2025年大数据技术与金融科技创新发展学术会议

会议手册

2025年9月13-14日

中国•南宁

主办单位：全国工业统计学教学研究会金融科技与大数据技术分会

管理科学与工程学会金融与风险管理分会

承办单位：广西财经学院中国—东盟统计学院、金融与保险学院

协办单位：华东师范大学统计交叉科学研究院　广西科达信科技有限公司

欢迎致辞

今年是推动金融科技与大数据技术深度融合、服务经济高质量发展的关键之年。为更好地促进学术交流与技术应用，全国工业统计学教学研究会金融科技与大数据技术分会与管理科学与工程学会金融与风险管理分会将于 2025 年 9 月 13 日至 9 月 14 日联合举办“2025 年大数据技术与金融科技创新发展学术会议”。

此次会议由全国工业统计学教学研究会金融科技与大数据技术分会、管理科学与工程学会金融与风险管理分会联合主办，广西财经学院中国—东盟统计学院和金融与保险学院承办，华东师范大学统计交叉科学研究院和广西科达信科技有限公司协办。本次会议旨在汇聚全国大数据技术与金融科技领域的专家学者、教育工作者及行业精英，将围绕机器学习算法、大数据统计分析、金融计量学、资产定价与风险管理、数字金融创新等核心议题展开深入交流。同时，会议将回顾金融科技与大数据技术融合的最新进展，展望未来发展方向，共同推动中国金融科技理论与实践的创新发展，为金融业数字化转型与风险防控提供坚实的智力支撑。

会务组

刘旭 18721376590 邮箱:liu.xu@sufe.edu.cn

蒋丽 13617811801 邮箱:244845923@qq.com

孙瑞 15651030390 邮箱:sunrui0390@126.com

钱琳 15201708957 邮箱:qianlin86@163.com

会议信息

酒店

维也纳国际酒店（动物园店）

地址：广西壮族自治区南宁市西乡塘区大学东路 158 号动物园对面（地铁一号线动物园 D 出口）

会议时间、地点及餐食

日期	时间	主要内容	地点
9 月 12 日	14:00-21:00	报到注册	维也纳酒店一楼
		办理入住	维也纳酒店一楼
	17:30-18:30	晚餐	维也纳酒店三楼维景厅
9 月 13 日	7:30-8:30	早餐	维也纳酒店三楼维景厅
	8:30-9:00	开幕式	维也纳酒店六楼维也纳厅
	9:00-10:30	大会报告 1，2，3	
	10:30-10:50	茶歇	
	10:50-12:20	大会报告 4，5，6	
	12:20-13:30	午餐	维也纳酒店三楼维景厅
	13:30-15:00	大会报告 7，8，9	维也纳酒店六楼维也纳厅
	15:00-15:20	茶歇	维也纳酒店六楼维也纳厅
	15:20-17:40	分组报告（一）	维也纳酒店六楼维也纳厅
	15:20-18:00	分组报告（二）	维也纳酒店四楼莫扎特厅
	15:20-18:00	分组报告（三）	维也纳酒店四楼锦程厅
	18:00-20:00	晚餐	维也纳酒店三楼维景厅
9 月 14 日	7:30-8:30	早餐	维也纳酒店三楼维景厅
	8:30-11:50	分组报告（四）	维也纳酒店六楼维也纳厅
	8:30-11:30	分组报告（五）	维也纳酒店四楼莫扎特厅
	8:30-11:30	分组报告（六）	维也纳酒店四楼锦程厅
	9:50-10:10	茶歇	维也纳酒店四楼肖邦厅
	12:00-13:00	午餐	维也纳酒店三楼维景厅

会议日程安排

9 月 13 日上午 8：30-9：00 开幕式

地点：维也纳酒店六楼维也纳厅

时间	内容	主持人
8:30-9:00 开幕式	主持人介绍与会嘉宾	**唐红祥** 广西财经学院 党委常委、副校长
	校领导致辞 广西财经学院党委副书记、校长 **范祚军教授**	
	会议召集人致辞 华东师范大学统计学院院长 **周勇教授**	
	合影留念（原地）	

9 月 13 日上午 9：00-15：20 大会报告

地点：维也纳酒店六楼维也纳厅

时间	内容	主持人
9:00-9:30 大会报告 1	**报告人：王维国教授** （原东北财经大学党委常委、副校长，现任东北财经大学统计学部主任） **题目：**数字金融发展的碳减排效应——来自家庭消费侧的微观调查证据	**方方** 华东师范大学 统计学院
9:30-10:00 大会报告 2	**报告人：涂云东教授** （北京大学博雅特聘教授） **题目：** The Factor Tree	
10:00-10:30 大会报告 3	**报告人：石芸教授** （华东师范大学） **题 目：** Beyond Forward-Looking: Enhancing Market Timing with Option-Implied Probability Weighting	
10:30-10:50	**茶歇**	

9 月 13 日上午 9：00-15：20 大会报告

地点：维也纳酒店六楼维也纳厅

时间	内容	主持人
10:50-11:20 大会报告 4	**报告人：朱仲义教授** （复旦大学） **题目：** Transfer Learning for High-dimensional Regression	**张军舰** 广西师范大学数学与统计学院
11:20-11:50 大会报告 5	**报告人：刘旭教授** （上海财经大学/广西财经学院） **题目：** Statistical Modeling and Inference in High-dimen--sional Data	
11:50-12:20 大会报告 6	**报告人：张新雨教授** （中国科学院数学与系统科学研究院） **题目：** Portfolio Optimization via Transfer Learning	
12:20-13:30	**午餐：** 维也纳酒店三楼维景厅	
13:30-14:00 大会报告 7	**报告人：陈雪蓉教授** （西南财经大学） **题目：** Robust group identification and membership Prediction	**曹杰** 合肥工业大学管理学院
14:00-14:30 大会报告 8	**报告人：喻达磊教授** （西安交通大学） **题目：** Unified optimal model averaging with a general loss function based on cross-validation	
14:30-15:00 大会报告 9	**报告人：严晓东教授** （西安交通大学） **题目：** 最优动态决策序列和静态数据共生的统计推断	
15:00-15:20	**茶歇**	

9 月 13 日下午 15：20-18：00 分组报告（一）

地点：维也纳酒店六楼维也纳厅

组别	时间	内容	主持人
Session 1	15:20-15:40	**范佳怡**（上海财经大学） 题目：Bilevel Network Estimation with a Deep Neural Network under Hierarchical Structured Sparsity	**吴梦云** 上海财经大学 统计与数据科学学院
	15:40-16:00	**段园家**（上海财经大学） 题目：Modeling Heterogeneous Recurrent Processes with Informative Censoring: A Threshold-Based Approach	
	16:00-16:20	**周蒙**（上海财经大学） 题目：Integrated Bayesian non-parametric spatial modeling for cross-sample identification of spatially variable features	
	16:20-16:40	**谭颂华**（中国科学技术大学） 题目：Linear Functional Quantile AR process	
Session 2	16:40-17:00	**曾奕程**（中山大学） 题目：A Theory-Driven Approach to Inner Product Matrix Estimation for Incomplete Data: An Eigenvalue Perspective	**朱学虎** 西安交通大学 数学与统计学院
	17:00-17:20	**朱学虎**（西安交通大学） 题目：Testing for large-dimensional covariance matrix under differential privacy	
	17:20-17:40	**李乐**（广西财经学院） 题目：金融科技对企业破产风险的影响研究——基于中国上市企业的经验证据	
	17:40-18:00	**鞠红梅**（北京物资学院） 题目：基于模糊隶属度与粒子群算法协同优化的双子支持向量机鲁棒分类方法	

9 月 13 日下午 15：20-18：00 分组报告（二）

地点：维也纳酒店四楼莫扎特厅

组别	时间	内容	主持人
Session 3	15:20-15:40	**苏田园**（上海财经大学） 题目：Collective Wisdom: Policy Ave ra-ging, with an Application to the Newsvendor Problem	**石芸** 华东师范大学 统计学院
	15:40-16:00	**颜廷进**（华东师范大学） 题目：Risk-Sensitive Reinforcement L e-arning with Information Costs	
	16:00-16:20	**苏涛**（华东师范大学） 题目：A Flexible Functional Approach to Volatility Prediction	
	16:20-16:40	**王小刚**（北方民族大学） 题目：连续门限分位数自回归模型的股市收益率预测	
Session 4	16:40-17:00	**管欣**（中南财经政法大学） 题目：Subgroup learning in functional regression models under the RKHS framework	**林存洁** 中国人民大学 统计学院
	17:00-17:20	**李鑫桐**（清华大学） 题目：Doubly Robust Estimation of Optimal Individual Treatment Regime in a Semi-supervised Framework	
	17:20-17:40	**陈宝麟**（首都经济贸易大学） 题目：我国二氧化碳排放的影响因素与可持续投资策略研究	
	17:40-18:00	**杨笑然**（山东工商学院） 题目：Large-scale quantile regression with covariates missing at random	

9 月 13 日下午 15：20-18：00 分组报告（三）

地点：维也纳酒店四楼锦程厅

组别	时间	内容	主持人
Session 5	15:20-15:40	**乔高秀**（西南交通大学） 题目：Prediction of asset prices integrating topological data analysis and decomposition approaches: Insights from cryptocurrency, crude oil, and stock markets	**陈雪蓉** 西南财经大学 统计研究中心
	15:40-16:00	**吴量**（西南财经大学） 题目：Stability-based generalization analysis based on stochastic algorithms	
	16:00-16:20	**潘雅婷**（云南财经大学） 题目：基于均值-协方差建模的纵向数据混合增长曲线模型	
	16:20-16:40	**刘斌**（西南财经大学） 题目：图机器学习在产业链、供应链上的应用	
Session 6	16:40-17:00	**康俊卿**（中山大学） 题目：Too Much of a Good Thing? Copy Trading and the Limits of Price Efficiency	**陈峥** 广东工业大学 管理学院
	17:00-17:20	**彭幸春**（武汉理工大学） 题目：Stochastic Stackelberg differential investment and reinsurance game with ambiguous correlation	
	17:20-17:40	**王美玲**（北京物资学院） 题目：基于深度学习的物流快递包裹损坏识别方法研究	
	17:40-18:00	**黄菲**（广西大学） 题目：健康投资的长期回报：少年参加新农合对青年期收入的影响	

