

应用统计 与数据科学

Applied Statistics and Data Science



ART AND DESIGN PRESS INC.

(626 810 4480)

119 S Atlantic Blvd, Suite 300D

Monterey Park, CA 91754

Copyright © 2025 by ART AND DESIGN PRESS INC.

Complimentary Copy



编委会名单

主 编



王国长 (Guochang Wang), 暨南大学 (Jinan University),
邮箱: wanggc023@amss.ac.cn
研究兴趣: 函数型数据分析、时间序列、充分性降维、统计与机器学习
functional data analysis, time series, sufficiency dimension reduction, statistics and machine learning

执行副主编



张兴发 (Xingfa Zhang), 广州大学 (Guangzhou University),
邮箱: xingfazhang@gzhu.edu.cn
研究兴趣: 金融统计、环境统计、机器学习
financial statistics, environmental statistics, machine learning

编 委



胡志勇 (Zhiyong Hu), 广州大学 (Guangzhou University)
邮箱: zhyhu65@163.com
研究兴趣: 大数据、人工智能、财务与会计
big data, artificial intelligence, finance and accounting



雷田礼 (Tianli Lei), 深圳职业技术大学 (Shen Zhen Polytechnic University)
邮箱: ltl@szpu.edu.cn
研究兴趣: 数量经济、高职数学教育
auantitative economics, higher vocational mathematics education



舒连杰 (Lianjie Shu), 澳门大学 (University of Macau),
邮箱: ljshu@um.edu.mo
研究兴趣: 量化金融, 统计学习, 质量控制及管理
quantitative finance, statistical learning, quality control and management



王纬 (Wei Wang), 山东财经大学 (Shandong University of Finance and Economics)
邮箱: wangwei_0115@sdufe.edu.cn
研究兴趣: 统计机器学习、应用统计、计量经济
statistical machine learning, applied statistics, econometrics



杨凯 (Kai Yang), 长春工业大学 (Changchun University of Technology)
邮箱: yangkai@ccut.edu.cn
研究兴趣: 时间序列分析、高维数据分析、贝叶斯分析
time series analysis, high-dimensional data analysis, bayesian analysis



周彦 (Yan Zhou), 深圳大学 (Shenzhen University)
邮箱: zhouy1016@163.com
研究兴趣: 生物统计, 数据科学
Biostatistics, data science



李永明 (Yongming Li), 上饶师范学院 (Shangrao Normal University)
邮箱: lym1019@163.com
研究兴趣: 非参数统计, 金融统计, 极限理论与统计大样本性质, 小波方法
nonparametric statistics, financial statistics, limit theory and statistical large sample theory, wavelet method



刘照德 (Zhao de Liu), 广东财经大学 (Guangdong University of Finance & Economics)
邮箱: lzhaode@163.com
研究兴趣: 经济统计、多元统计、计量分析
economic statistics, multivariate statistics, econometric analysis



王小刚 (Xiaogang Wang), 北方民族大学 (North Minzu University)
邮箱: wongone@163.com
研究兴趣: 经济与社会统计
economic and social statistics



严晓东 (Xiaodong Yan), 西安交通大学 (Xi'an Jiaotong University)
邮箱: yanxiaodong@xjtu.edu.cn
研究兴趣: 统计机器学习、智能体统计学
statistical machine learning, agent statistics



朱柯 (Ke Zhu), 香港大学 (The University of Hong Kong)
邮箱: mazhuke@hku.hk
研究兴趣: 统计建模、金融时间序列分析、计量经济、金融大数据
statistical modeling, financial time series analysis, econometrics, financial big data

应用统计与数据科学

Applied Statistics and Data Science

第1卷 第7期 2025年9月刊

主办单位: ART AND DESIGN PRESS INC

(United States)

承办单位: 广州统计师事务所

指导单位: 广东省现场统计学会、广州市统计学会

主 编: 王国长

执行副主编: 张兴发

编 委: 胡志勇、李永明、雷田礼、刘照德、
舒连杰、王小刚、王 纬、严晓东、
杨 凯、朱 柯、周 彦

地 址: 119 S Atlantic Blvd, Suite 300D Monterey
Park, CA 91754

网 址: <https://www.artdesignnp.com>

本刊说明:

凡向本刊所投稿件, 全体作者需签署论文著作权
转让声明书和论文发表承诺书, 声明、承诺及相关事
项如下:

- 作者将论文的复制权、发行权、网络传播权、翻
译权、汇编权、信息网络传播权、改编权等著作
权在世界范围内免费转让给本刊。
- 论文不侵犯他人著作权和其他权利, 否则作者将
承担由此产生的全部责任, 并赔偿由此给出版单
位造成的全部损失。
- 论文署名作者享有该作品的完全著作权, 署名作
者的身份真实。
- 论文未曾以任何形式公开发表过。
- 作者所投本刊稿件, 本刊编辑部拥有修改权。



市场调研 | MARKET RESEARCH

- 001 “职”引未来——福建省 AI 训练师
的市场现状调研 肖婧, 刘柏慧, 郑佳莹, 尚文然, 刘久畅, 李木易
"Job" Leads to the Future — A Study on the Current Market
Situation of AI Trainers in Fujian
Province Xiao Jing, Liu Baihui, Zheng Jiaying, Shang Wenran, Liu Jiuchang, Li Muiyi
- 007 广州市临期食品销售现状及消费者购买
意愿影响因素调查 颜欣如, 陈伟钰, 吴语桐, 李庆驹
A Survey on the Sales Status of Near-Expired Food and Influencing
Factors of Consumers' Purchase
Intention in Guangzhou Yan Xinru, Chen Weiyu, Wu Yutong, Li Qingju
- 013 广东佛山木版年画认知现状调查
与活态传承意愿分析 黄冬燕, 张兆芹, 李爱萍, 冼卓毅
A Survey on the Current Cognitive Status and Willingness for Living
Inheritance of Guangdong Foshan Woodblock
New Year Paintings Huang Dongyan, Zhang Zhaoqin, Li Aiping, Xian Zhuoyi
- 019 当代广州市大学生对疗愈产品的购买意向
及影响机制研究 张桢岚, 张少华, 李思妍, 吴永桢, 李宇轩, 黄校英
A Study on the Purchase Intention and Influencing Mechanisms
of Healing Products Among Contemporary University
Students in Guangzhou Zhang Zhenlan, Zhang Shaohua, Li Siyan, Wu Yongzhen,
Li Yuxuan, Huang Jiaoying
- 024 关于“国潮风”下传统中医药零食消费者
购买意向调查 李凤娇, 曾雪怡, 张俊超, 吴芯怡, 李雯琴, 黄家良
A Survey on the Purchase Intentions of Traditional Chinese Medicine Snacks Among
Consumers under the "National Trend" Li Fengjiao, Zeng Xueyi, Zhang Junchao,
Wu Xinyi, Li Wenqin, Huang Jialiang
- 029 以广州市为例探讨数字人民币的采纳动因
与场景渗透策略 庞晓晴, 何彦哲, 刘韵, 陶钰, 严梓维, 王孟欣
A Study on Adoption Motivations and Scenario Penetration Strategies of
e-CNY: A Case Study of Guangzhou City Pang Xiaoping, He Yanzhe, Liu Yun, Tao Yu,
Yan Ziwei, Wang Mengxin
- 035 中老年“两病”患者基层医疗机构首诊意愿影响因素分析——对广东省中山市
639位中老年“两病”患者的调查 邱伊汶, 张玥, 杨美婵
Analysis of Influencing Factors on the Willingness of Middle-Aged and Elderly Patients
with "Two Diseases" to Seek Initial Treatment at Primary Medical Institutions
—A Survey of 639 Middle-Aged and Elderly Patients with "Two Diseases" in Zhongshan
City, Guangdong Province Qiu Yiwen, Zhang Yue, Yang Meichan

经济与金融统计 | ECONOMIC AND FINANCIAL STATISTICS

- 040 ESG 绩效对农林牧渔业企业融资约束的影响分析——以企业声誉
为中介变量 庄丽凤, 陈晓彤, 黄洁洁, 裴祺阳, 陈永亨, 陈立宇, 周燕
Analysis of the Impact of ESG Performance on the Financing Constraints of Agriculture,
Forestry, Animal Husbandry and Fishery Enterprises — Taking Corporate Reputation as
a Mediating Variable Zhuang Lifeng, Chen Xiaotong, Huang Jiejie, Pei Qiyang,
Chen Yongning, Chen Liyu, Zhou Yan
- 046 化解运营困境: 广州市地下综合管廊 PPP 项目
资金缺口成因与实战对策研究 陈晓坚
Resolving Operational Challenges: A Study on the Causes and Practical
Countermeasures for Funding Shortfalls in Guangzhou's Urban Utility
Tunnel PPP Project Chen Xiaojian

