have led to success in many classical models, they often fail to hold in specific and critical scenarios, limiting their practical applicability. This work addresses a common yet challenging classification scenario: the target domain comprises a small and meticulously corrected dataset with clean labels, while the auxiliary domain consists of a large dataset potentially contaminated by label noise. Without imposing any similarity assumptions between the auxiliary and target domains, a connection between the distributions of noisy and clean data is established by investigating the generation mechanism of noisy labels. This insight is ingeniously utilized to extract valuable information from the noisy data and identify “Bayes-optimal samples”. Additionally, we integrate active learning to further improve the classification performance of the small target dataset. As a general framework, our method does not require specifying the form of the classification model in any procedure, making it highly versatile. Theoretical analysis provides rigorous justification for the proposed approach, and extensive experiments on both synthetic and real-world data demonstrate its effectiveness and broad applicability.

报告人简介：任明旸，上海交通大学长聘教轨副教授、博士生导师，入选上海白玉兰海外高层次人才计划。主要研究方向为异质数据分析、图模型和生物统计。2022 年博士毕业于中国科学院大学，期间于耶鲁大学联合培养，2022-2024 于香港中文大学从事博士后研究。曾获中国科学院院长奖、ICSA 青年研究者荣誉提名奖、International Biometric Society ENAR 杰出学生论文奖。

机器学习专场二报告 5

报告时间：6 月 8 日上午　报告地点：博智楼 70203

报告题目：Some progress in transfer learning.

汇报人：王小舟

摘要：Transfer learning can leverage knowledge acquired from related source domains to assist in solving tasks in a target domain. In this talk, I will introduce some research results on transfer learning. The algorithms and the corresponding theoretical properties are presented. Numerical studies are conducted to demonstrate the effectiveness of the proposed methodologies.

报告人简介：王小舟，华东师范大学统计学院副教授，2015 年于上海交通大学致远学院获得本科学位，2020 年于上海交通大学数学科学学院获得博士学位。主要研究方向为统计机器学习，研究成果发表在 Journal of Machine Learning Research、Science China Mathematics、IJCAI 等国际期刊和会议。主持国家自然科学基金青年科学基金项目，入选上海市“晨光计划”、上海市“扬帆计划”。

机器学习专场二报告 6

报告时间：6 月 8 日上午　报告地点：博智楼 70203

报告题目：A Doubly Robust Matched Learning Approach for Optimal Multi-Category Treatment Regimes in Observational Studies

报告人：李灿辉

摘要：Personalized medicine has recently received increasing attention because of the significant heterogeneity of patient responses to the same medication. The estimation of optimal individualized treatment regime or individualized treatment rule is an important part of personalized medicine. Individualized treatment regimes are designed to recommend treatment decisions to patients based on their individual characteristics and to maximize the overall clinical benefit to the patient. However, most of the existing statistical methods are mainly concerned with the estimation of optimal individualized decision rules for the two categories of treatment options and rely heavily on data from randomized controlled trials. There has been a relative lack of research work on the selection of multicategory treatment options in real-world settings. We address this challenge and propose a machine learning approach (ACML) to estimate optimal multi-category treatment regimes. This new learning approach allows for more accurate assessment of individual treatment response and alleviation of confounding, more importantly, ACML is doubly robust, efficient and easy to interpret. We first introduce the concordance-based value function that measures weighted concordance for each patient by matching imputation. We then propose a novel surrogate loss and employ an angle-based method to maximize the concordance-based value function that directly handles the problem of optimization with multicategory treatment options. Furthermore, an extension of ACML can be applied to ordinal treatment settings. The theoretical results show that proposed method is doubly robust. We further obtain that the resulting estimator of the treatment rule is consistent. Through a large number of simulation studies, we demonstrate that ACML outperforms existing methods. Lastly, the proposed method is illustrated in an analysis of AIDS clinical trial data.

报告人简介：李灿辉，河南大学数学与统计学院讲师，硕士生导师，东北师范大学统计学博士。研究兴趣有精准医疗的统计推断、统计机器学习、教育统计等，在 Biometrika, Journal of Applied Statistics 等杂志发表学术论文多篇。

机器学习专场二报告 7

报告时间：6 月 8 日上午　报告地点：博智楼 70203

报告题目：Multiplier bootstrap meets high-dimensional PCA

报告人：虞龙

摘要：In this paper, we examine the feasibility (i.e., both the good advantages and the bad limitations) and the adaptivity (i.e., the potential for beneficial modifications) of employing multiplier bootstrap to analyze the asymptotic distributions of the largest eigenvalues of potentially spiked high-dimensional sample covariance matrices. Our findings and proposed algorithms demonstrate that multiplier bootstrap remains valid, provided the multipliers are appropriately chosen and the bootstrap procedures are applied multiple times with suitable corrections. First, for non-spiked sample covariance matrices, we propose a novel algorithm to replicate the asymptotic distribution (i.e., the Tracy-Widom law) of their largest eigenvalues. This is achieved by repeatedly bootstrapping the entire sample covariance matrix using carefully designed bounded multipliers that satisfy certain concentration properties. We highlight that unbounded multipliers fail in this setting, as the bootstrapped eigenvalues asymptotically follow a Fréchet or Gumbel distribution. Second, for spiked sample covariance matrices, while both bounded and unbounded multipliers can recover the asymptotic normality of the largest eigenvalues, they may introduce additional bias, particularly when the spikes are not strong. To mitigate this, we apply a modified multiplier bootstrap multiple times to correct the bias. Finally, leveraging our modified multiplier bootstrap procedures, we propose a novel and straightforward distribution-based test for selecting common factors in the factor model.

报告人简介：虞龙，现任上海财经大学统计与管理学院助理教授；于 2020 年在复旦大学管理学院获得理学博士学位；2018 年至 2019 年访问美国密歇根大学安娜堡分校统计系完成博士联合培养项目；2020 年-2022 年访问新加坡国立大学统计与数据科学系完成博士后研究员工作。研究方向为多元统计分析、计量经济学、随机矩阵理论等，重点关注基于因子模型的数据建模、数据降维、稳健统计等方向，学术成果发表在 Biometrika, Journal of Econometrics, Journal of Business and Economic Statistics, Bernoulli, Journal of Multivariate Analysis, Biostatistics 等期刊。

机器学习专场二报告 8

报告时间：6 月 8 日上午　报告地点：博智楼 70203

报告题目：Generalization error bounds of gradient descent for learning a three-layer vision Transformer

汇报人：程学伟

摘要：This paper investigates the generalization performance and gradient descent training dynamics of a three-layer vision Transformer (ViT) within a neighborhood of random initialization. Theoretically, we establish two key contributions: (i) a generalization error bound for the ViT under suitable initialization conditions, revealing fundamental connections between model architecture, data distribution, and generalization capability; and (ii) convergence guarantees for gradient descent optimization, demonstrating efficient parameter learning with global approximation guarantees. Together, these results provide an end-to-end learning framework characterizing both optimization and generalization for gradient-descent-trained ViTs. Numerical experiments validate the non-vacuousness of our generalization bounds, confirming their practical relevance in standard vision tasks. Our analysis offers new theoretical insights into the optimization dynamics and generalization behavior of Transformer architectures.