9 月 14 日上午 8：30-11：50 分组报告（四）

地点：维也纳酒店六楼维也纳厅

组别	时间	内容	主持人
Session 7	8:30-8:50	**陈欣**（上海立信会计金融学院） 题目：Off-Policy Evaluation with Irre-gularly Spaced, Outcome-Dependent Observation Times	**韩东啸** 南开大学 统计与数据科学学院
	8:50-9:10	**吴捷**（安徽大学） 题目：Transfer Learning for High-dim--ensional Accelerated Failure Time Models	
	9:10-9:30	**逯文琪**（南开大学） 题目：Tensor Additive Quantile Regression	
	9:30-9:50	**闫引桥**（北京工业大学） 题目：Spatially aware adjusted Rand index for evaluating spatial transcriptomics clustering	
	9:50-10:10	**茶歇**	
Session 8	10:10-10:30	**吕茏**（华中师范大学） 题目：Linear spline index regression model: Interpretability, nonlinearity and dimension reduction	**曲连强** 华中师范大学 数学与统计学院
	10:30-10:50	**邓力航**（华中师范大学） 题目：Individualized Treatment Plan Estimation Based on Transfer Learning	
	10:50-11:10	**宋皓月**（华中师范大学） 题目：Semiparametric analysis for paired comparisons with covariates	
	11:10-11:30	**余光英**（广西财经学院） 题目：数字普惠金融如何影响农业新质生产力 ——基于系统 GMM 分析	
	11:30-11:50	**李玉凤**（上海财经大学） 题目：Subgroup Testing for Change-Plane Cox Models in Heterogeneous Credit Risk Data	

9 月 14 日上午 8：30-11：30 分组报告（五）

地点：维也纳酒店四楼莫扎特厅

组别	时间	内容	主持人
Session 9	8:30-8:50	**韩博**（云南大学） 题目：Large-scale survival analysis with a cure fraction	**郝美玲** 对外经济贸易大学 统计学院
	8:50-9:10	**刘晓玉**（暨南大学） 题目：Single-index Semiparametric Transformation Cure Models with Interval-censored Data	
	9:10-9:30	**杨玥含**（中央财经大学） 题目：Target Stratum Analysis in Stratified Randomized Experiments with Transfer Learning and Regression-Adjusted Average Treatment Effect Estimates	
	9:30-9:50	**张思亮**（华东师范大学） 题目：A Latent Variable Framework for Multiple Imputation with Non-ignorable Missingness: Analyzing Perceptions of Social Justice in Europe	
	9:50-10:10	**茶歇**	
Session 10	10:10-10:30	**胥静茹**（云南民族大学） 题目：云南省产业转移承接能力评价及提升路径	**刘常彪** 广西财经学院 中国—东盟统计学院 副教授
	10:30-10:50	**潘生**（云南大学） 题目：Robust transfer regression with corrupted labels	
	10:50-11:10	**马雨菡**（西安交通大学） 题目：A new estimator for conditional expectile-based value-at-risk of a linear predictive regression	
	11:10-11:30	**伍震寰**（云南大学） 题目：Non-Asymptotic Analysis of the Wasserstein Variational Bayes Method	

9 月 14 日上午 8：30-11：30 分组报告（六）

地点：维也纳酒店四楼锦程厅

组别	时间	内容	主持人
Session 11	8:30-8:50	**温建宁**（上海立信会计金融学院） 题目：科创板资产配置及风险管理策略研究	**霍海峰** 广西科技大学理学院 副院长
	8:50-9:10	**覃泓荪**（广西民族大学） 题目：仁义难两全——企业 ESG 表现会抑制企业的社保投入吗?	
	9:10-9:30	**李懿媛**（上海财经大学） 题目：Change-plane Analysis in Functional Response Quantile Regression	
	9:30-9:50	**陈晶莹**（广西财经学院） 题目：数字普惠金融对广西蚕桑产业的影响机制研究	
	9:50-10:10	**茶歇**	
Session 12	10:10-10:30	**许萌萌**（云南大学） 题目:Specification Testing with Com--plex Survey Data	**于桂海** 贵州财经大学 大数据统计学院 副院长
	10:30-10:50	**李惠芳**（广西民族大学） 题目：数字经济发展对地方政府隐性债务风险的影响是非线性的吗?	
	10:50-11:10	**宋涛**（广西民族大学） 题目：政府耐心资本培育对股价同步性的影响研究	
	11:10-11:30	**高斌、韦喜月、覃泓荪**（广西民族大学） 题目：ESG Report Sentiment Analysis and Environmental Violation Prediction of Chinese Listed Companies Based on FinBERT: Integrating Traditional Empirical and Machine Learning Methods	

大会邀请报告专家简介

王维国

数字金融发展的碳减排效应——来自家庭消费侧的微观调查证据

坚持绿色低碳发展是中国式现代化的显著特征，但受技术条件和公众环保意识等多重原因，消费侧碳减排激励明显不足，而数字金融的快速发展则为促进消费侧碳减排创造了良好契机。本文基于微观调查数据测算家庭消费碳排放，构建理论模型并实证分析数字金融发展对家庭消费碳排放的影响。研究发现，数字金融发展有助于降低家庭消费碳排放，而家庭消费结构升级和公众环保意识提高是重要影响途径。但受“数字鸿沟”的影响，数字金融的这一效应主要体现在地区数字基础设施更完善、数字设备可及度更高以及数字技能和偏好更强的家庭。而且数字金融发展也有助于降低邻近地区家庭碳排放，呈现空间溢出特征。此外，数字金融发展还能够推动家庭能源消费结构清洁化转型。本文研究结论为数字化时代下健全绿色消费激励机制，促进经济绿色低碳转型提供政策参考。

王维国，原东北财经大学党委常委、副校长，现任东北财经大学统计学部主任，博士生导师，全国模范教师、国务院特殊津贴专家、国务院学位委员会第八届统计学科评议组成员，教育部经济学类专业教学指导委员会委员，教育部实验室建设与实验教学指导委员会委员，国家级教学团队带头人，中国数量经济学会副会长。担任《数量经济技术经济》《中国管理科学》《系统工程理论与实践》等期刊编委。相关成果发表在 Economic Model、Journal of Evolutionary Economics、EJOR、IJPE 等 SCI/SSCI 国际期刊，以及《经济研究》《管理世界》《统计研究》《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》等国内重要学术期刊 200 余篇。研究领域为经济计量分析、人口资源环境与可持续发展、宏观经济统计分析等。成果获得第九届高等学校科学研究优秀成果奖(人文社会科学)二等奖。

朱仲义

Transfer Learning for High-dimensional Regression

Information from related source studies can often enhance the findings of a target study.However, the distribution shift between target and source studies may severely impact theefficiency of knowledge transfer. In the high-dimensional regression setting, existing transfer mainly focus on the parameter shift. In this paper, we focus on the high-dimensional quantile regression with knowledge transfer under three types of distribution shift: parameter shift, covariate shift, and residual shift. We propose a novel transferable set and a new transfer framework to address the above three discrepancies. Nonasymptotic estimation error bounds and source detection consistency are established to validate the availability and superiority of our method in the presence of distribution shift. Additionally, an orthogonal debiased approach is proposed for statistical inference with knowledge transfer, leading to sharper
asymptotic results. Extensive simulation results as well as real data applications further demonstrate the effectiveness of our proposed procedure.

朱仲义，复旦大学统计与数据科学系教授，博士研究生导师；曾任中国概率统计学会第八、九届副理事长，国际著名杂志“Statistica Sinica”副主编；“应用概率统计”，“中国科学：数学”杂志编委；现为国际数理统计学会当选会员，担任“数理统计与管理”杂志编委和国际顶级统计杂志 JASA 的副主编。专业研究方向为：纵向数据(面板数据)模型；分位数回归模型，机器学习等。主持完成国家自然科学基金面上项目七项、国家社会科学基金一项，作为子项目负责人完成国家自然科学基金重点项目二项，重大项目子项目一项，目前主持国家自然科学基金重点项目一项。近几年发表论文 100 多篇(其中包括在国际四大统计和机器学习顶级刊物等 SCI 论文八十多篇)。2015 年获得教育部自然科学二等奖。

涂云东

The Factor Tree

Threshold factor models are pivotal for capturing rapid regime-switching dynamics in high-dimensional time series, yet existing frameworks relying on a single pre-specified threshold variable often suffer from model misspecification and unreliable inferences. This paper introduces a novel factor tree model that integrates classification and regression tree (CART) principles with high-dimensional factor analysis to address structural instabilities driven by multiple threshold variables. The factor tree is constructed via a recursive sample splitting procedure that maximizes reductions in a loss function derived from the second moments of estimated pseudo linear factors. At each step, the algorithm selects the threshold variable and cutoff value yielding the steepest loss reduction, terminating when a data-driven information criterion signals no further improvement. To mitigate overfitting, an information criterion-based node merging algorithm consolidates leaf nodes with identical factor representations. Theoretical analysis establishes consistency in threshold variable selection, threshold estimation, and factor space recovery, supported by extensive Monte Carlo simulations. An empirical application to U.S. financial data demonstrates the factor tree's effectiveness in capturing regime-dependent dynamics, outperforming traditional single-threshold models in decomposing threshold effects and recovering latent factor structures. This framework offers a robust data-driven approach to modeling complex regime transitions in high-dimensional systems.

涂云东，北京大学博雅特聘教授，联合受聘于光华管理学院商务统计与经济计量系和北京大学统计科学中心。入选“日出东方”北大光华青年人才，北京大学优秀博士学位论文指导教师(2017,2021,2024)，北京大学优秀研究生导师(2024)，教育部“长江学者奖励计划”青年长江学者，国家杰出青年科学基金获得者。先后获武汉大学理学学士学位(2004)和经济学硕士学位(2006)、加州大学经济学博士学位(2012，河滨分校)。环亚太青年计量经济学者(YEAP)会议发起人和主要组织者。50 余篇学术论文发表在多个国际国内知名专业杂志。著作教材《时间序列分析》由人民邮电出版社于 2022 年 9 月出版。研究领域涵盖时间序列分析、非参数计量方法、大数据分析、金融计量和预测等。

石芸

Beyond Forward-Looking: Enhancing Market Timing with Option-Implied Probability Weighting

This paper introduces a novel approach to asset allocation by incorporating time-varying probability weighting derived from option-implied distributions. Traditional forward-looking models typically assume risk aversion under Expected Utility Theory, overlooking investors' behavioral biases in probability perception. By embedding a probability weighting mechanism into the pricing kernel, we capture how investors overweight or underweight tail risks over time. Empirically, we apply this method to the S&P 500 index and options market, demonstrating that the adjusted real-world return distributions significantly improve portfolio performance across Sharpe ratios, opportunity cost, and utility-based measures. The gains are especially pronounced during crisis periods, where conventional models often fail. Robustness checks across alternative weighting functions, optimization schemes, and dynamic settings affirm the strategy's stability. These findings underscore the significance of incorporating the time-varying probability weighting feature when utilizing option information for asset allocation. This is a joint work with Shaocong Peng.