管理统计 | MANAGEMENT STATISTICS

- 050

加强基层辅助调查员管理的几点思考

Reflections on Strengthening the Management of Grassroots Auxiliary Investigators

曹秀平

Cao Xiuping
- 053

政府投资项目跟踪审计与投资效益提升的协同机制探索

Exploration of the Synergistic Mechanism between Government Project Tracking Audit and Investment Efficiency Enhancement

简结倩

Jian Jieqian

统计与教育教学 | STATISTICS AND EDUCATION TEACHING

- 057

大湾区经管类概率统计课程的就业与升学导向教改研究

Research on Employment and Further Study-Oriented Teaching Reform of Probability and Statistics Courses for Economics and Management Majors in the Greater Bay Area

刘超, 彭舒心, 林诗敏, 张钰笛

Liu Chao, Peng Shuxin, Lin Shimin, Zhang Yudi
- 061

高中“双新”驱动的“五维联动”生涯教育实践

The Practice of "Five-Dimensional Linkage" Career Education Driven by "Double New Initiatives" in Senior High Schools

李春波, 陈璟

Li Chunbo, Chen Jing
- 065

融合 Shiny 与大语言模型的数理统计教学创新研究

Innovative Teaching of Mathematical Statistics Integrating Shiny and Large Language Models

赖廷煜, 胡志明

Lai Tingyu, Hu Zhiming

数据科学技术、方法与应用 | DATA SCIENCE TECHNOLOGY, METHODS AND APPLICATIONS

- 070

融合深度学习与动态因子分析的省际经济差异测度分析

Measuring Provincial Economic Disparities with Integrating Deep Learning and Dynamic Factor Analysis

刘家欢, 范金宇

Liu Jiahuan, Fan Jinyu
- 075

基于注意力机制的 Transformer 模型并行计算架构设计研究

Research on Parallel Computing Architecture Design of Transformer Model Based on Attention Mechanism

苏泽宇

Su Zeyu

统计方法及应用 | STATISTICAL METHODS AND APPLICATIONS

- 079

基于日内 GARCH 模型 LAD 估计的混成检验

Portmanteau Test Based on LAD Estimation of Intraday GARCH Model

朱华锋, 陈燕珊, 张兴发

Zhu Huafeng, Chen Yanshan, Zhang Xingfa
- 085

广东省放心消费“双承诺”活动满意度测评报告——以江门市为例

Guangdong Province Assured Consumption "Double Commitment" Activity Satisfaction Evaluation Report — Take Jiangmen City as an Example

王润良

Wang Runliang
- 091

基于高维协方差矩阵估计的投资组合优化策略比较与分析

Comparison and Analysis of Portfolio Strategies Based on High-Dimensional Covariance Matrices Estimation

孙章爽, 张婷, 万宇雷, 王国强

Sun Zhangshuang, Zhang Ting, Wan Yulei, Wang Guoqiang
- 097

基于 T 检验的现场展示教学模式影响规律探索——以统计学专业实验案例教学为例

Exploration of Influence Patterns of Field Demonstration Teaching Mode Based on T-test — Taking Experimental Case Teaching of Statistics as an Example

车金星, 陈锦文, 曹寒问, 董炜, 张毓华, 万冰蓉, 曾康

Che Jinxing, Chen Jinwen, Cao Hanwen, Dong Wei, Zhang Yuhua, Wan Bingrong, Zeng Kang
- 102

关于构建促进高效能社会治理体系的政务数据目录的研究

Research on the Construction of Government Data Directory for Building a High-Performance Social Governance System

谷剑芳

Gu Jianfang

“职”引未来——福建省 AI 训练师的市场现状调研

肖婧, 刘柏慧, 郑佳莹, 尚文然, 刘久畅, 李木易

厦门大学 经济学院, 福建 厦门 361000

DOI:10.61369/ASDS.2025070001

摘 要 : 随着 AI 技术在医疗、金融、制造、教育等领域的深入应用, AI 训练师已成为技术领域的热门职位。本文结合问卷调查和网络爬虫文本挖掘, 探索福建省专本研学生对 AI 训练师的了解程度、潜在从业意愿以及相关企业对 AI 训练师岗位的需求情况。得出结论: 学生对 AI 训练师职业的认知与从业意愿受多因素影响; 福建省 AI 训练师市场供需不平衡; AI 训练师职业发展路径不明确; 企业对 AI 训练师的技能要求与人才储备之间存在差距; AI 训练师职业的薪资与福利水平有待提升。由此, 我们提出优化教育体系; 加强政策支持与引导; 完善职业培训与认证体系; 优化薪资福利体系; 构建职业示范与社群支持体系; 强化企业责任, 推动数据隐私保护与职业发展等建议。

关 键 词 : AI 训练师; 人工智能; 二元 Logistic 回归; 因子分析

"Job" Leads to the Future — A Study on the Current Market Situation of AI Trainers in Fujian Province

Xiao Jing, Liu Baihui, Zheng Jiaying, Shang Wenran, Liu Jiuchang, Li Muiy

School of Economics, Xiamen University, Xiamen, Fujian 361000

Abstract : With the in-depth application of AI technology in the fields of healthcare, finance, manufacturing, education, etc., the AI trainer has become a popular position in the technology field. This paper combines a questionnaire survey and web crawler text mining to explore the degree of understanding of AI trainers, potential willingness to practice, etc., among specialized undergraduate students in Fujian Province and the demand for AI trainer positions in related enterprises. It is concluded that: students' knowledge of AI trainer career and their willingness to practice are affected by multiple factors; the supply and demand of AI trainer market in Fujian Province is unbalanced; the career development path of AI trainer is unclear; there is a gap between the skill requirements of enterprises for AI trainer and the talent pool; the salary and welfare level of the AI trainer career needs to be improved. As a result, we propose optimizing the education system; strengthening policy support and guidance; improving the vocational training and certification system; optimizing the salary and welfare system; constructing a vocational demonstration and community support system; and reinforcing the responsibility of enterprises to promote the protection of data privacy and career development.

Keywords : AI trainer; artificial intelligence; binary Logistic regression; factor analysis

引言

人工智能技术的广泛应用深刻地改变了各行各业, 尤其对劳动力市场产生了巨大影响。AI 逐渐替代常规型传统职业的同时, 也创造出一些新职业、新岗位, AI 训练师就是其中的代表。AI 训练师是使用智能训练软件, 在人工智能产品实际使用过程中进行数据库管理、算法参数设置、人机交互设计、性能测试跟踪及其他辅助作业的人员^[1]。AI 训练师的市场潜力巨大。《中国青年报》报道中显示, 随着数据行业的发展, 未来人工智能行业的门槛会越来越高, 数据标注可能将从劳动密集型产业转向技术型产业^[2]。因此, 随着数据分类的精细化、复杂化, 人工智能行业需要更加专业、高学历的数据标注师来提供数据服务。

本文聚焦于福建省 AI 训练师市场现状, 通过对 AI 训练师的供给与需求情况的分析, 为相关政策和职业培训提供相应的策略建议。

项目信息: 第十五届全国大学生市场调查大赛福建赛区获奖作品。

作者简介:

肖婧, 女, 安徽安庆人, 厦门大学经济学院在读本科生, 研究方向: 财政学;

刘柏慧, 女, 河南郑州人, 厦门大学经济学院在读本科生, 研究方向: 金融学;

郑佳莹, 女, 福建泉州人, 厦门大学经济学院在读本科生, 研究方向: 财政学;

尚文然, 女, 上海人, 厦门大学经济学院在读本科生, 研究方向: 经济统计学;

刘久畅, 女, 河南漯河人, 厦门大学经济学院在读本科生, 研究方向: 经济统计学;

李木易, 女, 安徽滁州人, 厦门大学经济学院教授, 研究方向: 统计学。

一、调查方案设计与实施

（一）方案设计

本文基于文献综述、问卷调查及网络舆论调查方法，调查 AI 训练师潜在从业者的就业认知及意愿，并针对该职业的市场需求进行深入分析，旨在针对 AI 训练师的技术需求为当代大学生就业提供参考依据，同时为相关部门和企业优化市场布局、提高服务质量提供对策建议。