报告人简介：程学伟，硕士研究生导师。2023 年博士毕业于中南大学(统计学专业，理学博士)，2022 年 9 月-2023 年 5 月，公派赴新加坡国立大学统计与数据科学系联合培养。主要从事统计机器学习的理论与应用、生物统计等方面的研究工作。在《Neurocomputing》、《Computer Methods and Programs in Biomedicine》、《Computational Statistics》、《数量经济技术经济研究》、《数理统计与管理》等国内外刊物上发表学术论文 10 余篇。主持湖南省自然科学基金青年项目和教育厅优秀青年基金项目各 1 项，参与国家自然科学基金面上项目和社会科学基金一般项目各 1 项。

机器学习专场二报告 9

报告时间：6 月 8 日上午　报告地点：博智楼 70203

报告题目：Robust personalized federated learning with sparse penalization

报告人：章晓菲

摘要：Federated learning is an emerging topic due to its advantage in collaborative learning with distributed data. Due to the heterogeneity in the local data-generating mechanism, it is important to consider personalization when developing federated learning methods. In this work, we propose a personalized federated learning method to address the robust regression problem. Specifically, we aim to learn the regression weight by solving a Huber loss with the sparse fused penalty. Additionally, we designed our personalized federated learning for robust and sparse regression (PerFL-RSR) algorithm to solve the estimation problem in the federated system efficiently. Theoretically, we show that the proposed PerFL-RSR reaches a convergence rate of O(1/T), and the proposed estimator is statistically consistent. Thorough experiments and real data analysis are conducted to corroborate the theoretical results of our proposed personalized federated learning method. Supplementary materials for this article are available online including a standardized description of the materials available for reproducing the work.

报告人简介：章晓菲，中南财经政法大学统计与数学学院讲师。博士毕业于美国爱荷华州立大学统计系。研究兴趣包括分布式统计推断和高维数据分析等。在 JASA, TIFS，ICLR 等国际杂志和国际会议发表论文数篇

统计计算与优化专场报告 1

报告时间：6 月 8 日上午　报告地点：博智楼 70204

报告题目：Consistent specification testing of linear censored quantile regression

报告人：张飞鹏

摘要：This article proposes non-parametric tests for the correct specification of a linear quantile regression for censored data. The proposed specification tests are based on an orthogonal projection on the tangent space of nuisance parameters. The asymptotic distributions of the proposed tests are derived, and the validity of a multiplier bootstrap approximation is established. Monte Carlo simulation studies and an empirical application demonstrate desirable finite sample performance of the proposed tests.

报告人简介：张飞鹏，西安交通大学经济与金融学院教授、博士生导师。美国宾夕法尼亚州立大学博士后，英国肯特大学访问学者。担任中国优选法统筹法与经济数学研究会理事、数据科学分会副理事长，中国管理科学与工程学会理事、金融计量与风险管理分会常务理事，全国工业统计学教学研究会理事等。主持完成国家自科基金、国家社科基金重大专项子课题、省自科基金、英国 UKRI 基金等多项课题。研究领域为复杂大数据统计学习、大语言模型、金融计量经济学、数字金融等。已在综合性科学期刊《Nature Comm.》、《Genome Research》、《Science China Math.》、《Biometrics》，经管期刊《Energy Economics》、《Technological Forecasting and Social Change》、《International Review of Financial Analysis》、《Oxford Bulletin of Economics and Statistics》、《数量经济技术经济研究》、《中国管理科学》、《系统工程理论与实践》，以及机器学习国际顶级会议 *NeurIPS* 等发表 60 余篇论文。研究成果多次被国际知名学者在 Nature、Science、Cell、PNAS 等国际顶级期刊广泛引用和正面评价。

统计计算与优化专场报告 2

报告时间：6 月 8 日上午　报告地点：博智楼 70204

报告题目：Influence Diagnostics for Tensor Regression Models

报告人：郝程程

摘要：Tensor regression models have found widespread application across diverse fields; however, influence diagnostics within these models remain underexplored. This study extends local influence analysis and the case-deletion method to the generalized CP tensor model. One-step approximations of generalized Cook's distance are derived using Hessian and Fisher information matrices. Three perturbation schemes (i.e. case-weighted, single-explanatory-variable, and group-explanatory-variable) are analyzed. . Simulations and empirical analyses demonstrate that the four diagnostic methods can accurately identify influential observations when the appropriate rank assumption is met.

报告人简介：郝程程，瑞典斯德哥尔摩大学统计学博士。现任上海对外经贸大学研究生院常务副院长，统计与信息学院副教授。入选上海高校青年东方学者、上海市浦江人才。研究领域为多元纵向数据分析与异常点诊断，在《Scandinavian Journal of Statistics》《Statistical Methods in Medical Research》等国内外权威期刊发表第一作者/通讯作者论文 10 余篇；主持完成国家自然科学基金项目 1 项，省部级项目 2 项，瑞典皇家科学院项目 2 项。

统计计算与优化专场报告 3

报告时间：6 月 8 日上午　报告地点：博智楼 70204

报告题目： A Recursive Stochastic Algorithm for Real-Time Online Parameter Estimation in Item Response Theory: Enhancing Computational Efficiency for Dynamic Educational Assessment

报告人：陆婧

摘要： Traditional large-scale educational data is often static and updated periodically, making it difficult to reflect the dynamic changes in the educational environment in real time. With advancements in technology, many online exam platforms and educational assessment tools (such as adaptive tests) can collect students' response data and behavioral data in real time during the examination process. Most estimation methods in Item Response Theory (IRT), while highly accurate, are mostly used for parameter estimation in offline environments. When real-time data continuously arrives and online parameter estimation is needed, these methods become computationally impractical. To address this issue, this paper proposes a recursive stochastic algorithm--average truncated stochastic Newton algorithm--for the quick online estimation of parameters in the IRT model. The proposed algorithm can also provide parameter estimates in offline environments and demonstrates significant improvements in computational efficiency compared to the EM algorithm in the mirt package in R software. In situations requiring faster approximation methods, our proposed approach can serve as a powerful alternative to the offline EM method. Furthermore, we investigate the asymptotic properties of the algorithm, proving its almost sure convergence and asymptotic normality. Numerical experiments on both simulated and real data also demonstrate the practicality of the proposed method.