石芸，华东师范大学统计学院和统计交叉科学研究院，教授，博士生导师。研究领域为行为金融与金融工程。在国内外专业学术期刊上发表论文 20 余篇，其中包括在 Operations Research，INFORMS Journal on Computing，IEEE Transactions on Automatic Control, Journal of Economic Dynamics & Control,《管理科学学报》等期刊上发表的多篇论文。入选上海市浦江人才计划，主持国家自然科学项目两项。现任中国运筹学会金融工程与金融风险管理分会常务理事，管理科学与工程学会理事，上海市运筹学会青年委员等。

张新雨

Portfolio Optimization via Transfer Learning

Recognizing that asset markets generally exhibit shared informational char acteristics, we develop a portfolio strategy based on transfer learning that lever ages cross-market information to enhance the investment performance of the market of interest. Our strategy asymptotically identifies and utilizes the in formative datasets, selectively incorporating valid information while discarding the misleading information. This enables our strategy to achieve the maximum Sharpe ratio asymptotically, while maintaining a variance no greater than that of the strategy solely relying on the data from the market of interest. The promising performance is demonstrated by numerical studies and empirical analysis of investments in different industries.

张新雨，中国科学院数学与系统科学研究院研究员。长期从事统计和计量经济学理论与应用方面的研究工作，与合作者解决了模型平均研究中的多个难题，并将模型平均与迁移学习、随机森林等方法融合提出了具有创新性的新方法，同时将所提出的预测方法应用于实际问题为相关部门的决策提供了参考依据。先后主持国家自然科学基金委杰出及其延续项目等，曾获中国青年科技奖。

喻达磊

Unified optimal model averaging with a general loss function based on cross-validation

Studying unified model averaging estimation for situations with complicated data structures, we propose a novel model averaging method based on cross-validation (MACV). MACV unifies a large class of new and existing model averaging estimators and covers a very general class of loss functions. Furthermore, to reduce the computational burden caused by the conventional leave-subject/one-out cross validation, we propose a SEcond-order-Approximated Leave-one/subject-out (SEAL) cross validation, which largely improves the computation efficiency. As a useful tool, we extend the Bernstein-type inequality for strongly mixing random variables that are not necessarily identically distributed. In the context of non-independent and non-identically distributed random variables, we establish the unified theory for analyzing the asymptotic behaviors of the proposed MACV and SEAL methods, where the number of candidate models is allowed to diverge with sample size. To demonstrate the breadth of the proposed methodology, we exemplify four optimal model averaging estimators under four important situations, i.e., longitudinal data with discrete responses, within-cluster correlation structure modeling, conditional prediction in spatial data, and quantile regression with a potential correlation structure. We conduct extensive simulation studies and analyze real-data examples to illustrate the advantages of the proposed methods.

喻达磊，西安交通大学数学与统计学院教授，博士生导师，入选国家高层次青年人才计划。研究领域为统计预测、估计理论和统计极限理论等，一些成果发表在 JRSS-B、JASA、JMLR、JBES 和中国科学：数学等期刊上。先后主持三项国家自然科学基金项目和一项国家重点研发计划项目课题。

严晓东

最优动态决策序列和静态数据共生的统计推断

在人工智能飞速发展的时代，传统的统计推断方法，由于对静态数据分布特性或预设模型的依赖，已难以充分应对复杂动态决策场景。本报告指出，任务驱动统计学习范式正带来根本性变革：它将统计推断从单纯的静态数据驱动提升至动态决策序列驱动，核心在于实现最优动态决策序列与静态数据的紧密共生。这意味着，我们不仅要从静态数据中提取信息，更要在此基础上，通过序列化的最优决策迭代，反过来优化和利用数据，形成一个动态反馈的推断闭环。

然而，当前统计学界在这一融合了动态决策与静态数据共生的任务驱动学习领域，尚缺乏一套完备的统计理论支撑。为此，本报告基于非线性期望理论，提出并系统阐释了全新的策略极限理论。该理论建立了“策略大数定律”与“策略中心极限定理”，首次揭示了在任务导向的最优动态决策序列下，其统计行为所遵循的内在规律，为突破传统静态统计推断的局限、实现动态数据下的高效统计推断奠定了坚实的理论基石。

在此框架下，报告将深入探讨策略极限理论在统计估计与统计检验两大核心任务中的数学原理与独特优势，并通过“滴滴出行”和“华为业务”的典型应用案例，具体展现任务驱动统计思维如何有效应对“最优动态决策序列”与“静态数据”共生带来的复杂挑战。实践结果有力证明。

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，荣获“华为火花奖”，“滴滴盖亚学者”，研究方为统计决策、统计推断和统计计算等。学术成果发表在著名期刊 JRSSB,AOS,JASA,JOE 以及人工智能顶级会议 ICML，AAAI，AISTAT 等 47 篇。 在“高等教育出版社出版”以独立主编出版了《机器学习》、《数据科学实践基础-基于 R》两部教材。

陈雪蓉

Robust group identification and membership Prediction

In this paper, we propose a novel quantile regression modeling framework with a latent group structure, allowing samples drawn from a population consisting of groups with different conditional quantiles along with certain covariates. Different from most conventional modeling approaches for group identification, such as finite mixture models and threshold models, our new model is distribution free, allows the same individual to have different group affiliations on different covariates. Furthermore, it permits the group numbers and group structure of regression coefficients to be the same or different for different covariates. We identify the potential group
structure for the quantile regression coefficients using the K-NN fused penalized method and recover the group boundaries using SVM method, after artificially assigning appropriate labels to different groups. The computational burden of our approach is significantly lower than the pairwise fused regularization method in Ma and Huang (2017). Moreover, unlike existing regularization methods, our method can analyze and explain the reasons for grouping and predict the group membership of new individuals based on the estimated group boundary. We establish the theoretical properties of the proposed estimators for group parameters and boundary parameters. Simulation studies and a real data analysis illustrate that the proposed methods perform well.

陈雪蓉，西南财经大学青年杰出教授、博士生导师，国家级青年人才计划入选者，省级高层次人才入选者。中科院数学与系统科学研究院博士（联合培养），美国密苏里大学统计系、乔治城大学生物统计博士后，美国密歇根大学、香港城市大学、香港大学访问学者。论文发表于 JASA，JBES，JCGS 等统计学、计量经济学权威期刊。主持国家自然科学基金面上项目 2 项、青年项目、国家自然科学基金重点项目子课题、国家重点研发计划课题子课题各 1 项。曾荣获教育部“第八届高等学校科学研究优秀成果奖青年成果奖”。中国应用统计学会理事，资源与环境分会等分会常务理事。

刘旭

Statistical modeling and inference in high-dimensional data

High-dimensional data is frequently collected in the big-data age, including healthcare, agriculture, economy. This talk investigates the statistical modeling and testing for high-dimensional data. First, we develop response best-subset selector and tensor-decomposition-based methods for multivariate high-dimensional models. Second, we provide ridge-based methods to estimate error's variance in high-dimensional data. Third, we construct efficient confidence intervals and propose hypothesis test method for high-dimensional quantile regressions, including the cases when the high or low-dimensional nuisance parameter presents.

刘旭，上海财经大学统计与数据科学学院的教授，担任广西财经学院中国—东盟统计学院副院长（挂职）。研究兴趣为复杂大数据分析、人工智能大模型、深度学习，及其在经济管理中的应用。在国际顶级统计学、计量经济学以及人工智能期刊 Journal of the American Statistical Association, Biometrika, Journal of Econometrics, Journal of Machine Learning Research 发表多篇论文。主持两项国家自然科学基金面上项目，负责一项国家自然科学基金重点项目子课题。获得上海市第十六届哲学社会科学优秀成果二等奖。

分组报告摘要

Session one：

Modern Modeling for High-Dimensional and Structured Data

Bilevel Network Estimation with a Deep Neural Network under Hierarchical Structured Sparsity

范佳怡（上海财经大学）

Accurate network estimation serves as the cornerstone for understanding complex systems across scientific domains, spanning from decoding gene regulatory networks in systems biology to identifying social relationship patterns in computational sociology. Modern applications demand methods that simultaneously address two critical challenges: capturing nonlinear dependencies between variables and reconstructing inherent hierarchical structures where higher-level entities coordinate lower-level components (e.g., functional pathways organizing gene clusters). Traditional Gaussian graphical models fundamentally fail in these aspects due to their restrictive linear assumptions and flat network representations. We propose DeepBLNet, a deep learning framework for bi-level network inference. The core innovation lies in hierarchical selection layers that enforce structural consistency between high-level coordinating groups and their constituent low-level connections via adaptive sparsity constraints. This architecture is integrated with compositional deep architectures that learn cross-level association patterns through constrained nonlinear transformations, explicitly preserving hierarchical dependencies while overcoming the representational limitations of linear methods. Crucially, we establish formal theoretical guarantees for the consistent recovery of both high-level connections and their internal low-level structures under general statistical regimes. Extensive validation demonstrates DeepBLNet's effectiveness across synthetic and real-world scenarios, achieving superior F1 scores compared to competitive methods and particularly benefiting complex systems analysis through its interpretable bi-level structure discovery.

Modeling Heterogeneous Recurrent Processes with Informative Censoring: A Threshold-Based Approach

段园家（上海财经大学）

Recurrent event data are commonly encountered in longitudinal follow-up studies and provide valuable insights into disease progression. Additionally, the observation of recurrent events may be terminated by informative dropouts or failure events, such as death. Despite the extensive literature on this data structure, few studies have addressed the complex population heterogeneity in recurrent event processes, which may not be

adequately captured by conventional regression models. In this paper, we propose a single-index threshold multiplicative intensity (SIT-MI) model to identify latent subgroups by determining whether a linear combination of classification covariates exceeds certain threshold values. Under the Poisson assumption for recurrent process, we construct a smoothed objective function to estimate the model parameters. Under mild regularity conditions, we establish the consistency of all parameter estimators and show that the slope parameter estimators are asymptotically normal. An extensive simulation study shows that the proposed methods perform well in practical applications. The approach is applied to real-world data from a study on readmission among chronic heart failure patients, obtained from the clinical data repository of the University of Virginia Health System.