基于文献综述，我们了解到学界和媒体普遍认可 AI 技术对就业的创造效应^[9]。但具体到 AI 训练师领域，从业者技能要求与人才储备脱节等问题仍待解决，这为问卷设计内容提供思路。

基于问卷调查，本团队广泛查阅统计数据和相关文献，设计了针对高校学生为代表的潜在从业者的多维度问卷，通过网络发放问卷并对多填、漏填等无效问卷进行筛查，以保证问卷数据真实性和有效性。

基于网络舆论分析，通过在社交媒体上爬取 AI 训练师已从业者对自身职业发展现状的评论，分析该职业的市场现状以及发展前途。

（二）调查内容

AI 训练师在国家职业技能标准中被明确定义为在 AI 产品实际使用过程中进行数据库管理、算法参数设置等其他辅助作业的人员。国家早在 2017 年已将人工智能发展上升为国家战略，近年来发布的政策文件中也在不断加强对 AI 训练师市场的监督力度^[4]。

本文以福建省为例，选取福建省专科、本科及硕士学生为主要调查对象，旨在分析福建省高校学生对 AI 训练师的认知程度、职业兴趣以及潜在就业意愿，并探究影响潜在从业者从业意愿的因素。

（三）抽样设计

1. 预调查

在正式调查之前，通过在线问卷平台（问卷星）发布预调查问卷，主要针对本校学生进行。通过收集一定数量的问卷，对问卷设置的问题进行修正，并为问卷的正式开展打基础。

2. 样本量的确定

本次调查选取福建省内各高校的专科生、本科生和研究生作为研究对象。采用分层抽样和二阶段不等概率抽样的调查方式，根据福建省内不同地区的经济发展水平、教育资源分布以及 AI 产业的发展状况等标准进行抽样，得到的结果相对真实合理。依据样本量计算公式最终得到最佳样本量为 471。

（四）问卷设计

1. 问卷结构

问卷面向福建省专科、本科、研究生，分为 5 个部分，共 24 题。涵盖基本信息、对 AI 训练师的了解情况、从业意愿、培训教育了解程度和开放性问题。积极收集问卷并分析数据，了解被调查者对 AI 训练师的认知情况以及对该领域职业发展的态度。

2. 数据处理

使用 SPSS 进行描述性统计、二元 Logistic 回归、因子分析等工具进行数据处理。将选择题编号（如 Q1-Q24），转化为

Excel 文件。剔除例如未通过测试题、逻辑矛盾、填写不完整的一些无效问卷。对多选题全选 / 全不选、作答时间异常的问卷重点核查。

3. 问卷信度和效度检验

（1）信度检验

信度分析主要用来考察问卷中量表所测结果的稳定性以及一致性，即用于检验问卷中量表样本是否可靠可信^[9]。本调查对影响因素题（Q18-Q24）进行 Cronbach's α 检验，系数为 0.618（>0.6），证明问卷结构合理。

（2）效度检验

效度，即有效性，它是指测量工具或手段能够准确测出所需测量的事物的吻合程度^[9]。本调查将问卷进行效度分析，进行 KMO 和 Bartlett 的检验。KMO 值为 0.684（>0.6），研究数据适合提取信息，效度很好。进而可以进行问卷的信度检验。Bartlett 球形度检验中近似卡方 290.706，自由度 21，显著性 <0.001，变量间存在相关性。

二、AI 训练师供给现状

（一）潜在从业者分析

1. 描述性统计分析

正式调查共得到问卷 576 份，有效问卷 498 份，有效回收率 86.26%。

（1）性别分布

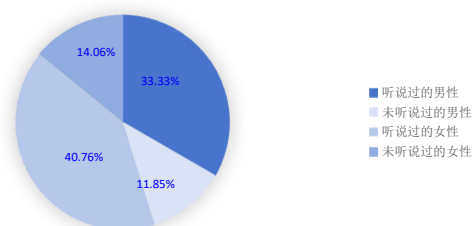


图1 受访者性别分布图

受访者性别比例接近 1:1，性别对职业认知无显著影响。

（2）学历情况

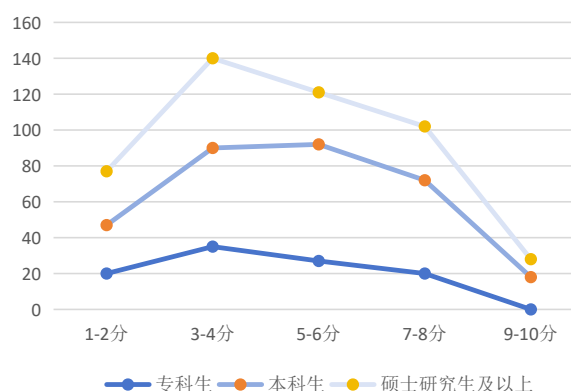


图2 受访者学历分布图

专科生、本科生、硕博生占比分别为 20.48%、49.60%、

29.92%。

专科学历平均意愿得分4.50分，本科4.90分，硕博5.33分，从业意愿随学历提升显著增强。

(3) 专业分布

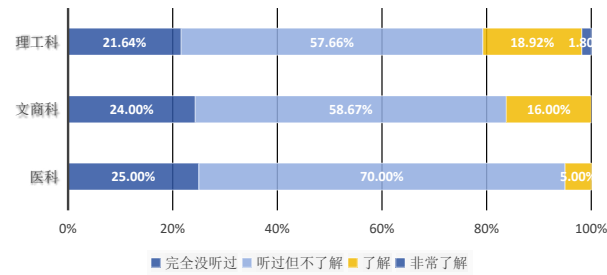


图3 受访者专业分布图

工科学生对AI训练师的了解程度显著高于文商科和医科，这种差异可能来自于理工科课程与AI技能的高度相关性。而医科学生因专业聚焦，相关认知度最低。

(4) 地区分布



图4 问卷回收数量地区分布图

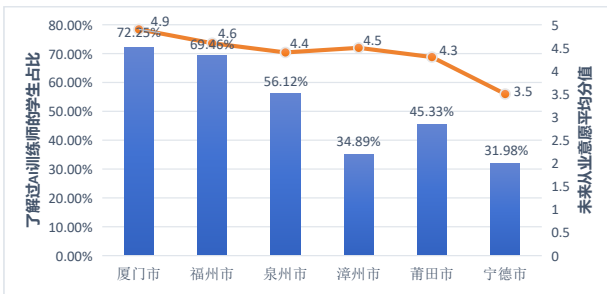


图5 受访者地区分布图

样本集中分布于AI产业较集聚的厦门、福州和泉州。产业发达地区学生对AI训练师行业认知更深、从业意愿更强。主要原因包括：

- ①产业集聚，教育资源与产业需求紧密结合
- ②市场需求旺盛
- ③政策支持，薪酬待遇较高

2. 情感分析

我们对问卷收集到的语句集进行共现频次计算，使用 Gephi 软件构建共词网络^[7]如下：

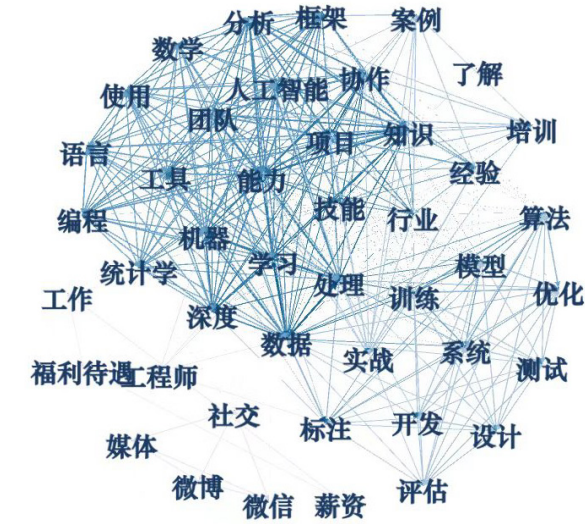


图6 情感共词网络图

当代大学生认为AI训练师需要较强的个人能力和团队协作意识，同时要精通数字化模型和算法。因此，对潜在从业者进行相关培训至关重要。

3. 二元 Logistic 回归分析^[8]

表1 关键驱动因素表

变量	回归系数	z 值	p 值	OR 值	95%CI
薪资水平符合预期 (x_1)	0.485	2.875	0.004	1.624	[1.167,2.261]
身边从业普遍性 (x_2)	0.622	4.279	0.000	1.863	[1.401,2.477]
福利待遇好 (x_3)	0.441	2.039	0.041	1.554	[1.017,2.373]
截距	-4.839	-4.838	0.000	0.008	[0.001,0.056]