报告人简介：陆婧，现为东北师范大学数学与统计学院副教授，博士生导师。研究方向为教育统计与心理测量。主持并完成国家自然科学基金青年项目一项，中国博士后科学基金特别资助站中项目及中国博士后科学基金面上项目各一项。在 Psychometrika、Multivariate Behavioral Research、British Journal of Mathematical and Statistical Psychology 、 Journal of Mathematical Psychology 、 Journal of Educational and Behavioral Statistics 等学术期刊发表论文多篇。目前研究兴趣主要有作答反应时间的建模及其应用、作答反应行为分析、缺失数据分析、测量模型的相关算法开发、认知诊断模型等。

统计计算与优化专场报告 4

报告时间：6 月 8 日上午　报告地点：博智楼 70204

报告题目：An algorithm for distributed parameter estimation in modal regression models

报告人：马学俊

摘要：In this paper, we propose a new algorithm to handle massive data sets, which are modelled by modal regression models. Differing from the existing methods regarding distributed modal regression, the proposed method combines the divide-and-conquer idea and a linear approximation algorithm. It is computationally fast and statistically efficient to implement. Theoretical analysis for the resultant distributed estimator under some regularity conditions is presented. Simulation studies are conducted to assess the effectiveness and flexibility of the proposed method with finite sample size. Finally, an empirical application to the chemical sensors data is analyzed for further illustration.

报告人简介：马学俊，苏州大学数学科学学院的副教授，主要研究方向包括海量数据分析、统计计算、变量选择、经验似然、非参数回归等统计模型及其应用。主持和参与了多项科研项目，包括：国家自然科学基金青年基金项目“海量数据参数模型的统计推断及其应用”、江苏省青年自然基金项目“高维成分数据的变量选择及其应用研究”等。被评为苏州大学优秀青年学者。

统计计算与优化专场报告 5

报告时间：6 月 8 日上午　报告地点：博智楼 70204

报告题目：Semi-parametric inference on inequality measures with non-ignorable non-response using callback data

报告人：王淳林

摘要：Measuring income inequality is vital in economics and official statistics. In practice, the household survey data sets that are typically used to measure income inequality inevitably suffer from non-ignorable non-response, that is, the response probabilities depend on the missing actual income values. Existing methods assuming fully observed income data may not be applicable or lead to distorted statistical inference based on the complete-case analysis. In this paper, we consider efficient and reliable estimation and inference of popular measures of inequality in the presence of non-ignorable non-response. We exploit the callback data routinely collected along with many surveys to tackle the model identifiability issue and correct the biased sample through a semi-parametric modeling strategy that does not require any parametric specification on the income distribution. A full likelihood approach is developed to estimate various inequality measures. To circumvent complex numerical optimization, we further devise a novel expectation-maximization algorithm for stable and convenient computation. The asymptotic properties of the proposed estimators are established, which enable valid statistical inference of the inequality measures. The simulation results demonstrate that the proposed semi-parametric method corrects the non-response bias of the estimated inequality measures, is robust to income distributions, and leads to efficient inference results. We apply the proposed inference procedures of inequality measures to a real income survey data set with non-ignorable non-response for illustration.

报告人简介：王淳林，现任厦门大学经济学院统计学与数据科学系与王亚南经济研究院副教授，博士生导师。加拿大滑铁卢大学统计学博士。目前研究兴趣集中在基于经验似然和密度比模型的半参数与非参数统计推断方法和理论，及其在选择/抽样偏差等问题中的应用，如病例对照数据、不可忽略缺失数据、诊断医学、收入不平等、混合模型等。在 Statistica Sinica、JMVA、CSDA、Statistics in Medicine 等国际知名统计学期刊上发表多篇科研论文，并主持多项国家级和省部级科研项目。

统计计算与优化专场报告 6

报告时间：6 月 8 日上午　报告地点：博智楼 70204

报告题目：Sequential estimation of high-dimensional signal plus noise models under general elliptical frameworks

报告人：李艳鹏

摘要：High dimensional data analysis has attracted considerable interest and is facing new challenges, one of which is the increasingly available data with noise corrupted and in a streaming manner, such as signals and stocks. In this paper, we develop a sequential method to dynamically update the estimates of signal and noise strength in signal plus noise models. The proposed sequential method is easy to compute based on the stored statistics and the current data point. The consistency and, more importantly, the asymptotic normality of the estimators of signal strength and noise level are demonstrated for high dimensional settings under mild conditions. Simulations and real data examples are further provided to illustrate the practical utility of our proposal.

报告人简介：李艳鹏，哈尔滨工业大学数学学院博士后副研究员，2023 年毕业于哈尔滨工业大学，获统计学博士学位，曾获 2023 年哈尔滨工业大学“春雁英才计划”，主持黑龙江省博士后面上资助 1 项和参与中国高校产学研创新项目 1 项，发表 10 余篇论文，包括 Journal of Multivariate Analysis 等期刊。

统计计算与优化专场报告 7

报告时间：6 月 8 日上午　报告地点：博智楼 70204

报告题目：A case study on the share holder network effect of stock market data: An SARMA approach

报告人：张榕

摘要：One of the key research problems in financial markets is the investigation of inter-stock dependence. A good understanding in this regard is crucial for portfolio optimization. To this end, various econometric models have been proposed. Most of them assume that the random noise associated with each subject is independent. However, dependence might still exist within this random noise. Ignoring this valuable information might lead to biased estimations and inaccurate predictions. In this article, we study a spatial autoregressive moving average model with exogenous covariates. Spatial dependence from both response and random noise is considered simultaneously. A quasi-maximum likelihood estimator is developed, and the estimated parameters are shown to be consistent and asymptotically normal. We then conduct an extensive analysis of the proposed method by applying it to the Chinese stock market data.