Integrated Bayesian non-parametric spatial modeling for cross-sample identification of spatially variable features

周　蒙（上海财经大学）

Accurate network estimation serves as the cornerstone for understanding complex systems across scientific domains, spanning from decoding gene regulatory networks in systems biology to identifying social relationship patterns in computational sociology. Modern applications demand methods that simultaneously address two critical challenges: capturing nonlinear dependencies between variables and reconstructing inherent hierarchical structures where higher-level entities coordinate lower-level components (e.g., functional pathways organizing gene clusters). Traditional Gaussian graphical models fundamentally fail in these aspects due to their restrictive linear assumptions and flat network representations. We propose DeepBLNet, a deep learning framework for bi-level network inference. The core innovation lies in hierarchical selection layers that enforce structural consistency between high-level coordinating groups and their constituent low-level connections via adaptive sparsity constraints. This architecture is integrated with compositional deep architectures that learn cross-level association patterns through constrained nonlinear transformations, explicitly preserving hierarchical dependencies while overcoming the representational limitations of linear methods. Crucially, we establish formal theoretical guarantees for the consistent recovery of both high-level connections and their internal low-level structures under general statistical regimes. Extensive validation demonstrates DeepBLNet's effectiveness across synthetic and real-world scenarios, achieving superior F1 scores compared to competitive methods and particularly benefiting complex systems analysis through its interpretable bi-level structure discovery.

Linear Functional Quantile AR process

谭颂华（中国科学技术大学）

This paper proposes a novel linear functional quantile autoregressive (AR) model, in which traditional time series is studied and the functional predictor is unobserved. Specifically, the conditional quantile function is affected by the past conditional quantile function. This proposed model can capture the functional relationship within two conditional distributions, and can be viewed as a generalization of CaViaR model. We propose a simple and effective three-step approach of quantile regression estimation for our model based on functional principal component analysis (fPCA). Unlike functional data analysis, which usually focuses on the convergence rate of the estimator, we investigate the strong approximations for the three-step estimator. A Monte Carlo experiment is conducted to investigate the finite sample performance of the proposed estimator and its asymptotic behavior. An empirical example is presented to illustrate the usefulness of the new method.

Session two

A Theory-Driven Approach to Inner Product Matrix Estimation for Incomplete Data: An Eigenvalue Perspective

曾奕程（中山大学）

Addressing the critical challenge of data incompleteness in inner product matrix estimation, we introduce a novel eigenvalue correction method designed to precisely reconstruct true inner product matrices from incomplete data. Utilizing random matrix theory, our method adjusts the eigenvalue distribution of the estimated inner product matrix to align with the ground truth. This approach significantly reduces estimation errors for both inner product matrices and the associated Euclidean distance matrices, thereby enhancing the effectiveness of similarity searches on incomplete data. Our method surpasses traditional data imputation and similarity calibration techniques in both maximum inner product search and nearest neighbor search tasks, demonstrating marked advancements in managing incomplete data.

Testing for large-dimensional covariance matrix under differential privacy

朱学虎（西安交通大学）

The increasing prevalence of high-dimensional data across various applications has raised significant privacy concerns in statistical inference. In this paper, we propose a differentially private integrated test statistic for testing large-dimensional covariance structures, enabling accurate statistical insights while safeguarding privacy. First, we analyze the global sensitivity of sample eigenvalues for sub-Gaussian populations, where our method bypasses the commonly assumed boundedness of data covariates. For sufficiently large sample size, the privatized statistic guarantees privacy with high probability. Furthermore, when the ratio of dimension to sample size, $d/n \to y \in (0, \infty)$, the privatized test is asymptotically distribution-free with well-known critical values, and detects the local alternative hypotheses distinct from the null at the fastest rate of $1/\sqrt{n}$. Extensive numerical studies on synthetic and real data showcase the validity and powerfulness of our proposed method.

金融科技对企业破产风险的影响研究——基于中国上市企业的经验证据

李　乐（广西财经学院）

在复杂多变的环境中，金融科技对企业风险的影响机制还有待深入研究。本文以2011—2022 年中国沪深 A 股上市公司数据作为研究样本，探究金融科技对企业破产风险的影响及作用机制。研究结果表明，地区金融科技发展可以显著降低企业破产风险，且主要通过信息生态重塑、优化金融资源配给和数字化转型赋能三种核心渠道抑制企业破产风险。进一步分析发现，对于企业家精神、供应链金融水平与企业国际化较高的企业，地区金融科技发展抑制企业破产风险的效果更为显著。从地方监管和宏观政策来看，在金融监管投入较低的地区以及环境不确定性较高与中美贸易摩擦背景下，地区金融科技发展可以更为有效地降低企业破产风险。此外，企业风险的降低有利于提高企业 ESG 水平，且金融科技在强化企业风险缓释对 ESG 促进作用方面发挥着关键的协同创新功能。本研究拓展了金融科技与企业破产风险的研究边界，也为企业在复杂多变的国际贸易环境中提升风险管理能力，抵御企业破产危机提供启示。本研究从金融创新视角为企业破产风险的研究与管理提供了新的启示。

基于模糊隶属度与粒子群算法协同优化的双子支持向量机鲁棒分类方法

鞠红梅（北京物资学院）

现实世界分类任务中的噪声、异常值及类别失衡问题会削弱机器学习模型的鲁棒性. 为此, 本文提出了 PSO-FTSVM, 一种与粒子群算法(PSO)相结合的模糊隶属度驱动的双子支持向量机(TWSVM)混合框架. 该模型将基于类中心距离和核密度估计的方法相结合, 引入了混合模糊隶属度函数, 能够依据样本的全局分布与局部密度动态分配隶属度, 从而增强了关键样本的代表性; 同时利用 PSO 对模型关键参数进行自适应优化, 以提升其泛化能力. 在 10 个数据集上的实验表明, PSO-FTSVM 平均准确率达到 92.011%±1.863, 较基准 TWSVM 和 FTSVM 分别提升 14.080%和 10.587%. 该模型在处理高度不平衡数据集时表现优异且在噪声环境中仍保持稳定性能, 显示出卓越的鲁棒性.

Session three

Collective Wisdom: Policy Averaging, with an Application to the Newsvendor Problem

苏田园（上海财经大学）

We propose a Policy Averaging Approach (PAA) that synthesizes the strengths of existing approaches to create more reliable, flexible and justifiable policies for stochastic optimization problems. An important component of the PAA is risk diversification to reduce the randomness of policies. A second component emulates model averaging from statistics. A third component involves using cross-validation to diversify and optimize weights among candidate policies. We demonstrate the use of the PAA for the news-vendor problem. For that problem, model-based approaches typically use specific and potentially unreliable assumptions of either independently and identically distributed (i.i.d.) demand or feature-dependent demand with covariates or autoregressive functions. Data-driven approaches, including sample averaging and the use of functions of covariates to set order quantities, typically suffer from overfitting and provide limited insights to justify recommended policies. By integrating concepts from statistics and finance, the PAA avoids these problems. We show using theoretical analysis, a simulation study, and an empirical study, that the PAA outperforms all those earlier approaches. The demonstrated benefits of the PAA include reduced expected cost, more stable performance, and improved insights to justify recommendations. Extensions to consider tail risk and the use of stratified sampling are discussed. Beyond the newsvendor problem, the PAA is applicable to a wide variety of decision-making problems under uncertainty.

Risk-Sensitive Reinforcement Learning with Information Costs

颜廷进（华东师范大学）

We study a provably optimal exploration problem in risk-sensitive reinforcement learning (RL) with information acquisition costs. The agent can pay for signals revealing information about future states and act based on them in a finite-horizon Markov decision process (MDP). We investigate a general objective class named optimized certainty equivalence,which includes popular risk measures such as conditional value-at-risk, variance, and entropic risk. To handle both signal acquisition and action selection, we introduce a unified probabilistic decision policy and characterize the policy space under both constrained and unconstrained signal acquisition settings. We propose a novel bonus-driven value iteration algorithm for tabular MDPs and establish its regret bounds.

A Flexible Functional Approach to Volatility Prediction

苏　涛（华东师范大学）

This paper introduces a novel functional volatility forecasting approach that significantly extends the functional method, fVP, proposed by Tan et al.(2024). First, we accommodate the presence of jumps in both price and volatility dynamics, as widely documented in the financial econometrics literature. Second, our method considerably relaxes the restrictive structure imposed by fVP, which models volatility as a sum of independent smooth and nonsmooth components.Instead, we assume a general Ito semimartingale framework for the volatility process. Third, unlike fVP, our approach does not rely on the log transformation of squared high-frequency returns,thereby avoiding the additional complications associated with zero returns. Simulation and empirical evidence both confirm the proposed method's strong finite-sample performance.

连续门限分位数自回归模型的股市收益率预测

王小刚（北方民族大学）

连续门限分位数自回归模型既能捕捉模型中存在的非线性门限效应，同时又能精准刻画分布的尾部形态，为揭示数据分布的非线性、异质性及相关性等特征提供了重要研究工具。本文基于 CTQAR 模型研究了股票收益率的动态影响机制，通过格点搜索法得到了门限位置估计，运用加权 CUSUM 统计量检验了门限存在性，利用 AIC 信息准则解决了模型定阶问题，给出了门限分位数预测条件分布及密度预测方法。基于滚动窗口法的模拟结果表明估计方法具有高预测精度和可靠性。实证研究表明：CTQAR 模型不仅能够揭示股票收益率的门限效应、异质性及相关性，并且能提供比连续门限回归模型和分位数回归模型更高的预测精度，利用条件密度预测全面刻画分布特征，为金融市场稳定运行、揭示投资者决策行为提供实证支撑。

Session four：统计前沿方法探索与多领域应用实践

Subgroup learning in functional regression models under the RKHS framework

管　欣（中南财经政法大学）

Motivated by the inherent heterogeneity observed in functional data applications, this paper aims to develop subgroup learning methods for functional data analysis. We propose a functional regression model with a change hyperplane to capture heterogeneous relationships within subgroups. An iterative algorithm is developed to estimate unknown parameters and identify the subgroup structure defined by the hyperplane. The asymptotic theory for the coefficient functions is established using a vector- valued reproducing kernel Hilbert space method, and the asymptotic properties of the grouping estimators are derived through a smoothing technique. For statistical inference, we propose a supremum-type score test statistic to test the existence of subgroups and establish its asymptotic distribution under both the null hypothesis and the local alternative hypothesis. Numerical studies demonstrate the finite sample performance of the proposed estimation and testing methods. Finally, we illustrate the practical application using COVID-19 and air quality datasets.