二元 Logistics 回归分析结果显示，身边从业普遍性影响效应最强。模型通过显著性检验 ($\chi^2=55.230,p<0.001$)，无共线性问题 ($VIF<1.25$)。

4. 因子分析^[9]

表2 可靠性统计表

克隆巴赫 Alpha	基于标准化项的克隆巴赫 Alpha	项数
0.618	0.614	7

表3 KMO 和巴特利特检验表

KMO 和巴特利特检验		
KMO 取样适切性量数		0.684
巴特利特球形度检验	近似卡方	290.706
	自由度	21
	显著性	<0.001

Cronbach's $\alpha=0.618>0.6$ ，KMO=0.684>0.6，信效度达标，Bartlett 检验 $p<0.001$ ，满足因子分析前提。

表 4 总方差解释

因子	特征根			旋转前方差解释率			旋转后方差解释率		
	特征根	方差解释率 %	累加 %	特征根	方差解释率 %	累加 %	特征根	方差解释率 %	累加 %
1	2.344	33.486	33.486	2.344	33.486	33.486	2.295	32.790	32.790
2	1.145	16.359	49.845	1.145	16.359	49.845	1.194	17.055	49.845
3	0.953	13.619	63.464						
4	0.876	12.511	75.975						
5	0.701	10.021	85.996						
6	0.599	8.556	94.551						
7	0.381	5.449	100.000						

因子累计方差解释率为49.845%。

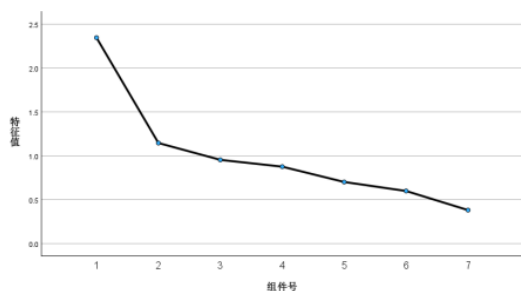


图 7 碎石图

碎石图结果显示特征根拐点位于因子2。

表 5 成分矩阵表

	成分	
	1	2
专业对口		0.727
兴趣爱好		0.663
高薪资	0.684	
福利待遇好	0.699	
工作时长短	0.679	
工作环境好	0.727	
职业发展前景好	0.580	

旋转后发现,因子1与工作环境好、福利待遇好和高薪资的系数较大,因子2与专业对口的系数最大。由此,得出大学生对选择AI训练师工作的意愿情况主要取决于工作环境、福利待遇、薪资待遇和专业对口。

（二）已从业者分析

1.AI 训练师职业关联要素



图 8 AI 训练师职业关联要素词云图

“本科”与“大专”表明这两个学历层次占据主导地位。“能力”和“技术”说明个人能力的重要性。“算法”“模型”“数据”是工作的核心内容。“公司”“团队”体现了所处的组织环境。“客户”“行业”“领域”则反映了工作的外部关联。

2.AI 训练师工作要求与福利



图 9 AI 训练师工作要求与福利词云图

“模型”“技术”强调专业技能的核心地位。“设计”“项目”“优化”展示承担的任务。“补贴”“绩效奖金”体现所能获得的回报。“公司”“团队”表明工作的组织场景。

三、AI 训练师需求现状

当下, AI 训练师在福建省的市场需求增长显著, 如厦门, 2024 年其被列为急需紧缺工种, 需求激增 415%^[10]。

（一）AI 训练师从业者人群分析

1. 性别分布

图 10 AI 训练师从业者性别分布图

AI 训练师从业者中男性占比较多, 约为 65%, 可能由于与 AI

训练师紧密相关的专业，男性的选择比例通常高于女性。

2. 年龄分布

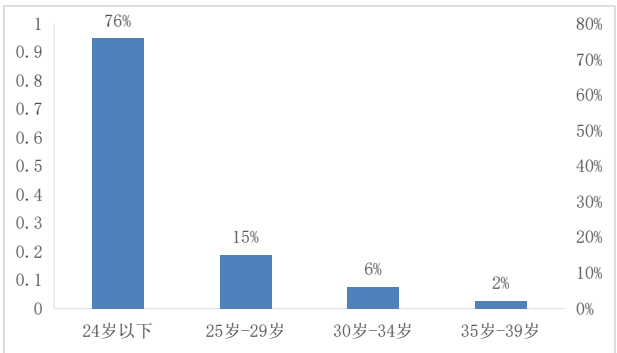


图 11 AI 训练师年龄分布图

AI 训练师从业者呈现出年轻态的年龄分布特征，其中 24 岁以下占比高达 76%。

3. 学历分布

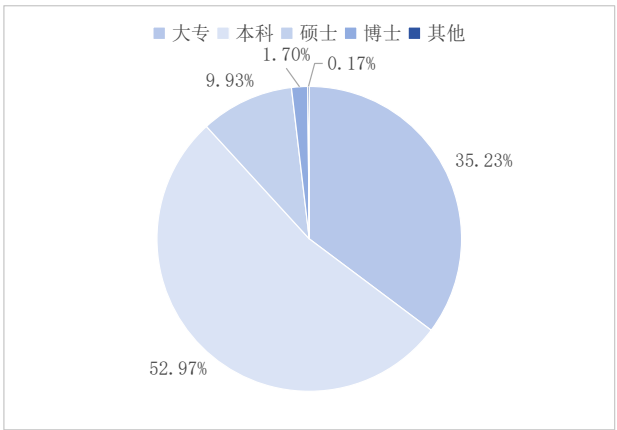


图 12 AI 训练师学历分布图

具有大专及以上学历的从业者达到 99% 以上，体现了 AI 训练师从业者较高的专业要求。

4. 公司标签

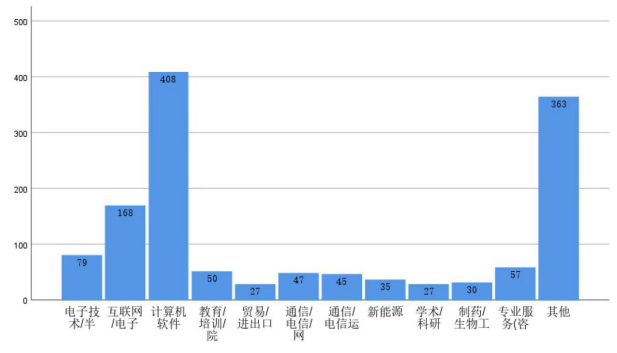


图 13 公司标签条形图

众多行业对 AI 训练师的需求均较高，其中计算机软件类公司对其需求在行业中占比最大。

5. 工资水平

AI 训练师工资集中在 6000-8000 元这一区间，说明行业成本控制使得薪资未达更高水准，但这一薪资范围对初入行者有一定吸引力，利于人才流入。

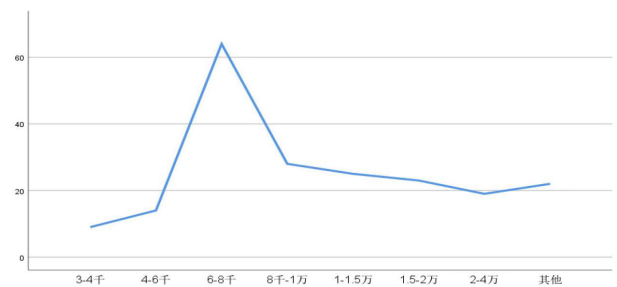


图 14 AI 训练师工资水平折线图

(二) AI 训练师职业相关要素

1. 工作经验

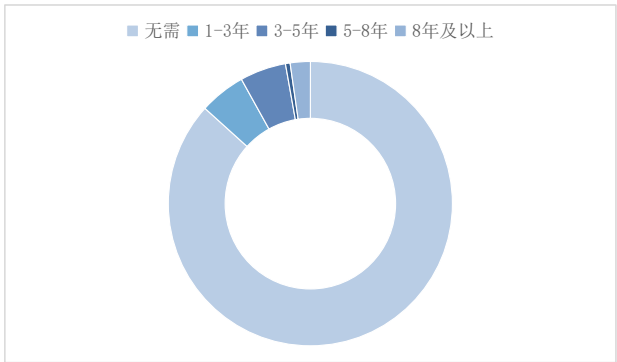


图 15 工作经验分布图

AI 训练师岗位对工作经验要求较低，“无需”的占比极高。说明该行业利于快速吸纳新鲜血液，推动发展。

2. 工作福利

AI 训练师工作福利整体较为丰富。取得相关证书可获 1000-2000 元补贴，同时不少企业提供带薪年假、健康体检、绩效奖金等福利。

3. 工作要求

AI 训练师工作要求兼具多方面能力。专业知识上，需掌握计算机科学、人工智能等；技术能力方面，需熟练运用 Python 等编程语言；实践中，需精通数据标注。此外，还应具备终身学习和自我提升能力。