报告人简介：张榕，女，汉族，1992 年生人，籍贯云南丽江，中共党员，2023 年毕业于云南大学，获理学博士学位，为云南大学唐年胜老师和北京大学王汉生老师联合执导的博士。目前为云南大学数学与统计学院讲师，论文发表于中国科学（数学），参加过多个科研项目，研究方向为研究方向为大规模数据分析、复杂网络数据分析。

统计计算与优化专场报告 8

报告时间：6 月 8 日上午　报告地点：博智楼 70204

报告题目：隐私约束下的分布式保形预测

报告人：王天真

摘要：保形预测可用于构建不依赖于分布的预测区间，并在有限样本下可以保证预测区间的覆盖率。现有的保形预测方法大多是在集中式环境中开发的。然而，在实际应用中，数据往往由于隐私保护或通信限制而分布在多个机器之间。当不同机器之间不能共享原始数据时，传统的保形预测方法将不再适用。为了解决这一问题，本文提出了一种分布式保形预测方法，该方法利用分布式随机森林而无需共享原始数据。为了构建该方法的预测区间，我们设计了一种新颖的基于二分法的分布式分位数估计算法。此外，我们提出了一种分布式保形分位数回归方法，该方法能够自适应地应对数据的异方差性。为了进一步提升预测精度，我们提出了一种基于密度的分布式保形预测方法，该方法在面对复杂误差结构(如偏态或多峰分布)时，能够生成更窄的预测区间。我们还建立了所提方法的理论性质，重点分析了预测区间覆盖概率的上下界。最后通过数值模拟和实际数据分析，展示了所提方法的有效性。

报告人简介：王天真，华东师范大学统计学院，博士研究生

复杂数据建模专场报告 1

报告时间：6 月 8 日上午　报告地点：博智楼 70205

报告题目：Data integration using covariate summaries from external sources

报告人：张宇谦

摘要：In modern data analysis, information is frequently collected from multiple sources, often leading to challenges such as data heterogeneity and imbalanced sample sizes across datasets. Robust and efficient data integration methods are crucial for improving the generalization and transportability of statistical findings. In this work, we address scenarios where, in addition to having full access to individualized data from a primary source, supplementary covariate information from external sources is also available. While traditional data integration methods typically require individualized covariates from external sources, such requirements can be impractical due to limitations related to accessibility, privacy, storage, and cost. Instead, we propose novel data integration techniques that rely solely on external summary statistics, such as sample means and covariances, to construct robust estimators for the mean outcome under both homogeneous and heterogeneous data settings. Additionally, we extend this framework to causal inference, enabling the estimation of average treatment effects for both generalizability and transportability.

报告人简介：张宇谦，中国人民大学统计与大数据研究院助理教授，博士生导师。2016 年本科毕业于武汉大学，2022 年博士毕业于美国加州大学圣地亚哥分校。主要研究方向包括因果推断、半监督学习、高维统计、机器学习理论、缺失数据、精准医疗等。文章发表于 Annals of Statistics、Biometrika、Information and Inference 等期刊。

复杂数据建模专场报告 2

报告时间：6 月 8 日上午　报告地点：博智楼 70205

报告题目：Exponential synchronization for the complex networks with intermittent event-triggered control under the second-order process

报告人：艾晓辉

摘要：This paper address the exponential synchronization problem of stochastic complex networks driven by the second-order moment processes, with an innovative intermittent event-triggered control (IETC) strategy introduced. In traditional synchronization studies of stochastic complex networks, Brownian motion is predominantly used to simulate white noise. However, white noise features infinite bandwidth and flat power spectral density, which is inconsistent with real physical systems. Therefore, introducing a second-order moment process to model colored noise interference in physical systems is more appropriate.

An IETC strategy is designed, which is implemented during working periods and suspended during resting periods to reduce communication and computational costs. Besides, the number of event triggers is reduced, saving costs and time by introducing the concept of an average control rate. Then, using graph theory and inequality techniques, a sufficient criterion for exponential synchronization of stochastic complex networks under IETC driven by second-order moment processes is obtained. Furthermore, the strategy's ability to avoid Zeno behavior and ensure stable system operation is demonstrated.

To validate the results, numerical simulations under Hopfield neural networks is presented. The results show that under environmental noise, blue light can activate cells, accelerate ion exchange, and enhance information processing, enabling the network to reach synchronization and memory states faster. The IETC strategy significantly improves the synchronization speed, offering new insights for Hopfield neural network research, particularly in potential applications for mental illness treatment.

报告人简介：艾晓辉，东北林业大学副教授，硕导，吉林大学博士后，中国现场统计研究会多元分析应用委员会理事，主要从事随机过程、随机微分方程，复杂网络等研究。发表论文 20 余篇。

复杂数据建模专场报告 3

报告时间：6 月 8 日上午　报告地点：博智楼 70205

报告题目：Estimation and Inference for Nonparametric Expected Shortfall Regression over RKHS

报告人：王月

摘要：Expected shortfall (ES) has emerged as an important metric for characterizing the tail behavior of a random outcome, specifically associated with rarer events that entail severe consequences. In climate science, the threats of flooding and heatwaves loom large, impacting natural environments and human communities. In actuarial studies, a key observation in modeling insurance claim sizes is that features exhibit distinct effects in explaining small and large claims. This article concerns nonparametric expected shortfall regression as a class of statistical methods for tail learning. These methods directly target upper/lower tail averages and will empower practitioners to address complex questions that are beyond the reach of mean regressionbased approaches. Using kernel ridge regression, we introduce a two-step nonparametric ES estimator that involves a plugged-in quantile function estimate without sample-splitting. We provide non-asymptotic estimation and Gaussian approximation error bounds, depending explicitly on the effective dimension,sample size, regularization parameters, and quantile estimation error. To construct pointwise confidence bands, we propose a fast multiplier bootstrap procedure and establish its validity. We demonstrate the finite-sample performance of the proposed methods through numerical experiments and an empirical study aimed at examining the heterogeneous effects of different air pollutants and meteorological factors on average and high PM2.5 concentration. Supplementary materials for this article are available online,including a standardized description of the materials available for reproducing the work.