Doubly Robust Estimation of Optimal Individual Treatment Regime in a Semi-supervised Framework

李鑫桐（清华大学）

In many health-care datasets like the electronic health record (EHR) dataset, collecting labeled data can be a laborious and expensive task, resulting in a scarcity of labeled data while unlabeled data is already available. This has sparked a growing interest in developing methods to leverage the abundant unlabeled data. We thus develop several types of semi-supervised (SS) methods for estimating optimal individulized treatment regime (ITR) that utilize both labeled and unlabeled data in a general model-free framework, with efficiency gains compared to supervised estimation methods. Our proposed method first utilizes a flexible imputation technique through single index kernel smoothing to exploit the unlabeled data, which performs well even in cases of multidimensional covariates, with a follow-up estimation to determine the optimal ITR by directly optimizing the imputed value function. Additionally, in cases where the propensity score function is unknown like in observational studies, we also develop a doubly robust SS estimation method based on a class of monotonic index models. Our estimators are shown to be consistent with the cube root convergence rate and exhibit a nonstandard asymptotic distribution characterized as the maximizer of a centered Gaussian process with a quadratic drift. Simulation studies demonstrate the efficiency and robustness of the proposed methods compared to supervised approach in finite samples. Additionally, a practical example from the ACTG 175 study illustrates its real-

world application.

我国二氧化碳排放的影响因素与可持续投资策略研究

陈宝麟（首都经济贸易大学）

随着全球气候变化越来越严重，CO_2排放成了导致气候变化的主要因素。而中国作为全球最大的碳排放国，经济增长和碳排放之间的关系仍然很紧密，而且不同地区的排放差异也很大。因此，研究各个地区和行业碳排放的规律，并制定不同的减排策略，以更好进行可继续投资，具有深刻的理论意义和现实价值。本研究通过分位数回归和 Expectile 回归方法，分析了中国农业、工业和服务业三大产业的CO_2排放驱动因素，以及这些因素是如何在不同地区产生不同影响的。我们采用了 2001 到 2022 年间中国 30 个省份的相关数据，结合经济增长、能源效率、城市化、财政能力、工业化等因素，分析它们对CO_2排放的影响。基于不同分位数和 Expectile 的角度，揭示了这些驱动因素在不同排放水平下的作用。研究结果表明，高排放地区和低排放地区的主要驱动因素差异很大，因此建议政策指定应该根据这些差异来进行有针对性的调整，才能达到更精确的减排效果，制定更完善的可持续投资则策略。

Large-scale quantile regression with covariates missing at random

杨笑然（山东工商学院）

Large-scale data has become increasingly prevalent across various domains due to modern technological advancements. Analyzing such an extensive dataset poses inherent challenges as it often exhibits significant skewness, heteroscedastic variance, and incomplete data. A widely adopted method for studying heterogeneous data is quantile regression (QR)｝. In this article, we fit a linear quantile regression model for large-scale data with randomly missing covariates. We integrate information from multiple machines by constructing a communication-efficient surrogate loss (CSL) function.This function combines a weighting method to address missing data with a convolution-type smoothing procedure to deal with the non-smooth quantile regression loss function. Under mild conditions, our proposed estimator demonstrates consistency and asymptotic normality. The proposed approach is evaluated through extensive simulation studies and applied to the Survey data from the Behavioral Risk Factor Surveillance System (BRFSS).

Session five

Prediction of asset prices integrating topological data analysis and decomposition approaches: Insights from cryptocurrency, crude oil, and stock markets

乔高秀（西南交通大学）

This study proposes a novel hybrid forecasting framework that integrates topological data analysis (TDA), data decomposition techniques, and deep learning, aiming to enhance prediction accuracy across diverse financial markets. Focusing on four representative market types, including cryptocurrencies, crude oil futures, stock indices and individual stocks, we extract six topological features (e.g., persistence entropy) from cross-asset prices within each market, capturing the interconnected relationships and structural correlations across assets, which are often overlooked by traditional methods. Then the predicted asset prices and topological features are decomposed by VMD, EMD and CEEMDAN approaches, respectively. The backpropagation (BP) and long short-term memory (LSTM) approaches are applied to the decomposed series, with predictions in each decomposed series aggregated to obtain the final predicted prices. The empirical results demonstrate that our newly constructed hybrid models (TDA-VMD-BP and TDA-VMD-LSTM) achieve higher prediction accuracy than comparison models and exhibit strong robustness and applicability across multiple market types. This research opens a new research path for topology-decomposition integrated models in the field of financial forecasting.

Stability-based generalization analysis based on stochastic algorithms

吴　量（西南财经大学）

In this talk, we want to introduce a fundamental analysis framework in learning theory based on stochastic algorithms, mainly including the excess generalization error and its decomposition. As an example, we further introduce the excess generalization of randomized coordinate descent (RCD) for pairwise learning. The excess generalization in this case is first analyzed by leveraging the powerful tool of algorithmic stability. This theoretical analysis develops optimal generalization bounds by trading off estimation and optimization, and further provides a principled guideline on how to stop the algorithm appropriately for the best generalization.

基于均值-协方差建模的纵向数据混合增长曲线模型

潘雅婷（云南财经大学）

混合增长曲线模型通过数据驱动的方式预测增长曲线模型中未知的组别矩阵，从而将纵向数据中无法观察的异质性纳入讨论。现有方法通常假设相同轨迹的观测值是独立的或其协方差是某种特定结构，再对各个子群体的均值建模，鲜少有研究探索异质性协方差结构的建模。我们引入了一个具有增长曲线模型框架的高斯混合联合模型，对均值和协方差同时建模，展示了个体内的相关性在纵向数据聚类时的重要性。迭代重加权最小二乘 EM（IRLSEM）算法被用来解决模型参数估计问题。在所有类别中，不同的均值轨迹和协方差结构都能够被有效地识别出来。数值模拟结果表明，在引入协方差建模后我们的方法表现良好，并提供了更准确的聚类结果。实际数据分析也说明了所提出方法的有效性。

图机器学习在产业链、供应链上的应用

刘　斌（西南财经大学）

当前产业链与供应链呈现复杂网络特征，传统分析方法难以精准刻画其内在运行规律，也无法有效量化外部冲击（如资源断供、物流停滞等）的传播效应，给产业调控与风险预警带来挑战。我们从图机器学习视角，结合多源数据构建动态图模型，探讨了产业链、供应链的内在运行经济学机制，以及外部冲击后的冲击效应传播机制，明确了网络结构对机制的关键影响。基于上述发现，我们在产业链性能预测、供应链断裂预测、投资组合及期货价格预测等下游任务中开展实践：产业链性能预测误差较传统模型降低 15%-20%，供应链断裂提前预警准确率超 85%，投资组合收益与期货价格预测稳定性也显著提升。从预测效果角度，有效验证了两种机制的科学性与实用性，为产业优化、风险防控及金融服务实体经济提供技术支持

Session six

Too Much of a Good Thing? Copy Trading and the Limits of Price Efficiency

康俊卿（中山大学）

We develop a rational expectations equilibrium (REE) model for copy trading, where households either trade directly as passive traders or become active followers by searching for informed leaders. Active followers determine their replication position on leaders' strategies instead of forming their own demand. When active followers neglect the replication risks, lower search costs increase funds allocated to informed leaders without necessarily enhancing market efficiency. While copy trading introduces informed capital by emulating leaders' strategies, it simultaneously injects systematic noise due to common replication errors. With sufficiently high neglect of replication risks and copy errors, excessive capital flows can amplify this noise, undermining the beneficial effects of informed capital.

Stochastic Stackelberg differential investment and reinsurance game with ambiguous correlation

彭幸春（武汉理工大学）

This paper investigates a Stackelberg investment and reinsurance game with the ambiguous correlation between the financial and insurance markets. The ambiguous correlation is modeled by constructing a set of non-equivalent prior probability measures under the G-framework. The reinsurer acts as the leader charging reinsurance premiums, while the insurer acts as the follower participating in proportional reinsurance. Both the reinsurer and the insurer invest in the financial market with the aim to maximize the expected utility of their terminal wealth under the worst-case scenario. Based on the dynamic programming principle under G-Brownian motion, the Stackelberg equilibrium strategies are derived by solving the Hamilton-Jacobi-Bellman-Isaacs (HJBI) equations. Through theoretical analysis and numerical simulation, the influence of ambiguity on the Stackelberg equilibrium strategies is analyzed. We find that both the insurer and reinsurer adopt more conservative strategies in the presence of market correlation ambiguity, with the insurer being more significantly affected by such ambiguity, while the reinsurer is less influenced. When both the insurer and the reinsurer maintain investment in the risky asset, the ambiguity faced by the insurer leads to higher reinsurance premiums, whereas the ambiguity faced by the reinsurer results in lower reinsurance premiums.