三、结论与建议

(一) 结论

1. 学生对 AI 训练师职业的认知与从业意愿受多因素影响

(1) 理工科学生由于专业背景与 AI 训练师工作内容的高相关性，对这一职业的了解程度普遍高于文商科和医科学生。

(2) 学历越高，学生对 AI 训练师的从业意愿越强。

2. 福建省 AI 训练师市场供需不平衡，教育体系与产业需求脱节

随着 AI 技术的快速发展，企业对 AI 训练师的需求迅速增长^[11]。然而，高校教育体系在 AI 训练师相关知识和技能的传授方面尚未完全适配市场需求，导致人才供给不足。

3. AI 训练师职业发展路径不明确，影响人才吸引力

调研中发现，AI 训练师作为一个新兴职业，其职业发展路径

尚不明确。这导致从业人员在职业规划上感到迷茫，不清楚自己的发展方向和晋升路径。这种不确定性可能会影响对潜在从业者的吸引力，尤其对那些对职业稳定性和发展空间有较高要求的学生群体。

4. 企业对 AI 训练师的技能要求与人才储备之间存在差距

从企业需求角度看，福建省内企业对 AI 训练师的技能要求较高，涵盖专业知识、技术能力、实践操作能力以及综合素质等多方面。然而，当前的人才储备难以完全满足企业需求。

5. AI 训练师职业的薪资与福利水平有待提升

AI 训练师的薪资集中在 6000–8000 元，虽然对初入行者有一定吸引力，但与 AI 行业的高技术门槛和职业发展潜力相比，薪资水平仍有提升空间。

(二) 建议

1. 供给方面

(1) 优化教育体系，加强课程设置与市场需求的适配性

高校应根据 AI 训练师的市场需求，调整课程设置，增加数据标注、模型开发、算法优化等实践课程和技能培训模块。同时，也需要鼓励学生跨学科选修课程，培养具备数学、计算机科学、统计学等多学科背景的复合型人才，以满足 AI 训练师岗位对跨学科知识的要求。高校也应加强与企业的合作，建立实习基地，为学生提供实践机会，同时为企业输送符合需求的人才。

(2) 加强政策支持与引导，推动人才与产业发展

政府应出台相关政策，吸引高端 AI 人才流入，缓解人才短缺

问题。同时，加大对 AI 产业的扶持力度，推动企业与高校、科研机构的合作，促进 AI 技术在更多行业的应用，为 AI 训练师创造更多的就业机会和发展空间。

2. 需求方面

(1) 完善职业培训与认证体系，提升职业规范化水平

学校和机构应当开发针对不同学历层次（专科、本科、研究生）的 AI 训练师培训课程，注重实践操作和案例分析，帮助学生和从业者快速掌握所需技能。

(2) 优化薪资福利体系，提升职业吸引力与竞争力

企业可以构建差异化薪酬结构，实施例如“基础薪资 + 绩效奖金 + 技能津贴”的复合激励模式，缓解 AI 训练师薪资水平与行业潜力不匹配的问题^[12]。应重点向数据标注、复杂模型开发等核心技能岗位倾斜薪资资源，激发从业者积极性。同步完善福利体系，增设 AI 技能认证补贴、弹性工作制及职业发展基金。

(3) 激活社会网络效应，构建职业示范与社群支持体系

基于“身边从业普遍性”对职业意愿的影响，建议政府可以联合企业打造 AI 训练师职业社群网络，鼓励资深从业者与高校学生结对，定期开展职业分享会与项目实践指导。

(4) 强化企业责任，推动数据隐私保护与职业发展

企业应加强数据隐私保护措施，确保数据使用的合法性和安全性。在数据收集、存储和使用过程中，严格遵守相关法律法规，保护用户隐私。

参考文献

- [1] 聂文琪. 人工智能技术冲击下的产业变革、就业影响及应对策略 [J]. 湖北社会科学, 2024, (08): 90–103. DOI: 10.13660/j.cnki.42-1112/c.016374.
- [2] 黄晖. 智能制造甘愿: 人工智能训练的劳动组织形式与控制策略 [J]. 社会, 2025, 45(02): 1–31. DOI: 10.15992/j.cnki.31-1123/c.2025.02.001.
- [3] 陈琳, 高悦蓬, 余林徽. 人工智能如何改变企业对劳动力的需求? ——来自招聘平台大数据的分析 [J]. 社会科学文摘, 2024, (09).
- [4] 贺德方, 陈涛, 杨芳娟, 等. 营造人工智能发展条件筑牢智慧时代发展根基 [J/OL]. 智库理论与实践, 1–7[2025–06–19]. <http://kns.cnki.net/kcms/detail/10.1413.N.20250519.0936.002.html>.
- [5] 安瑜, 范顺祥, 乔玉辉, 等. 基于管理者偏好的京郊农场生态景观建设与优化对策 [J]. 中国农业大学学报, 2024, 29(04): 288–300.
- [6] 陈曦, 胡中锋. 基于 DeepSeek 的智能评分: 效度、信度与可行性研究 [J]. 高教探索, 2025, (03): 62–67.
- [7] 叶佳鑫, 朱乐, 雷凌云, 等. 词汇功能赋能的知识关联挖掘方法 [J/OL]. 情报杂志, 1–10[2025–06–19]. <http://kns.cnki.net/kcms/detail/61.1167.G3.20250526.1401.010.html>.
- [8] 杜建刚, 吴章建, 朱丽雅, 等. 拥抱 AI 生成: 消费动机对人工智能生成内容偏好的影响 [J]. 管理科学, 2025, 38(01): 130–144.
- [9] 李洪晨, 赵星. 人工智能准备度、STARA 意识对人工智能增强科研创新的影响——基于信息资源管理学科的调查 [J/OL]. 图书馆论坛, 1–11[2025–06–19]. <http://kns.cnki.net/kcms/detail/44.1306.G2.20250604.0956.002.html>.
- [10] 林露虹, 何无痕. AI 赋能共启职业新赛道 [N]. 厦门日报, 2025–03–04(A04). DOI: 10.28890/n.cnki.nxmrb.2025.000570.
- [11] 姚树洁. 人工智能浪潮中的大国博弈: 中国的战略抉择与前景分析 [J]. 人民论坛·学术前沿, 2025, (09): 12–22. DOI: 10.16619/j.cnki.rmltxsqy.2025.09.002.
- [12] 周芷羽, 杨美菱. 应用型职业院校人才供需适配机制研究 [C]// 重庆市继续教育学会. 智能教学创新发展学术研讨会论文集 (数智教育专题). 四川天一学院, 2025: 182–186. DOI: 10.26914/c.cnkihy.2025.014644.

广州市临期食品销售现状及消费者购买意愿影响因素调查

颜欣如, 陈伟钰*, 吴语桐, 李庆驹
华南农业大学 经济管理学院, 广东 广州 510640
DOI:10.61369/ASDS.2025070002

摘 要 : “临期食品”指接近食品保质期, 仍在安全食用范围内的包装食品。近年来, 临期食品行业规模持续增长。然而, 当前临期食品行业仍处于发展初期, 行业存在企业食品安全乱象、消费者认知不足、营销策略缺乏针对性等困难。由此, 聚焦临期食品销售端现状、消费者临期食品认知情况和消费者临期食品购买动因, 结合网络数据挖掘和实地调查, 了解广州市988名消费者临期食品购买体验和购买动因, 并使用 CRITIC 权重法、结构方程和 K-means 聚类方法分析广州市消费者临期食品购买意愿影响因素, 进而从销售企业和政府部门层面提出针对性建议, 为推动临期食品行业高质量健康发展提供理论依据和实践参考。

关 键 词 : 临期食品; 市场现状; 反食品浪费; 购买意愿

A Survey on the Sales Status of Near-Expired Food and Influencing Factors of Consumers' Purchase Intention in Guangzhou

Yan Xinru, Chen Weiyu*, Wu Yutong, Li Qingju
College of Economics and Management, South China Agricultural University, Guangzhou, Guangdong 510640

Abstract : Near-expiry food refers to packaged food that is close to its expiration date but still within the safe consumption range. In recent years, the scale of the near-expiry food industry has been growing continuously. However, the current near-expiry food industry is still in the initial stage of development and faces challenges such as food safety chaos in enterprises, insufficient consumer awareness, and lack of targeted marketing strategies. Therefore, this study focuses on the current situation of the near-expiry food sales end, consumers' cognition of near-expiry food, and consumers' motivation for purchasing near-expiry food. Combining network data mining and field investigation, it investigates the purchase experience and motivation of 988 consumers in Guangzhou towards near-expiry food. It also uses binary logistics model, CRITIC weight method, structural equation and K-means clustering method to analyze the influencing factors of consumers' willingness to purchase near-expiry food in Guangzhou. Furthermore, targeted suggestions are put forward from the perspectives of sales enterprises and government departments, so as to provide theoretical basis and practical reference for promoting the high-quality and healthy development of the near-expiry food industry.