报告人简介：王月，中国科学技术大学管理学院副教授，2024 年毕业于香港城市大学获得博士学位，主要研究方向分布式学习，非参数统计，稳健性统计。先后在 Accepted by Journal of the American Statistical Association、Accepted by Journal of Computational and Graphical Statistics、Journal of Statistical Planning and Inference 等期刊发表论文。

复杂数据建模专场报告 4

报告时间：6 月 8 日上午　报告地点：博智楼 70205

报告题目：Matrix-quantile factor prediction for generalized matrix-variate regression

报告人：刘永欣

摘要：This paper proposes a latent matrix-factor regression to predict responses that may come from an exponential distribution with high-dimensional matrix-variates. The latent predictor is a low-dimensional factor extracted from the matrix quantile factor model, which provides a comprehensive and robust relationship between the high dimensional covariates and low-rank factor predictor. Our prediction modeling conducts dimension reduction that respects the geometry characteristic of intrinsic two-dimensional structure of the matrix covariate. A two-step algorithm is used to estimate the matrix quantile factor model and the generalized regression. We establish the convergence rate of the estimated matrix coefficient and prediction. Extensive simulation studies show that the prediction capability of the proposed method outperform existing penalized methods and latent matrix model that only extract mean factors. An empirical application illustrates the usefulness of LaGMQFR by the COVID-19 data.

报告人简介：刘永欣，南京审计大学副教授、硕士生导师。2019 毕业于山东大学获得理学博士学位。主要研究领域：高维统计、分位数模型、复杂数据统计分析。2021 年获江苏省“双创博士”，2022 年获南京审计大学第六批“润泽学者”。主持完成国家自然科学基金青年科学基金项目 1 项、江苏省高等学校自然科学研究面上项目 1 项，主持在研国家自然科学基金面上项目 4 项。

复杂数据建模专场报告 5

报告时间：6 月 8 日上午　报告地点：博智楼 70205

报告题目：Matrix-Factor-Augmented Regression

报告人：蔡雄

摘要：As matrix-variate observations are increasingly available, to incorporate the interplay between the multi-cross-sections, we introduce a matrix-factor-augmented regression model (M-FARM) that proposes to predict ahead of time with factors of matrix predictors augmented in the regression. We show that the estimation error in the factor matrices, estimated by the projection procedure in the first step, enters into the estimation error of the regression parameters and the prediction error of the response variable with an asymptotically negligible rate. The central limit theorems of the estimates of the regression parameters are established under some mild conditions. Forecasting intervals with a theoretical guarantee are given. Monte-Carlo simulations justify the theoretical results. We find empirically that the augmented matrix factors do help in forecasting macroeconomic variables relative to the benchmark matrix autoregressive model and vector-factor-augmented regression model (V-FARM).

报告人简介：报告人简介：蔡雄，南京审计大学统计与数据科学学院，讲师（校聘副教授、硕士生导师），研究领域为非参数与半参数统计、函数型数据分析、时间序列分析。2021 年博士毕业于北京工业大学统计学专业。在 Journal of Business & Economic Statistics, Statistica Sinica, Statistics in Medicine 等期刊发表论文 10 余篇。近年来主持项目两项，包括 1 项国家自然科学基金青年项目，参与课题 4 项。

复杂数据建模专场报告 6

报告时间：6 月 8 日上午　报告地点：博智楼 70205

报告题目：Stochastic Approximation MM Algorithms for Multiple Responses Mixed-effects Model and its Application

报告人：黄子健

摘要：Accurate blood pressure (BP) monitoring is crucial for cardiovascular risk stratification, yet conventional cuff-based methods remain constrained by discontinuous measurements, limited physiological specificity, and poor adaptability. Although photoplethysmography (PPG) provides continuous hemodynamic profiling through pulse wave analysis, existing computational approaches fail to resolve critical challenges. We present a Multiple Responses Mixed-Effects Model (MRMM) that synergistically addresses three limitations. First, our multivariate framework explicitly characterizes the covariance structure between systolic and diastolic BP (SBP/DBP) measurements, enhancing physiological fidelity. Second, hierarchical random effects in MRMM precisely quantify inter-subject variations and intra-individual hemodynamic fluctuations. Third, MRMM establishes statistically validated fixed-effect coefficients with an integrated significance testing framework, quantitatively linking characterized PPG waveform features to both SBP and DBP through physiologically interpretable associations. To overcome computational bottlenecks in large-scale analyses, we develop a Stochastic Approximation Maximization-Majorization (SMM) algorithm with accelerated convergence properties, accompanied by theoretical guarantees through asymptotic normality proofs for fixed effects and covariance estimators. Numerical simulations demonstrate our framework's superiority over state-of-the-art methods, achieving 60\% faster computation and 50\% lower RMSE. Clinical validations confirm significant accuracy gains while preserving physiological interpretability. This work establishes a statistical-physiological synthesis paradigm, advancing personalized cardiovascular monitoring through rigorous integration of hemodynamic coupling mechanisms and multi-level biosignal variability.

报告人简介：黄子健，哈尔滨工业大学（深圳）博士后。

复杂数据建模专场报告 7

报告时间：6 月 8 日上午　报告地点：博智楼 70205

报告题目： Efficient distributed estimation for expectile regression in increasing dimensions

报告人： 李晓妍

摘要： In this paper, we introduce an efficient surrogate loss method for large-scale expectile regression in non-randomly distributed scenarios. Specifically, a Poisson subsampling-based distributed asymmetric least squares estimator is proposed. Our theoretical analysis establishes the consistency and asymptotic normality as the dimensionality tends to ifinity, demonstrating that the proposed estimator achieves statistical efficiency comparable to that of the global estimator. A practical three-step algorithm is presented, offering an efficient implementation in practical applications. The proposed estimator exhibits two notable advantages: (i) it is communication-e ˙ icient, utilising all the data but only requiring the transmission of a small subsample and the local gradient from each local site; and (ii) it can effectively adapt to unevenly distributed data and non-randomly stored data. Within the Newton-Raphson algorithm, the initial value and the Hessian matrix are computed with enhanced robustness using the Poisson subsampling-derived subsample than using one local dataset or uniform subsampling-derived subsample. Both simulation studies and empirical results cofirm that the proposed estimator enhanced estimation efficiency relative to existing methods.

报告人简介： 李晓妍，重庆大学，博士研究生。

复杂数据建模专场报告 8

报告时间：6 月 8 日上午　报告地点：博智楼 70205

报告题目：A Nonparametric Control Chart for Monitoring Location based on Type II Censoring data

报告人：孙沣扬

摘要：This paper presents a novel nonparametric control chart based on the Wilcoxon score function specifically designed for monitoring location parameter shifts in processes with Type II censored data. The proposed method leverages rank-based statistics to effectively detect shifts in location parameters without requiring specific distributional assumptions. By adaptively integrating censored observations through modified Wilcoxon rank scores, the chart maintains sensitivity to location parameter deviations while circumventing distributional dependency. The performance of the proposed chart is evaluated using the median run length (MRL) metric through extensive Monte Carlo simulations. Results demonstrate that the Wilcoxon score-based control chart achieves competitive detection capabilities for both small and moderate shifts under Type II censoring.