基于深度学习的物流快递包裹损坏识别方法研究

王美玲（北京物资学院）

随着电子商务与物流行业的快速发展，快递包裹在运输与分拣环节中发生损坏的情况日益增多，高效、准确的损坏识别已成为提升服务质量和降低经济损失的关键。针对传统人工检测成本高、难以满足大规模包裹检测需求的问题，本文提出一种基于深度学习的快递包裹损坏自动识别方法。 研究选用 Kaggle 平台“Damaged and Intact Packages”图像数据集，采用迁移学习策略，利用 ImageNet 预训练参数初始化网络，以提高小样本条件下的收敛速度与识别精度。首先分别对 ResNet34、GoogLeNet 与 MobileNet V2 三种卷积神经网络进行对比实验，这三种模型分别代表了残差网络、多尺度融合网络和轻量化网络的典型设计思路，具有较强的代表性和可比性。结果表明 ResNet34 性能表现最佳，测试集准确率为 88.54%。在此基础上，本文对 ResNet34 进行优化改进：在深层特征提取阶段引入 CBAM 注意力机制，增强模型对损坏区域的关注；采用 LeakyReLU 激活函数替代 ReLU，缓解神经元失活并提升特征表达能力；在网络末端增加瓶颈结构与 DropBlock 正则化，以提升泛化性能；并结合标签平滑交叉熵、AdamW 优化器及余弦退火学习率调度，进一步优化训练。 实验结果表明，改进后的 ResNet34 在测试集上的 Accuracy 达到 95.83%，较原始 ResNet34 提升 7.29%，Precision、Recall 与 F1-Score 等指标亦显著改善。研究结果验证了该方法在快递包裹损坏识别中的有效性与优越性，为智能物流系统的自动化检测提供了可行的技术路径，对提升行业自动化水平和服务质量具有重要意义。

健康投资的长期回报：少年参加新农合对青年期收入的影响

黄　菲（广西大学）

提升国民人力资本是国家发展的核心战略，而少年时期的健康是其后续人力资本积累的关键基础。考察基础医疗保障制度对个体长期经济成就的影响具有重要的政策意义。本文利用中国家庭追踪调查（CFPS）的面板数据，采用双向固定效应模型，通过控制不随时间变化的省级区域特征以及共同的时间趋势冲击，识别了少年时期参与新型农村合作医疗对个体青年阶段劳动收入的长期影响。研究发现，少年时期参保使个体青年期收入显著提升了约 62.0%，该结论在多种稳健性检验（包括替换被解释变量、使用不同样本等）下依然成立。机制分析表明，新农合主要通过改善个体健康状况、保障家庭资源两个路径，促进长期人力资本的积累。本文的研究结论揭示了公共健康保险在促进人力资本积累和阻断贫困代际传递方面的长期价值，为持续优化和完善我国全民健康保障体系提供了经验证据。

Session seven

Off-Policy Evaluation with Irregularly Spaced, Outcome-Dependent Observation Times

陈　欣（上海立信会计金融学院）

While classic off-policy evaluation (OPE) literature typically assumes decision time points to be evenly spaced for simplicity, in many real-world scenarios such as those involving user-initiated visits, this assumption may not hold. To address OPE with irregularly spaced and potentially outcome-dependent decision times, this paper introduces a novel reinforcement learning (RL) framework that incorporates not only the state-action process but also an observation process dictating the time points at which actions are made. The proposed framework generalizes the semi-Markov decision process (SMDP) framework by allowing the gap times between observations to be correlated with previous states, actions, and gap times, and thereby better aligns with dynamic policies with outcome-dependent observation times. Furthermore, beyond the classic policy value function based on cumulative reward, we introduce a new policy evaluation criterion based on integrated reward, which is more suitable for settings with continuous observation times. Statistical inference of the newly defined value function is provided by positing a modulated renewal process model on the observation times. The validity of the proposed policy evaluation method is further supported by theoretical results, simulation studies, and a real-world application using electronic health records (EHR) to evaluate periodontal disease treatments.

Transfer Learning for High-dimensional Accelerated Failure Time Models

吴　捷（安徽大学）

Transfer learning is a powerful tool to leverage information from different but related source domain to boost performance of the target task. Despite the substantial developments in various high-dimensional data analysis, survival data are insufficiently accounted by current transfer learning approaches. Random censoring mechanism, the population heterogeneity of multiple datasets and high dimensionality of predictors pose major challenges. How to properly conduct transfer learning procedure onto survival datasets to improve the statistical inference at the target and simultaneously guarantee theoretical properties remain unclear. In this paper, we propose a new transfer learning method for the high-dimensional right-censored survival data under accelerated failure time (AFT) model. The proposed method applies mean imputation

techniques to adjust for censoring data, thus facilitating the subsequent pooling and debiasing transfer procedures. We show that under mild conditions, the proposed method achieves lower error bounds over the existing non-transfer learning methods even there exists obstacles from censoring. We also develop an algorithm based on sample splitting to detect informative source datasets and prove that the selected informative sources is consistent with the true one. Extensive simulation results and a real data case concerning TCGA micro-array gene expression are reported to support our theoretical claims and to demonstrate the empirical usefulness of the proposed transfer learning method. To the best of our knowledge, this is the first work on transfer learning for high-dimensional accelerated failure time modelling with theoretical guarantees.

Tensor Additive Quantile Regression

逯文琪（南开大学）

Additive nonparametric models are increasingly favored for analyzing tensor data, offering a flexible and parsimonious approach. This paper introduces a tensor additive quantile regression model, aiming to provide a more robust and detailed understanding of how covariates influence the response in tensor data. The component functions are estimated using basis function approximations. We stack the splines as an additional tensor dimension, which allows us to leverage the tensor structure and apply Tucker decomposition for dimension reduction. In high-dimensional settings, we propose a sparse tensor additive quantile regression model that incorporates a group penalty for variable selection. A key challenge addressed is connecting sparse tensor elements within the structure of Tucker decomposition. We propose an innovative approach that identifies a broader set of relevant features than the oracle, while enabling efficient algorithms to navigate the high-dimensional space. We establish the large sample properties of the proposed estimators, evaluate the finite sample performance of our method through Monte Carlo simulations and demonstrate its application on real-world datasets, including stock market and head pose image data.

Spatially aware adjusted Rand index for evaluating spatial transcriptomics clustering

闫引桥（北京工业大学）

The spatial transcriptomics (ST) clustering plays a crucial role in elucidating the tissue spatial heterogeneity. An accurate ST clustering result can greatly benefit downstream biological analyses. As various ST clustering approaches are proposed in recent years,

comparing their clustering accuracy becomes important in benchmarking studies. However, the widely used metric, adjusted Rand index (ARI), totally ignores the spatial information in ST data, which prevents ARI from fully evaluating spatial ST clustering methods. We propose a spatially aware Rand index (spRI) as well as spARI that incorporate the spatial distance information. Specifically, when comparing two partitions, spRI provides a disagreement object pair with a weight relying on the distance of the two objects, whereas Rand index assigns a zero weight to it. This spatially aware feature of spRI adaptively differentiates disagreement object pairs based on their distinct distances, providing a useful evaluation metric that favors spatial coherence of clustering. The spARI is obtained by adjusting spRI for random chances such that its expectation takes zero under an appropriate null model. Statistical properties of spRI and spARI are discussed. The applications to simulation study and two ST datasets demonstrate the improved utilities of spARI compared to ARI in evaluating ST clustering methods. The R package to compute the proposed spRI and spARI is available at

Session eight

Linear spline index regression model: Interpretability, nonlinearity and dimension reduction

吕　茏（华中师范大学）

Inspired by the complexity of certain real-world datasets, this article introduces a novel flexible linear spline index regression model. The model posits piecewise linear effects of an index on the response, with continuous changes occurring at knots. Significantly, it possesses the interpretability of linear models, captures nonlinear effects similar to nonparametric models, and achieves dimension reduction like single-index regression models. In the model, both the locations and number of knots are unknown. This not only enhances the adaptability of the model in practical applications but also presents challenges in estimating and inferring the unknown parameters. Here, the knots represent points where the effects of covariates on the response exhibit significant changes while still maintaining continuity. Combining the penalized approach with convolution techniques, we propose a new method to simultaneously estimate the unknown parameters and the number of knots. The proposed method allows the number of knots to diverge with the sample size. We demonstrate that the proposed estimators can identify the number of knots with a probability approaching one and estimate the coefficients as efficiently as if the number of knots is known in advance. We also introduce a procedure to test the presence of knots. Simulation studies and two real datasets are employed to assess the finite sample performance of the proposed method.

Individualized Treatment Plan Estimation Based on Transfer Learning

邓力航（华中师范大学）

Transfer learning has gained significant attention across various fields, addressing the challenge of limited individual study data for predictions. In precision medicine, formulating accurate individualized treatment plans is particularly crucial. This paper primarily explores estimation methods for individualized treatment plans using transfer learning. We constructed multiple auxiliary models utilizing shared parameters from other datasets. A smooth concordance index function was employed to derive candidate model parameters, combined with a leave-one-out cross-validation criterion to determine the optimal weights for the averaging process. Our proposed method allows for different model forms among the auxiliary models. Theoretically, we demonstrate that under mild conditions, the proposed method achieves the highest smooth concordance index asymptotically when the main model is misspecified and achieves

model weight consistency when the main model is correctly specified. We also prove the convergence of the proposed algorithm. Simulation and empirical studies demonstrate that the transfer learning-based method for estimating individualized treatment plans proposed in this paper exhibits superior performance in reducing bias and enhancing predictive accuracy.

Semiparametric analysis for paired comparisons with covariates

宋皓月（华中师范大学）

Statistical inference in parametric models (e.g., the Bradley–Terry model and its variants) for paired-comparison data has been explored in the high-dimensional regime, in which the number of items involving in paired comparisons diverges. However, parametric models are highly susceptible to model misspecification. To relax the assumption of known distributions and provide flexibility, we propose a semiparametric framework for modeling the merits of items and covariate effects (e.g., home-field advantage) by introducing latent random variables with unspecified distributions. As the number of parameters increases with the number of items, semiparametric inference is highly nontrivial. To address this issue, we employ a kernel-based least squares approach to estimate all unknown parameters. When each pair of items has a fixed number of comparisons and the number of items tends to infinity, we prove the consistency of all resulting estimators and derive their asymptotic normal distributions. To the best of our knowledge, this is the first study to conduct a semiparametric analysis of paired comparisons with an increasing dimension. We conduct simulations to evaluate the finite-sample performance of the proposed method and illustrate its practical utility by analyzing an NBA dataset.

数字普惠金融如何影响农业新质生产力 ——基于系统 GMM 分析

余光英（广西财经学院）

数字普惠金融有效缓解了农业发展过程中的融资难等问题，赋予农业新质生产力发展动力。文章基于 2012-2022 省级面板数据，构建系统 GMM 模型及门槛效应模型，分析数 字普惠金融对农业新质生产力的影响效应。结果显示：数字普惠金融促进农业新质生产力的 发展；数字普惠金融对农业新质生产力的影响因人口密集度不同而不同；数字普惠金融影响 农业新质生产力存在门槛效应。政策建议：优化数字普惠金融服务体系，进一步促进农业新 质生产力发展；实施差异化的数字普惠金融政策，因地制宜赋能农业新质生产力发展。

Subgroup Testing for Change-Plane Cox Models in Heterogeneous Credit Risk Data

李玉凤（上海财经大学）

Change-plane analysis has emerged as an effective tool to detect subgroups with distinct effects on the response of interest. Notably, the change-plane Cox model has gained significance in the subgroup analysis of survival data. Testing for the existence of a change plane can provide valuable insights into optimal decisions for the specific subgroups. However, classical supremum testing methods often suffer from limited efficiency in practical settings. To address this drawback, we propose a novel testing procedure designed to enhance statistical power. Our approach calculates the weighted average of the squared score test statistic (WAST) over the space of parameter that defines the subgroup, significantly improving power in practice. Moreover, we derive the asymptotic distributions of the test statistic under both the null and local alternative hypotheses. The performance of the proposed method is evaluated by extensive simulation studies, exhibiting more accurate size control and higher power than existing testing approaches. Additionally, we illustrate the practical application of our approach using the Lending club loan dataset and the German credit dataset, showing its ability to identify subgroups with different default risks.