Keywords : near-expiry food; market status; anti-food waste; purchase intention

引言

“临期食品”指接近食品保质期, 仍在安全食用范围内的包装食品。在第十三届全国人民代表大会常务委员会《中华人民共和国反食品浪费法》和国家发展改革委《反食品浪费工作方案》等政策引导下, 我国临期食品行业迎来快速发展^[1-4]。

近年来, 临期食品所引发的食物浪费与食品安全问题, 日益受到学界关注, 相关研究亦取得一定成果。然而, 与临期食品市场经济、监管等研究领域相比, 从消费者主观视角剖析临期食品购买意愿的研究起步较晚, 尚存诸多不足。当前文献多聚焦于临期食品法律

项目信息: 2025年“正大杯”全国大学生调查与分析大赛国家级一等奖获奖作品; 2025年省级大学生创新训练计划项目(S202510564229)。

作者简介:

颜欣如, 华南农业大学经济管理学院本科在读, 研究方向为市场调查;

吴语桐, 华南农业大学经济管理学院本科在读, 研究方向为市场调查;

李庆驹, 华南农业大学经济管理学院助教, 研究方向为大学生创新创业。

通讯作者: 陈伟钰, 华南农业大学经济管理学院本科在读, 研究方向为市场调查。

法规的完善、市场监管策略以及理性对待临期食品的路径，侧重关注学术上临期食品的社会价值，鲜少文献从市场现实出发思考当前临期食品销售困境^[5-13]。由此，本研究聚焦临期食品销售现状及消费者临期食品购买意愿影响因素，综合运用文献研究法、实地调查法和问卷调查法的方法组合进行系统分析，旨在为“推动临期食品行业高质量健康发展”这一调研目的提供理论依据和实践参考。

一、舆情分析

为给方案设计提供更多参考，研究从消费者和网民两个视角进行初步分析。消费者方面，利用爬虫技术爬取淘宝网关于临期食品的评论信息，经过数据清洗、中文分词后对评论数据进行高频词词云图绘制，并进一步结合 BosonNLP 情感词典进行情感倾向分析。在网民舆论方面，通过抖音旗下巨量算数平台进行关键词数据获取，并对获取信息进行探索分析，根据分析结果确认广州市为研究调研地。

（一）基于网络爬虫的消费者关注点分析

1. 高频词词云图

从下图的结果可以明显看出，消费者对于临期食品的评价大多是以“好吃”“实惠”等与好评相关的词汇为主，这说明了临期食品得到了大部分消费者的认可。出现频率中等的“性价比”“物美价廉”等词汇则突出了临期食品本身具有的临期、价格优惠、性价比高的特点。同时，频率占比较小的词汇如“过期”“差评”等也从另一方面反映了商家需要改进的方面。



图1：爬虫词云统计图

2. 临期食品评价的情感倾向分析——基于情感词典

文本情感倾向性分析通过对文本进行分句分词，并结合 BosonNLP 情感词典对情感词进行评分统计。若总分值大于0，则情感倾向为积极；反之则为消极。利用 Python 代码，通过 BosonNLP 情感字典分析评论情感倾向结果如下：

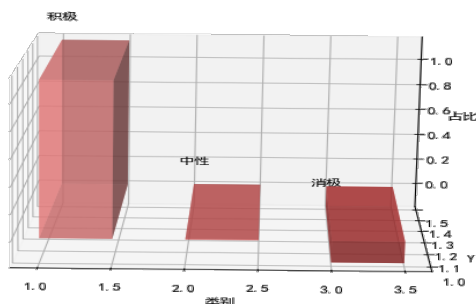


图2：网络评论情感占比

从图中可明显地看出消费者对于临期食品的态度大部分以积极为主，占比非常大，可见消费者对于临期食品的接受度较高，评价正面。而负面的评论占比较小，表现出当下临期食品销售端已趋向成熟，较小程度引起消费者不满。

（二）基于抖音指数的网络关注现状分析

为了解消费者对“临期食品”的关注程度及其市场发展现状，并为本次问卷设计和调研提供优化依据，本部分利用抖音旗下巨量算数平台获取相关网络数据。



图3：关键词搜索全国分布图

根据全国搜索指数分布的结果，显示山东、广东、江苏的搜索指数排名前三。反映了这些地区消费者对新兴消费模式的开放态度以及对高性价比商品的需求。

而基于巨量算数关键词分析，我们选取广州市作为我们的调研地：一是广东是临期食品搜索人群分布占比第二的省份，广州的搜索指数在广东省排名第一；二是广州作为广东省会城市，是改革开放的前沿城市，消费层次高且开放，同时也是国内较早发展临期食品专卖及连锁品牌的城市，研究其消费者购买意愿具有一定参考意义。

二、调查方案设计及实施

（一）调查方式

基于文献调查法，团队开展文献检索，梳理临期食品领域的学术论文、行业白皮书及政策文本，聚焦购买行为动因等核心维度，为问卷设计奠定基础。

基于问卷调查法，采用多阶段抽样法，选取广州市六区十二街道实施定点拦截，发放问卷以获取消费者购买行为、购买认知及购买意愿影响因素的一手数据，为后文分析提供数据支撑。

基于实地调查法，对选定街道上的商超展开实地调研，随机抽取货架上的食品进行日期统计，同时统计食品种类、摆放方式以及降价幅度等，以探究“是否存在临期食品混放”的销售端乱象。此外，记录商超对临期食品的处理方式，并将其与普通食品进行对比分析，以进一步明晰临期食品市场面临的问题，从而提

出相应的改进策略。

（二）问卷设计

1. 整体问卷设计

- （1）问卷结构：说明部分、甄别部分、主体部分、结束部分
- （2）问卷题型：单选题、多选题、开放题
- （3）甄别部分：本部分共设置两道题。其中一道题是注意力测试题，要求选择“33”，未选择“33”说明专注力不够，作为废卷处理。另一题是“近一年内是否购买过临期食品”来区分消费者和潜在消费者。

2. 量表设计

为充分了解广州市临期食品消费者现状，及临期食品购买意愿影响因素，本团队根据消费价值理论（TCV），基于已有关于临期食品购买意愿影响因素的文章，从价格、品质、消费观念、情境、社会责任认知、食品安全素养等六个方面设计问卷量表，问卷通过了信效度检验，题目有效。

3. 数据处理

在正式调查开始之前，我们利用微信等 app，向符合条件的人群发放了若干份预调查问卷，并回收有效问卷 206 份。

正式调查在广州市 6 区 12 街道通过街头拦访方式发放问卷 1093 份，回收有效问卷 988 份，有效回收率达 90.4%，满足最低样本量 722 份的统计要求。对问卷进行信效度检验，量表信度符合要求，调查结果可信。

表 1：正式调查信度检验表

类别	Cronbach's α 系数	标准化 Cronbach's α 系数	项数
整体量表	0.727	0.733	19
价格因素	0.659	0.663	2
品质因素	0.821	0.822	7
消费观念因素	0.758	0.759	4
情境因素	0.612	0.614	2
社会责任认知因素	0.771	0.773	2
食品安全素养因素	0.730	0.737	2

表 2：正式调查问卷 KMO 和 Bartlett 球形检验表

KMO 和 Bartlett 球形度检验			
KMO 值		0.817	
Bartlett 球形度检验	近似卡方	6912.067	
	df	171	
	p	0.000	

从以上结果分析，正式问卷的各个 Cronbach's α 系数值均大于 0.6，表示量表通过信度检验，正式调查问卷结果可信。检验结果 KMO 值 = 0.817，Bartlett 球形度检验 p 值小于 0.05，统计上显著。正式问卷具有结构效度，可以进行因子分析。