报告人简介：孙沣扬，辽宁大学，研究生。

复杂数据建模专场报告 9

报告时间：6 月 8 日上午　报告地点：博智楼 70205

报告题目： Iterative Generalised Method of Moments Estimation of Synthetic Difference-in-Differences

报告人：顾瑞芹

摘要：This study proposes an iterative Generalized Method of Moments estimation method for the Synthetic Difference-in-Differences (IGMM− SDID) framework, which is integrated with the Iterated Generalised Method of Moments (IGMM). The IGMM provides a robust method for causal inference in the presence of data heterogeneity and imbalance. The asymptotic normal distribution and variance of the model are consistently estimated through large-sample theoretical analyses, and the SDID is also identified as a special case of the IGMM−SDID from an estimation perspective. The IGMM − SDID method exhibits superior estimation accuracy, robustness, and adaptability compared to the standard SDID and DID methods. Concurrently, our empirical investigation into the correlation between literacy, education and fertility yields conclusions of superior statistical quality compared to the SDID estimates, highlighting the need for more targeted methodologies in policy evaluation.

报告人简介：顾瑞芹，辽宁大学，研究生。

高维数据与随机矩阵专场报告 1

报告时间：6 月 8 日上午　报告地点：博智楼 70206

报告题目：Sparse Asymptotic PCA: Identifying Sparse Latent Factors Across Time Horizon in High-Dimensional Time Series

报告人：高照省

摘要：This paper introduces a novel sparse latent factor modeling framework using sparse asymptotic Principal Component Analysis (APCA) to analyze the co-movements of high-dimensional panel data over time. Unlike existing methods based on sparse PCA, which assume sparsity in the loading matrices, our approach posits sparsity in the factor processes while allowing non-sparse loadings. This is motivated by the fact that financial returns typically exhibit universal and non-sparse exposure to market factors. Unlike the commonly used ℓ_1-relaxation in sparse PCA, the proposed sparse APCA employs a truncated power method to estimate the leading sparse factor and a sequential deflation method for multi-factor cases under ℓ_0-constraints. Furthermore, we develop a data-driven approach to identify the sparsity of risk factors over the time horizon using a novel cross-sectional cross-validation method. We establish the consistency of our estimators under mild conditions as both the dimension N and the sample size T grow. Monte Carlo simulations demonstrate that the proposed method performs well in finite samples. Empirically, we apply our method to daily S\&P 500 stock returns (2004–2016) and identify nine risk factors influencing the stock market.

报告人简介：高照省，电子科技大学数学科学学院特聘教授，主要研究方向包括高维时间序列分析、面板数据统计建模、金融计量经济学及机器学习等领域。

高维数据与随机矩阵专场报告 2

报告时间：6 月 8 日上午　报告地点：博智楼 70206

报告题目：Nonstationary Binary Factor Models: Maximum Likelihood Estimation

报告人：吴彬

摘要：This paper develops an asymptotic theory for general factor models incorporating observable covariates and unobservable factors generated by an integer process. This framework encompasses scenarios where single-indexes are nonstationary and cointegrated. For nonstationary single-indexes, the maximum likelihood estimator (MLE) of the coefficients is consistent under the condition $T^{1/4}\min(\sqrt{N},\sqrt{T})^{-1}\to0$, as both the cross-sectional dimension (N) and the temporal dimension (T) approach infinity. Similarly, the MLE of all nonstationary factors is consistent when $T^{\delta}\min(\sqrt{N},\sqrt{T})^{-1}\to0$; here, δ is determined by the property of the link function. Notably, the limiting distributions of the factors exhibit a novel phenomenon: they are correlated with time (t) due to the convergence of the Hessian matrix to zero as t increase. In the case of cointegrated single-indexes, the MLEs of both factors and coefficients remain consistent, with the rate of convergence improving to $\min(\sqrt{N},\sqrt{T})$. Additionally, a distinct phenomenon arises in the limiting distributions compared to nonstationary single-indexes: the dual rate of convergence of the coefficients increases from $(T^{1/4},T^{3/4})$ to $(T^{1/2},T)$, and the class of limiting distributions changes accordingly. Moreover, the limiting distributions of the factors no longer depend on t in the cointegrated case. Monte Carlo simulations verify the accuracy of the estimates. In an empirical application, we apply our model to the arrival of jumps in financial markets, where we extract the jump factor and find its excellent asset pricing performance.

报告人简介：吴彬，中国科学技术大学管理学院博士后。

高维数据与随机矩阵专场报告 3

报告时间：6 月 8 日上午　报告地点：博智楼 70206

报告题目： Fast Online Change Point Detection for High-dimensional Regression Models

报告人： 王籼入

摘要：This paper presents a fast online algorithm for detecting change points in high-dimensional regression models, a challenging task due to the lack of direct observations of the parameter of interest. We introduce a novel test statistic-based algorithm designed to operate in an online fashion, ensuring that storage requirements and computational complexity per new observation remain independent of the number of previous observations. A key innovation of our approach is the departure from traditional moving window-based methods, which are limited by a fixed window size that can negatively affect detection performance. To address this, we propose an online change point detection method that evaluates all possible window sizes when constructing the test statistics. Although this increases computational complexity, we develop an efficient algorithm that can compute the optimal window size. Theoretical results are provided, showing that the average run length has a lower bound and that the detection delay is bounded from above, ensuring the reliability and effectiveness of the proposed method.

报告人简介： 王籼入，西南财经大学统计学院，副教授。

高维数据与随机矩阵专场报告 4

报告时间：6 月 8 日上午　报告地点：博智楼 70206

报告题目：Multiplier Bootstrap Meets High-dimensional PCA

报告人：虞龙

摘要：In this paper, we examine the feasibility (i.e., both the good advantages and the bad limitations) and the adaptivity (i.e., the potential for beneficial modifications) of employing multiplier bootstrap to analyze the asymptotic distributions of the largest eigenvalues of potentially spiked high-dimensional sample covariance matrices. Our findings and proposed algorithms demonstrate that multiplier bootstrap remains valid, provided the multipliers are appropriately chosen and the bootstrap procedures are applied multiple times with suitable corrections. First, for non-spiked sample covariance matrices, we propose a novel algorithm to replicate the asymptotic distribution (i.e., the Tracy-Widom law) of their largest eigenvalues. This is achieved by repeatedly bootstrapping the entire sample covariance matrix using carefully designed bounded multipliers that satisfy certain concentration properties. We highlight that unbounded multipliers fail in this setting, as the bootstrapped eigenvalues asymptotically follow a Fréchet or Gumbel distribution. Second, for spiked sample covariance matrices, while both bounded and unbounded multipliers can recover the asymptotic normality of the largest eigenvalues, they may introduce additional bias, particularly when the spikes are not strong. To mitigate this, we apply a modified multiplier bootstrap multiple times to correct the bias. Finally, leveraging our modified multiplier bootstrap procedures, we propose a novel and straightforward distribution-based test for selecting common factors in the factor model.