Session night

Large-scale survival analysis with a cure fraction

韩　博（云南大学）

With the advent of massive survival data with a cure fraction, large-scale regression for analyzing the effects of risk factors on a general population has become an emerging challenge. This article proposes a new probability-weighted method for estimation and inference for semiparametric cure regression models. We develop a flexible formulation of the mixture cure model consisting of the model-free incidence and the latency assumed by the semiparametric proportional hazards model. The susceptible probability assesses the concordance between the observations and the latency. With the susceptible probability as weight, we propose a weighted estimating equation method in a small-scale setting. Robust nonparametric estimation of the weight permits stable implementation of the estimation of regression parameters. A recursive probability-weighted estimation method based on data blocks with smaller sizes is further proposed, which achieves computational and memory efficiency in a large-scale or online setting. Asymptotic properties of the proposed estimators are established. We conduct simulation studies and a real data application to demonstrate the empirical performance of the proposed method.

Single-index Semiparametric Transformation Cure Models with Interval-censored Data

刘晓玉（暨南大学）

In this paper, we propose a class of flexible single-index semiparametric transformation cure models for interval-censored data, where a single-index model and a semiparametric transformation model are utilized for the uncured and conditional survival probability, respectively, encompassing both the proportional hazards cure and proportional odds cure models as specific cases. We approximate the single-index function and cumulative baseline hazard functions via the kernel technique and splines, respectively, and develop a computationally feasible expectation-maximisation (EM) algorithm, facilitated by a four-layer gamma-frailty passion data augmentation. Simulation studies demonstrate the satisfactory performance of our proposed method, compared to the spline-based approach and the classical logistic-based mixture cure models. The application of the proposed methodology is illustrated using the Alzheimer's dataset.

Target Stratum Analysis in Stratified Randomized Experiments with Transfer Learning and Regression-Adjusted Average Treatment Effect Estimates

杨玥含（中央财经大学）

This paper introduces a novel covariate adjustment method based on design-based analysis, leveraging information from auxiliary strata to enhance the efficiency of average treatment effect estimation in sparse settings. The proposed estimator is particularly effective for specific targets with limited samples or studies by utilizing insights from related but potentially correlated auxiliary strata. We develop an estimator for the average treatment effect and provide a Neyman-type conservative variance estimator for hypothesis testing and confidence interval construction. We establish the asymptotic properties of the estimator and outline conditions under which it outperforms or matches existing methods in efficiency. The proposed approach achieves higher precision and narrower confidence intervals than conventional estimators. Its advantages are further demonstrated through simulations and empirical data analysis.

A Latent Variable Framework for Multiple Imputation with Non-ignorable Missingness: Analyzing Perceptions of Social Justice in Europe

张思亮（华东师范大学）

This paper proposes a general multiple imputation approach for analyzing large-scale data with missing values. An imputation model is derived from a joint distribution induced by a latent variable model, which can flexibly capture associations among variables of mixed types. The model also allows for missingness which depends on the latent variables and is thus non-ignorable with respect to the observed data. We develop a frequentist multiple imputation method for this framework and provide asymptotic theory that establishes valid inference for a broad class of analysis models. Simulation studies confirm the method's theoretical properties and robust practical performance. The procedure is applied to a cross-national analysis of individuals' perceptions of justice and fairness of income distributions in their societies, using data from the European Social Survey which has substantial nonresponse. The analysis demonstrates that failing to account for non-ignorable missingness can yield biased conclusions; for instance, complete-case analysis is shown to exaggerate the correlation between personal income and perceived fairness of income distributions in society.

Session ten

云南省产业转移承接能力评价及提升路径

胥静茹（云南民族大学）

近年来，我国区域发展不平衡问题突出，产业梯度转移成为促进经济协调增长的重要路径。本研究聚焦云南省承接产业转移的能力评价与产业选择策略，旨在破解区域发展失衡背景下产业转移的适配性问题。通过构建涵盖产业吸引力、支撑力、发展力的综合评价体系，结合产业动静集聚指数、熵值法测度及产业梯度系数模型，分析东部沪苏浙闽粤五省市的产业转移动态、云南省 11 个州市的承接能力差异和云南产业梯度。研究发现：东部地区食品制造、文教工美用品、计算机通信设备等产业转移趋势显著；云南省承接能力呈现区域分化，昆明依托基础设施与创新资源稳居首位，曲靖、红河等地区次之，而保山、昭通等边缘地区受制于基建滞后与产业单一，承接能力较弱。结合产业梯度系数筛选，云南宜重点承接食品制造、医药、计算机通信设备等产业。云南省需要差异化推进区域基建与产业配套，破解边缘地区承接瓶颈；构建“科技+产业”双轮驱动机制，提升重点产业附加值；建立“省际-跨境”双向协同机制，优化产业承接生态。

Robust transfer regression with corrupted labels

潘　生（云南大学）

In this paper, we introduce a robust transfer regression method designed to handle corrupted labels in target data, under the scenarios that the corruption affects a substantial portion of the labels and the locations of these corruptions are unknown. Our theoretical analysis decomposes the estimation error into three interpretable components: (1) source data, (2) domain shift, and (3) label corruption. This framework guarantees that our method consistently outperforms target-only estimation. We validate our method through numerical experiments focused on reconstructing corrupted compressed signals, showing robustness even when a high fraction of labels are corrupted, especially when some source data exhibit structural similarities to the target data. Additionally, we apply our method to analyze the association between O6-methylguanine-DNA methyltransferase (MGMT) methylation and gene expression in Glioblastoma (GBM) patients.

A new estimator for conditional expectile-based value-at-risk of a linear predictive regression

马雨菡（西安交通大学）

Since it is the only elicitable law-invariant coherent risk measure, the expectile-based value-at-risk (EVaR) is a recently recommended risk measure in financial risk management. This paper considers the large sample statistical inference problem of conditional EVaR under a linear predictive regression model. Based on the least-squares residuals, we propose a novel least-squares residual estimator for the conditional EVaR of a linear predictive regression. The asymptotic properties of the proposed estimator are investigated in the context of dependence. We illustrate that the proposed estimator is computationally efficient and has desirable finite sample performance through numerical studies and an empirical application to risk assessment.

Non-Asymptotic Analysis of the Wasserstein Variational Bayes Method

伍震寰（云南大学）

This paper develops a novel variational Bayesian framework that minimizes the 1-Wasserstein distance, formulated through the Kantorovich Rubinstein dual representation, between the variational distribution and the posterior. Unlike the Kullback Leibler divergence, the 1-Wasserstein distance provides symmetry, robustness to support mismatch, and a geometry-aware comparison of probability measures. We establish non-asymptotic theoretical guarantees for the resulting Wasserstein variational posterior, deriving explicit contraction rates, tail probability bounds, and Gaussian approximation results relative to the maximum likelihood estimator. These results demonstrate that the 1-Wasserstein variational posterior achieves reliable finite-sample accuracy and retains the concentration properties of the exact posterior. Simulation studies on linear regression and logistic regression further confirm the stability and accuracy of the method compared to mean-field variational Bayes. Overall, this work introduces the Kantorovich Rubinstein perspective into variational inference, establishing a theoretically rigorous and practically effective framework for scalable Bayesian computation, with strong potential for broader applications in high-dimensional statistical modeling.

Session eleven

科创板资产配置及风险管理策略研究

温建宁（上海立信会计金融学院）

从 2018 年 11 月习总书记宣告设立科创板，到 2019 年 7 月科创板成功开板，发展科创板培育科技硬实力，推动科创经济全面提速发展，成为国家竞争力体现的重大战略部署。我们发现从六年前科创板首批 25 家公司共同上市为标志，开启了科创板直接融资支持新质生产力的发展过程，到 2025 年 8 月科创板已成为金融市场重要的投资场所，其中著名的算力芯片股票武纪（688256）成为市场价值投资的标杆，股价因受到市场投资者的一致认可，价格不断上涨越过千元大关，并继续加速涨升，一度达到 1595.88 元，超越老牌价值投资棋手贵州茅台（600519），短暂成为证券市场的第一高价股，打开了重新评价和认知科技股票投资价值的全新研究课题，同时也提出了科创经济时期资产配置的全新实践话题。我们围绕“科技立国、金融强国”的新发展理念，聚焦建设中国式现代化的时代命题，统计了上海及长三角地区集成电路上市公司的融资总额和技术特征等指标，结合上海和长三角科创板和科创经济之间发展现状分析，发现了一些进行资产配置的可行性条件和经济运行特征，给出金融市场投资风险管理的一些策略。

仁义难两全——企业 ESG 表现会抑制企业的社保投入吗？

覃泓荪（ 广西民族大学）

本文以我国上市公司为样本，实证研究企业 ESG（环境、社会与治理）表现对社会保险投入水平的影响。基于资源有限理论和合法性理论，本文提出 ESG 表现可能通过资源分配权衡和合法性动机影响企业社保投入。研究采用固定效应面板回归模型，并引入融资约束的中介效应和成本压力的调节效应进行机制分析。结果发现：第一，企业 ESG 表现与社保投入水平显著负相关，ESG 评级越高的企业，其营业收入占比社保费用投入越低。第二，融资约束在二者关系中发挥部分中介作用，ESG 提升缓解了企业融资约束，从而影响社保投入。第三，成本费用率对 ESG 与社保投入的关系具有显著负向调节效应，高成本压力下 ESG 对社保投入的抑制作用更强。第四，非重污染行业和中西部地区的企业中，ESG 的影响更为显著，而重污染行业、东部地区企业中相关效应较弱。上述结论在更换 ESG 度量、样本分组、剔除特殊年份、加入高维固定效应以及 Heckman 两阶段和工具变量等多种稳健性检验后依然成立。研究建议监管部门完善 ESG 信息披露和评价体系，将员工社保等纳入 ESG 考核，鼓励企业在提升 ESG 绩效的同时切实履行对员工的社保责任，实现环境治理与社会保障的协同发展。研究结论为政策

制定者通过 ESG 引导企业社会责任实践、完善社保政策提供了经验证据支持.