三、销售端现状实地调查分析

（一）广州市商超临期食品日期抽样调查研究

在商超食品日期抽样抽查中，我们重点关注了问卷中“您更倾向于购买以下哪些品类的临期食品”中的同类食品与在临期货架中经常出现的商品类型，同时运用 excel 计算各商品的上市时

间，并定义了以下公式：

$$\text{临期率} = \frac{\text{上市天数}}{\text{保质期}} \times 100\%$$

本次调查覆盖广州市 6 区 28 家商超，共抽取 801 件商品，涵盖休闲食品类、饮料类、乳制品类、速冻食品类、烘焙食品类、粮油副食类等，样本具有一定代表性。

计算各商超的抽查商品平均临期率，具体如图 4。虽然抽查数量较少，但也反映出部分信息，如：好特卖这样的食品折扣店临期率较高，部分规模较小的社区超市临期率偏高，而大型连锁超市和量贩零食店的平均临期率较低。这一差异与不同业态的库存周转能力密切相关，临期食品折扣店以临期商品为主要销售对象，而大型商超和量贩零食店依托高效供应链能更快消化库存。

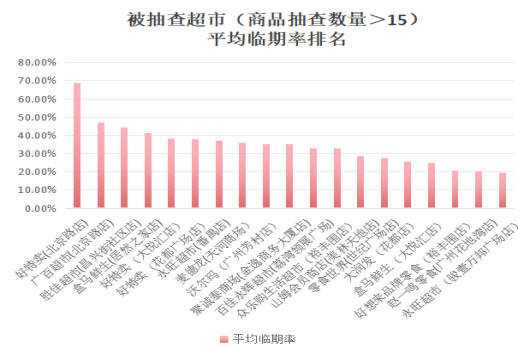


图 4：被抽查超市临期率分布

对于所有抽查商品，我们根据临期食品界定表，界定临期率 > 85% 为临期食品，由图 5 可知本次调研超市存在临期食品混杂于商品货架销售的情况，且比例并不算低，需要予以重视。

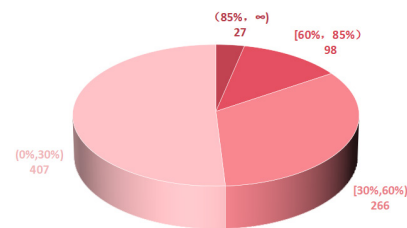


图 5：抽查商品临期率分布饼图

（二）广州市商超临期食品价格对比分析

经过问卷调查得出，价格是影响消费者临期食品购买意愿的关键因素，其定价可能会刺激潜在消费者的消费，相关数据如下图所示。

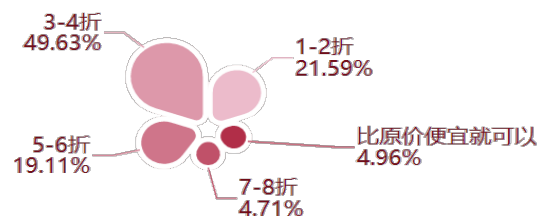


图 6：潜在消费者价格期望

在调研过程中，发现广州市临期食品销售主要存在专门临期食品货架、折扣货架、促销货架、捆绑货架、夹杂商品间销售、明显标签等类型。其中，前三类较为显眼，而后几类则较不明

显，临期食品夹杂在其他商品之间，需要辨别。

通过统计商超对于临期食品的定价，有助于进一步拓展临期食品发展路径。由图7结果显示，抽取调研商超的临期食品平均降价幅度约为50%，而依据图6，认为应降价60%以上的潜在消费者占到了71.22%，表现临期食品定价策略还有一定改进空间。

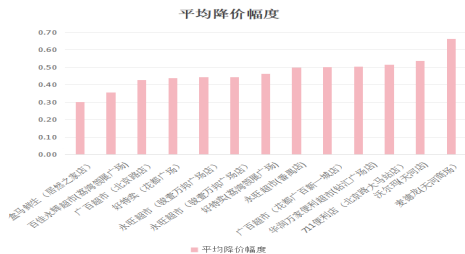


图7：各超市临期食品降价幅度对比图

四、统计建模与分析

（一）描述性统计

1. 消费者基本信息

本次受访消费者的基本信息如下图表所示。

表3：消费者基本信息

调查内容	百分比	调查内容	百分比
性别	男性 47.06%	学历	高中及以下 23.89%
	女性 52.94%		专科 30.16%
年龄	18岁以下 2.53%		本科 39.68%
	18-23岁 27.94%	年龄	研究生及以上 6.27%
	24-30岁 12.75%		2000元以下 31.18%
	31-40岁 14.68%		2001—5000元 25.30%
	41-50岁 14.27%		5001—8000元 20.45%
	51-60岁 13.46%		8001—12000元 17.91%
	60岁以上 14.37%		12000元以上 5.16%

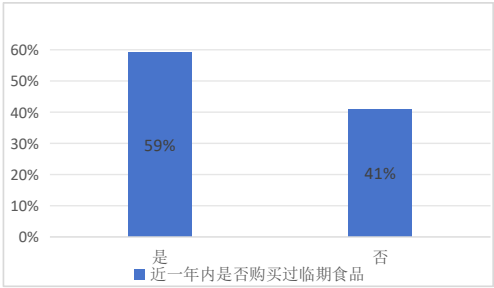


图8：消费者临期食品购买情况

从以上图表数据可知，本次调查覆盖不同年龄阶段、学历层次和社会经济地位的人群，具有较强的代表性和统计意义。样本性别分布均匀，男女消费热情相当。月收入主要集中在12000元以内，可体现社会中占比较大的中低收入人群的购买习惯。年龄段以18岁以上为主，他们消费决策更具理性。学历集中在研究生以下，不同知识层次人群都有消费。有59%的受访对象近一年内购买过临期食品。

2. 消费者购买途径及购买黏性分析

（1）信息获取渠道与购买途径

如图9所示，在线下端，消费者主要以店内信息标语、店员指

引为信息获取渠道，而在线上端，则依靠电商平台推荐和推送。消费者主要在超市临期食品折扣货架、临期食品折扣店，便利店中购买临期食品。

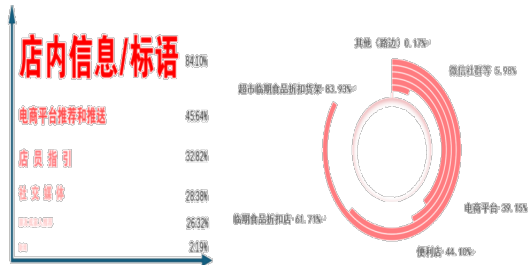


图9：临期食品信息渠道分布图

（2）购买黏性分析

分析图10购买频率和金额发现，两者走势基本一致，呈现先升后降的趋势。这表明消费者依赖度与消费能力呈负相关，且低频低金额群体明显高于高频高金额消费群体。

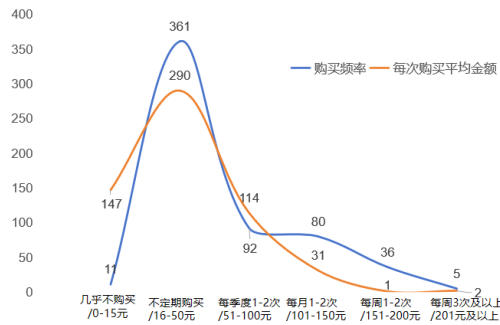


图10：临期食品购买黏性双线图

3. 消费者消费倾向

各类临期食品中，消费者更倾向于购买休闲食品和饮料，其次是烘焙食品、乳制品和速冻食品，其中最少的粮油副食类。

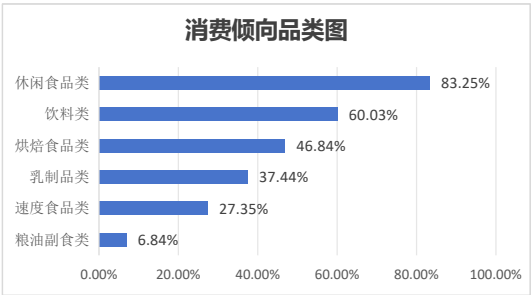


图11：临期食品消费倾向品类图

（二）基于 CRITIC 权重法的购买意愿影响因素权重分析

CRITIC 权重法是一种基于标准差和指标冲突性的客观权重赋权方法，通过衡量指标取值的分散程度和指标间的冲突性来确定各指标的权重。

根据 CRITIC 方法，计算得每个指标的分析结果如下表所示：

表5：CRITIC 权重法分析结果

指标	指标变异性	指标冲突性	信息量	权重 (%)
价格	0.72	5.518	3.971	11.693
质量	1.053	5.819	6.127	18.043
环境	0.893	5.323	4.752	13.993
面子	0.839	5.722	4.799	14.133