报告人简介：虞龙，现任上海财经大学统计与管理学院助理教授；于 2020 年在复旦大学管理学院获得理学博士学位；2018 年至 2019 年访问美国密歇根大学安娜堡分校统计系完成博士联合培养项目；2020 年-2022 年访问新加坡国立大学统计与数据科学系完成博士后研究员工作。研究方向为多元统计分析、计量经济学、随机矩阵理论等，重点关注基于因子模型的数据建模、数据降维、稳健统计等方向，学术成果发表在 Biometrika, Journal of Econometrics, Journal of Business and Economic Statistics, Bernoulli, Journal of Multivariate Analysis, Biostatistics 等期刊。

高维数据与随机矩阵专场报告 5

报告时间：6 月 8 日上午　报告地点：博智楼 70206

题目：Asymptotic properties of a multicolored random reinforced urn model with an application to multi-armed bandits

报告人：杨丽 (西安交通大学)

摘要：The random self-reinforcement mechanism, characterized by the principle of "the rich get richer", has demonstrated significant utility across various domains. One prominent model embodying this mechanism is the random reinforcement urn model. In this work, we investigate a multicolored, multiple-drawing variant of the random reinforced urn model. We establish the limiting behavior of the normalized urn composition and demonstrate strong convergence upon scaling the counts of each color. Furthermore, we derive strong convergence estimators for the reinforcement means and prove their joint asymptotic normality. Additionally, we explore the parallels between the reinforced mechanisms in random reinforced urn models and multi-armed bandits, addressing hypothesis testing for expected reward in the latter context.

报告人简介：杨丽，西安交通大学统计系的助理教授，属于姜丹丹教授的科研团队。研究方向：罐子模型、高维统计推断。

高维数据与随机矩阵专场报告 6

报告时间：6 月 8 日上午　报告地点：博智楼 70206

报告题目： Application of the KOO Method in High-Dimensional Discriminant Analysis

报告人： 李玉玲

摘要： In high-dimensional discriminant analysis, an essential challenge is to effectively identify and eliminate irrelevant variables to improve classification accuracy and interpretability. Traditional variable selection methods often struggle with scalability and consistency when the number of variables far exceeds the sample size. In this paper, we explore the application of the knock-one-out (KOO) method for variable selection in high-dimensional discriminant analysis. By leveraging tools from random matrix theory, we establish the almost sure limits and central limit theorems of the KOO statistics under mild moment conditions. These results demonstrate the strong consistency of a KOO-based variable selection rule with a general threshold in the high-dimensional setting. Furthermore, to enhance the robustness of the selection procedure, we propose a bootstrap-based thresholding method for the KOO statistics. Simulation studies confirm the superior performance of our approach, showing higher selection accuracy compared to classical criteria such as the Akaike Information Criterion, Bayesian Information Criterion, and Mallow's Cp. Applications to benchmark datasets, including gene expression data and image recognition tasks, further illustrate the practical advantages of the KOO approach in high-dimensional discriminant analysis.

报告人简介： 李玉玲，长春工业大学数学与统计学院讲师。

高维数据与随机矩阵专场报告 7

报告时间：6 月 8 日上午　报告地点：博智楼 70206

报告题目： Asymptotic distributions of four linear hypotheses test statistics under generalized spiked model

报告人：刘芷君

摘要：In this paper, we establish the central limit theorem (CLT) for linear spectral statistics (LSSs) of a large-dimensional generalized spiked sample covariance matrices, where the spiked eigenvalues can be bounded or tend to infinity. Then the new CLT is applied to give asymptotic distributions of linear hypotheses test statistics under generalized spiked model, which can also be referred to as Wilks' likelihood ratio test statistic U, Lawley-Hotelling trace test statistic W and Bartlett-Nanda-Pillai trace test statistic V. Since the complexity of the test functions, in the final asymptotic results, we use series expansions to approximate the theoretical results. We also present asymptotic power functions for U,W,V and make a comparison with Roy's largest root test R under some certain scenarios. Finally, some numerical simulations are provided to confirm the accurate approximation of our asymptotic results.

报告人简介：刘芷君，东北大学理学院讲师，主要研究领域为大维随机矩阵、高维统计分析。

高维数据与随机矩阵专场报告 8

报告时间：6 月 8 日上午　报告地点：博智楼 70206

报告题目： On the rate of convergence in the CLT for LSS of large-dimensional sample covariance matrices；

报告人： 崔健 (东北师范大学)

摘要：The convergence rate is a nature probability problem after the establishment of CLT. A direct issue when applying the CLT is controlling the empirical size given sample size n, which requires convergence rates to quantify the discrepancy between the true and limiting distributions. Let $\bbB_n=\frac{1}{n}\bbT_p^{1/2}\bbX_n\bbX_n^*\bbT_p^{1/2},$, where $\bbX_n= (x_{ij}) $ is a $p \times n$ matrix with independent and identically distributed real or complex entries, $\bbT_p$ is a $p\times p$ nonrandom Hermitian nonnegative definite matrix with spectral norm uniformly bounded in p. In this paper, based on Stein's method, we establish that if $\mathbb{E}|x_{ij}|^{10}<\infty$ and $p/n\to y>0$ as $n\to\infty$, the rate of the normalized linear spectral statistics of the sample covariance matrices $\bbB_n$ converging in distribution to the standard normal distribution is $O(n^{-1/2+\kappa})$ for any fixed $\kappa>0$.