Change-plane Analysis in Functional Response Quantile Regression

李懿媛（上海财经大学）

Change-plane analysis is a pivotal tool for identifying subgroups within a heterogeneous population, yet it presents challenges when applied to functional data.

In this paper, we consider a change-plane model within the framework of functional response quantile regression to identify and test subgroups in non-Gaussian functional responses with scalar predictors. We employ an alternating direction method of multipliers algorithm to estimate the function coefficients and grouping parameters, which enables population division into distinct subgroups.To test for subgroup existence, we develop a weighted average of the squared score test statistic, which has a closed form and reduces computational burden. The asymptotic theory for the estimates is established based on the reproducing kernel Hilbert space, and the asymptotic distributions of the proposed test statistic are derived under both the null and alternative hypotheses.Simulation studies are conducted to evaluate the performance of the proposed approach in identifying and testing subgroups.Additionally, we apply the methods to two real datasets from Chinese stock markets and the COVID-19 pandemic.

数字普惠金融对广西蚕桑产业的影响机制研究

陈晶莹（广西财经学院）

文章以数字普惠金融为切入点，探讨其对广西蚕桑产业发展的推动作用及其作用机制。基于 2011—2022 年广西 12 个地级市的面板数据，构建固定效应模型和中介效应模型进行实证分析。结果表明，数字普惠金融的总指数及其构成的三个维度，即覆盖广度、使用深度和数字化程度均对蚕桑产业发展具有显著的正向影响，此中农业技术进步和基础设施建设发挥了重要的中介作用。进一步分析发现，在蚕桑主产区，数字普惠金融覆盖广度的影响作用比非主产区的要大，而总指数、使用深度和数字化程度的影响作用在非主产区要比主产区的大。为此提出要加强数字普惠金融与蚕桑产业的深度融合，强化政策支持与引导，充分发挥数字普惠金融在不同产区的差异化推动作用的建议。

Session twelve

Specification Testing with Complex Survey Data

许萌萌（云南大学）

Specification analysis for regression models has received considerable methodological attention over the past two decades. Existing approaches, however, remain largely confined to traditional model-based frameworks and are often ill-suited for analyzing complex survey data due to inherently intricate dependency structures induced by complex sampling designs and weighting schemes.

In this paper, we develop a novel specification testing procedure for linear quantile models within the design-based framework for complex survey data. We propose a nonparametric test grounded in an orthogonal projection onto the tangent space of nuisance parameters to assess the correct specification of a linear conditional quantile function over a continuum of quantile levels. The proposed test statistic serves two key purposes: it enables the assessment of model validity for design-based inference regarding the conditional effects on outcome distributions; and it effectively accommodates the survey design features and provides valid inference for both the particular finite population and the related superpopulation. We derive the asymptotic distribution of the test statistic under both the null and alternative hypotheses, showing in particular that the estimation of unknown model parameters has no asymptotic influence on the validity of the proposed test. For practical implementations of the test, we introduce a multiplier bootstrap procedure and establish its theoretical validity. The finite-sample performance of the proposed method is examined via simulation studies, and its practical utility is illustrated through a real data application.

数字经济发展对地方政府隐性债务风险的影响是非线性的吗？

李惠芳（广西民族大学）

本文基于 2011—2022 年我国 256 个地级市的面板数据，采用双向固定效应模型与中介效应模型，实证检验数字经济发展对地方政府隐性债务风险的影响及其作用机制。研究结果表明：数字经济发展对地方政府隐性债务风险呈先促进后抑制的 U 型影响特征，即发展初期会推高债务风险，而随着数字经济水平跨越一定阈值后则呈现显著抑制效应。机制分析发现，数字经济通过非线性路径影响金融发展水平，并经由线性机制作用于财政分权程度，进而对隐性债务风险产生传导效应。异质性分析显示，东部与西部地区影响显著，而中部地区结果不显著且呈现正 U 型关系，该差异可能源于区域政策导向及地方产业结构特征。本研究为理解数字经济与地方政府债务风险之间的复杂关系提供了实证依据，对差异化防

范与化解地方债务风险具有政策启示。

政府耐心资本培育对股价同步性的影响研究

宋　涛（广西民族大学）

积极培育耐心资本既是建设科技强国与金融强国的战略支撑，更是推动经济高质量发展的关键制度安排。本研究选取 2011-2023 年省级政府工作报告与沪深 A 股上市公司数据，采用文本分析法构建区域政府耐心资本培育强度指标，实证检验其对资本市场定价效率的影响。研究发现：政府培育耐心资本显著降低企业股价同步性，该结论在替换变量、调整模型设定等稳健性检验后依然成立。机制分析表明，政策效应通过增强分析师研报跟踪、机构投资者关注度与媒体信息披露三条渠道缓解信息不对称。异质性检验显示，产权性质、行业竞争、行业政策扶持和金融市场发展水平四个因素对培育耐心资本的股价同步减弱效应有显著影响。本文为完善资本市场基础制度提供了新的经验证据。

ESG Report Sentiment Analysis and Environmental Violation Prediction of Chinese Listed Companies Based on FinBERT: Integrating Traditional Empirical and Machine Learning Methods

高斌、韦喜月、覃泓荪（广西民族大学）

This study examines the relationship between sentiment expression in corporate ESGreports and environmental violations using data from Chinese listed companies from 2006to 2023. Employing the FinBERT model for sentiment analysis, we investigate how corporate emotional expression affects environmental violation behavior through a "sentimentcontagion-supervisory pressure-behavioral constraint" mechanism. The research innovatively extends FinBERT sentiment analysis from continuous scores to a five-level sentimentdistribution (extremely positive, positive, neutral, negative, extremely negative) and constructs heterogeneity indicators such as extreme sentiment ratios.

By integrating traditionaleconometric methods with machine learning approaches, we establish a comprehensive empirical framework including mediation effects, moderation effects, and heterogeneity analysis.

Our findings reveal that positive sentiment expression in ESG reports significantly reduces environmental violation probability by enhancing public trust, increasing investor attention, and strengthening regulatory recognition. A one standard deviation increase in sentiment score reduces environmental violation probability by approximately 12.3%, with this effect being more pronounced in companies with high media attention and excellent ESG ratings. The extreme positive sentiment ratio shows a significant negative correlation with violation rates, validating the importance of sentiment distribution heterogeneity. Machine learning models, particularly Random Forest, demonstrate superior performance in environmental violation prediction with 84.1% accuracy and 0.891 AUC, with SHAP value analysis revealing FinBERT sentiment features as the most important predictive factors.

The theoretical contribution of this study lies in constructing a complete "sentimentsupervision-behavior" transmission mechanism. The methodological contribution involvesthe first application of FinBERT multi-level sentiment analysis to the Chinese ESG context. The practical contribution provides AI-based environmental risk early warning tools for regulators and investors. The research validates the information authenticity hypothesis, confirming that positive sentiment expression genuinely reflects superior environmental performance, providing important theoretical and empirical support for improving ESG evaluation systems and environmental regulatory policies.

交通指引

报到地点：维也纳国际酒店(广西大学路动物园地铁站店)
具体位置：广西壮族自治区南宁市西乡塘区大学东路 158 号动物园对面(地铁 1 号线动物园 D 出口)

来程		
方向	交通方式	推荐路线
南宁吴圩机场-酒店	计程车	南宁吴圩机场-维也纳酒店单程约 38 公里，乘坐计程车正常情况下大约需要 40 分钟，车费为 60-80 元左右。
	机场大巴	乘坐机场大巴 3 号线，上车点为 T2 航站楼 6 号出口直达酒店。票价：20/人，全程大约 1 小时 10 分钟。(运营时间：8:15-20:15；大约 45 分钟一辆)。
南宁东站-酒店	计程车	南宁东站-维也纳国际酒店单程约 18 公里，乘坐计程车正常情况下大概需要 22 分钟，车费为 35-45 元左右。
	地铁	站内步行至火车东站地铁站 F 口，乘坐地铁 1 号线（石埠方向，2/人）到达动物园地铁站 D 口，步行 400 米左右到维也纳国际酒店，全程约 50 分钟。(运营时间：6:30-23:00)
南宁站-酒店	计程车	南宁站-维也纳国际酒店单程约 6 公里，乘坐计程车正常情况下大约需要 13 分钟，车费约为 15-25 元左右。
	地铁	站内步行至火车站地铁站 A 口，乘坐地铁 1 号线（石埠方向，2/人）到达动物园地铁站 D 口，步行 400 米左右到维也纳国际酒店，全程约 25 分钟。(运营时间：6:30-23:00)

返程		
方向	交通方式	推荐路线
酒店-南宁吴圩机场	计程车	维也纳酒店-南宁市吴圩机场单程约 38 公里，乘坐计程车正常情况下大约需要 40 分钟，车费为 60-80 元左右。
	机场大巴	乘坐机场大巴 3 号线，上车点酒店门口直达南宁吴圩机场 T2 航站楼 6 号出口。票价：20/人，全程大约 1 小时 10 分钟。(运营时间：7:00-21:00）；大约 45 分钟左右一辆）
酒店-南宁东站	计程车	维也纳国际酒店-南宁东站单程约 22 公里，乘坐计程车正常情况下大概需要 24 分钟，车费为 45-60 元左右。
	地铁	步行 400 米到动物园地铁站 D 口，乘坐地铁 1 号线（火车东站方向，2/人）到达火车东站地铁站 F 口，步行 100 米左右到南宁东站，全程约 50 分钟。(运营时间：6:30-23:00)
酒店-南宁站	计程车	维也纳国际酒店-南宁东站单程约 22 公里，乘坐计程车正常情况下大概需要 24 分钟，车费为 45-60 元左右。
	地铁	步行 400 米到动物园地铁站 D 口，乘坐地铁 1 号线（火车东站方向，2/人）到达火车站地铁站 A 口，步行 200 米左右到南宁站，全程约 25 分钟。(运营时间：6:30-23:00)

注意：1.航班起飞前 45 分钟停止办理值机手续，敬请留意，以便合理安排您的出发时间
2.如您乘坐高铁或动车抵达南宁，建议您选择终点站-南宁站，以节省时间。

笔记

2025年大数据技术与金融科技创新发展学术会议

2025年9月13-14日

中国•南宁

主办单位：全国工业统计学教学研究会金融科技与大数据技术分会

管理科学与工程学会金融与风险管理分会

承办单位：广西财经学院中国－东盟统计学院、金融与保险学院

协办单位：华东师范大学统计交叉科学研究院　广西科达信科技有限公司