报告人简介： 崔健，东北师范大学教师。

08 中国现场统计研究会简介

中国现场统计研究会是中国科协领导下的，由我国热心于统计学和数据科学工作的专业科技工作者自愿组成的具有公益性，群众性依法登记成立的学术团体，具备法人社团资格，是中国科协的组成部分，是发展我国统计科技事业的一支重要力量。本会的英文译名 Chinese Association for Applied Statistics，简称 CAAS。本会业务主管单位为中国科学技术协会，登记管理机关为中华人民共和国民政部。挂靠单位是北京工业大学。

中国现场统计研究会成立于 1979 年 8 月，是在邓小平同志批准 下成立的全国性学术组织。学会以马克思列宁主义、毛泽东思想、邓 小平理论、“三个代表”重要思想、科学发展观、习近平新时代中国特 色社会主义思想为指导，团结和组织致力于统计学、数据科学及相关 学科的科技工作者，促进统计学科的发展，促进应用统计的普及和发 展，促进统计和数据科学人才的成长和提高，促进统计在工农业、医 药卫生、科学技术及经济等领域和社会科学中的应用，为科技工作者 服务、为创新驱动发展服务、为提高全民科学素质服务、为党和政府 科学决策服务，推动开放型、枢纽型、平台型科协组织建设，成为领 导下团结联系广大科技工作者的社会团体。

现有分支机构（24 个）：试验设计分会、质量分会、医药与生物统计分会、气象水文地质分会、统计调查分会、生存分析分会、工 程概率统计分会、可靠性工程分会、教育统计与管理专业委员会、医药食品优化专业委员会、多元分析应用专业委员会、资源与环境统计 分会、统计综合评价研究分会、高维数据统计分会、空间统计学分会、计算统计分会、经济与金融统计分会、大数据统计分会、数据科学与人工智能分会、旅游大数据分会、风险管理与精算分会、统计交叉科研究分会、因果推断分会、机器学习分会。公开出版刊物（1 种）：《数理统计与管理》（Application of Statistics and Management）是中国科协主管、中国现场统计研究会主办的学术期刊。

本会主旨为弘扬科学精神、捍卫科学尊严、推广先进技术，致力于普及科学知识、传播科学思想和科学方法、开展青少年科学技术教育活动、提高全民族科学文化素质。四十多年来，在各届理事会的带领下，学会广大会员为我国统计学事业的发展做出了应有的贡献。

09 重庆工商大学数学与统计学院简介

重庆工商大学数学与统计学院设有经济统计系、数理统计系、数据科学系、应用数学系、大学数学教学部、统计智能计算与监测重庆市重点实验室等教学科研机构。现有教职工 103 人，专任教师 91 人，其中教授 28 人、副教授 20 人，具有博士学位的教师 71 人,博士生导师 7 人，硕士生导师 48 人，重庆市学术技术带头人 5 人，重庆市学术技术带头人后备人选 2 人，巴渝学者 4 人，重庆市优秀中青年骨干教师 4 人。

“统计学”是重庆市重点学科，学院拥有“统计学”一级学科博士点、“统计学”一级学科硕士点（含经济统计和数理统计两个研究生专业）、“数学”一级学科硕士点和 “应用统计”专业学位点。拥有“经济统计学”“统计学”“数据科学与大数据技术”“数学与应用数学”“数据计算及应用”5 个全日制本科专业，其中“经济统计学”“数学与应用数学”入选国家一流专业建设点，“统计学”“数学与应用数学”“经济统计学”入选重庆市一流专业建设点，“统计学”还是重庆市特色专业，在校本科生、研究生共计学生 1400 余人。学院拥有 3 门市精品课程、2 门市精品资源共享课程、1 门市研究生优质课程、4 门市级一流课程。

近五年教师在《Journal of Functional Analysis》《中国科学》《统计研究》《Discrete Mathematics》等刊物上发表论文 200 余篇，其中被 SCI 收录 180 余篇。主持省部级以上项目 60 余项，其中国家基金项目 20 项，科研经费近 1200 万元；为地方经济社会服务项目 150 余项，经费 2000 万元以上。获省部级以上科研成果奖 8 项。

学院积极开展国内外学术交流与合作，近几年先后召开了“中—加环境核算国际研讨会”、“NBS—OECD 中国供给使用与核算国际研讨会”、全国高端数学论坛“重庆分析研讨会”、中国现场统计研究会资源与环境统计分会第四届会员代表大会等高端学术会议，与美国、新加坡、加拿大及国内著名高校建立了长期友好的学术交流关系。

学院积极组织和引导学生参加课内外科技文化活动，培养学生的创新意识、创新能力。近五年学生获全国及重庆市各类学科竞赛奖 300 多项，其中全国大学生数学建模竞赛、全国大学生市场调查与分析大赛等全国二等及以上奖励 62 项。学生毕业率、授位率和就业率一直保持在 90%以上，历届毕业生均表现出较高的综合素质和专业能力，受到用人单位普遍好评。

10 统计智能计算与监测重庆市重点实验室简介

统计智能计算与监测重庆市实验室是重庆市唯一的统计学科重点实验室，拥有重庆学术技术带头人 2 人，重庆市中青年骨干教师 2 人，重庆市巴渝学者 5 人，博士生导师 7 人，硕士生导师 34 人。教授占比 51.1%，博士占比 82.14%，45 岁以下占比 64.29%。近五年发表论文 200 余篇，主持省部级以上项目 60 余项，其中国家基金项目 11 项；地方服务项目 80 余项，总经费 1000 余万元。研究成果荣获重庆市科学技术奖、重庆市社会科学优秀成果奖、重庆市发展研究奖等奖项。实验室发挥自身优势，为国家及区域发展战略提供有力的支撑，为第七次全国人口普查设计事后质量抽查工作方案，国内率先完成的“更高质量就业”评估标准，助力重庆两江新区建成国内首个“更高质量就业”城区。

实验室依托重庆工商大学重庆市重点学科和猪八戒股份有限公司大数据技术力量，通过实验室的建设，形成西部地区的数据科学与应用统计领域研究中心；培养一批应用统计相关领域创新型人才创新团队；产出一批围绕立足重庆、面向全国经济社会发展需求的成果；加强省级重点学科“统计学”以及相关特色鲜明的“交叉学科生态群落”建设，使实验室在科学研究、人才培育、社会服务等方面发挥优势学科引领作用，建成重庆乃至西部地区特色鲜明、水平先进的统计学高地，助力西部科学城和国家数字经济试验区建设。

实验室的“十四五”目标是聚焦数据科学与统计学中的重大理论问题，深入开展大数据模型理论、数字经济统计方法、生态环境统计模型等领域的基础和应用研究，并实现研究成果在经济、社会、环境等领域的应用转化，为政府和企业提供决策和咨询服务，使实验室成为重庆市一流、全国知名的研究中心。

11 参会人员

会议便笺

会议便笺


重慶工商大學
Chongqing Technology and Business University