环保意识	0.764	5.66	4.323	12.731
服务	0.919	5.217	4.793	14.113
尝鲜心态	0.92	5.646	5.193	15.294

模型结果可得，影响消费者购买临期食品购买意愿指标重要性由高到低排序为：质量、尝鲜心态、面子、服务、环境、环保意识、价格。

与舆情分析结论不同的是，价格在本模型排名最低。对此进行解释，我们发现，价格的平均重要性最高，但因为分数比较集中导致标准差较小，故在本模型中排序较低，但我们仍然不可以忽视它的重要性。

（三）基于结构方程模型的购买意愿影响因素路径分析

基于文献研究与问卷调查结果，本研究发现价格敏感度、质量情绪、消费情景及判断能力是影响消费者购买临期食品的主要因素。为深入探究这四个因素对消费者购买意愿的具体影响程度及其作用机制，本研究采用结构方程模型进行分析。将上述四个维度与“临期食品购买意愿”设为潜变量，根据对消费者购买临期食品意愿影响因素的问卷结果，以及对相关消费者的深度访谈，本研究提出了临期食品购买意愿、价格敏感度、质量情绪、消费情景、判断能力五个范畴的研究假设关系。

表6：结构方程假设汇总表

假设汇总
H1: 价格敏感度对临期食品购买意愿有显著的正向关系。
H2: 质量情绪对临期食品购买意愿有显著的负向关系。
H3: 消费情景对临期食品购买意愿有显著的正向关系。
H4: 判断能力对临期食品购买意愿有显著的负向关系。
H5: 价格敏感度对质量情绪有显著的正向关系。
H6: 消费情景对质量情绪有显著的正向关系。
H7: 价格敏感度对消费情景有显著的正向关系。
H8: 消费情景对判断能力有显著的正向关系。
H9: 质量情绪对判断能力有显著的正向关系。

构建结构方程模型，采用极大似然法进行估计，并根据模型结果进行合理修正后，结果如下：

表7：结构方程估计结果表

假设路径	路径系数	标准化路径系数	p
价格敏感度→购买意愿	0.776	0.699	0.000***
质量情绪→购买意愿	-0.047	-0.051	0.161
消费情景→购买意愿	0.146	0.113	0.010**
判断能力→购买意愿	-0.151	-0.088	0.012**
价格敏感度→质量情绪	1.347	1.113	0.000***
消费情景→质量情绪	0.102	0.072	0.003***
价格敏感度→消费情景	0.258	0.301	0.000***
消费情景→判断能力	0.055	0.103	0.010**
质量情绪→判断能力	0.259	0.344	0.000***

据结构方程分析结果可知：

1. 价格敏感度和消费情景正向影响临期食品购买意愿，而质量情绪和判断能力负向影响购买意愿。消费者对价格变动越敏感，越可能因价格优势而选择临期食品；消费情景（如购物环境和人员服务）的优化也能推动购买行为。而对商品质量的担忧和人们对安全知识的掌握会阻碍人们购买临期食品。

2. 价格敏感度和消费情景都正向影响质量情绪。消费者对价

格越敏感，越倾向于关注价格而非质量，可能因为低价而降低对质量问题的担忧；而舒适的消费环境和专业的人员服务，也会缓解消费者对临期食品质量的疑虑，进一步提升其质量情绪。

3. 消费情景和质量情绪都正向影响判断能力。消费情景中强调环保或可持续发展，消费者可能会逐渐认同这些价值观，并将其内化为自己的判断能力；而对食品质量越担忧，消费者越可能提升自己食品安全方面的相关能力。

（四）基于K-means聚类的两类消费者人物画像分析

采用K-means聚类对问卷受访者进行聚类分析，根据分析结果将消费者划分为试探型消费者、理智尝鲜型消费者和忠诚型消费者三类。结果如表8所示。

表8：消费者聚类类别表

类别	因子	类别1	类别2	类别3	F	p
基本信息	性别	1.396 ± 0.49	1.479 ± 0.501	1.6 ± 0.492	7.501	0.001***
	年龄	3.227 ± 2.112	3.889 ± 2.327	3.757 ± 2.171	5.652	0.004***
	职业	2.193 ± 1.454	2.639 ± 1.804	3.261 ± 2.048	17.838	0.000***
	月收入	3.258 ± 2.104	4.215 ± 2.263	3.896 ± 2.154	11.056	0.000***
	频率	1.748 ± 1.232	2.083 ± 1.361	2.226 ± 1.332	7.33	0.001***
影响因素	质量问题	3.387 ± 1.013	3.847 ± 0.839	3.635 ± 0.949	12.072	0.000***
	推荐意愿	3.632 ± 0.887	3.819 ± 0.816	3.748 ± 0.836	2.586	0.076*
	临时需求	3.095 ± 0.999	3.076 ± 0.997	2.696 ± 1.053	7.05	0.001***
	尝鲜心态	3.638 ± 0.886	3.396 ± 0.887	3.278 ± 0.996	8.106	0.000***
	购买意愿	4.0 ± 0.647	4.056 ± 0.843	4.226 ± 0.593	4.555	0.011**
	满意度	7.242 ± 1.427	7.403 ± 1.543	7.809 ± 1.123	6.927	0.001***
评价	喜爱程度	6.936 ± 1.673	7.312 ± 1.632	7.6 ± 1.262	8.308	0.000***

1. 消费者主要特征

（1）试探型消费者

该类消费者购买频率最低，主要基于临时需求或尝新动机进行购买，对临期食品的质量关注度较低。其购买意愿和推荐意愿均相对较低，对临期食品的满意度和喜爱程度也偏低，但在消费者群体中占比最大。

（2）理性尝鲜型消费者

该类消费者购买频率中等，较为关注产品质量，可能因临时需求而选择临期食品，同时存在一定尝鲜心态。其购买意愿、推荐意愿以及对临期食品的满意度和喜爱程度均处于中等水平。

（3）忠诚型消费者

该类消费者购买频率最高，已形成稳定的购买习惯，对临期食品的质量关注度适中。其对临期食品的整体满意度和喜爱程度较高，不仅愿意持续购买，还乐于向他人推荐，但在消费者群体中占比最小。

五、结论与建议

（一）结论

1. 临期食品销售端现状分析

渠道上，大型连锁商超临期率低，小型社区超市、折扣店临期率高且临期食品常混杂普通货架，部分商超未设独立临期专区、区位隐蔽、标识模糊，影响消费者信任与选购；定价上，平均降价约50%，但71.22%消费者期望降价超60%，且标签未与原价对比，吸引力不足；沟通上，货架多仅标“临期”，未能消除消费者对质量顾虑。

2. 消费者认知情况

消费者对临期食品存在认知参差，多数认知不足且环保理念薄弱；食品安全为主要顾虑，价格敏感显著，性价比是核心驱动力；偏好休闲食品、饮料等，其临期食品依赖度与消费能力负相关；行业已初步建立信任，消费接受意愿呈上升趋势。

3. 消费者购买动因

行为层面，年龄偏大、学历较高的中低收入群体更倾向购买，凸显理性决策特征；价格折扣是核心驱动力，对购买行为呈显著正向影响。品质顾虑构成主要阻碍，质量相关因素强烈抑制购买意愿。心理层面，理性品质化消费观念与食品安全素养促进购买，环保理念影响尚浅但关注度提升，而“面子”顾虑负向作用显著。此外，消费场景优化（如环境、服务）既直接推动购买，还能缓解质量担忧，通过双重路径提升购买意愿。

（二）建议

1. 销售企业

销售企业可通过规范陈列与智能引导解决线下商超渠道分化问题：在商超出入口设置统一标识的“临期食品专区”，严禁混放，并通过广播、电子屏及购物小票附加优惠券等方式增加临期食品曝光度；同时引入智能库存系统，自动预计临期食品并安排分配至专区，降低过期风险，减少消费者购买顾虑。

针对定价与消费者预期不符问题，销售企业可实施阶梯式定价机制：按保质期剩余时间采取5-7折“品质特惠”和2-5折“清仓特惠”策略，进一步满足消费者对60%以上降价幅度的期望。同时在标签上标注原价、现价及折扣比例，如“原价20元→现价8元，降价60%”，还可以参考好特卖的标签设计补充保质期等细节，从而激发消费者购买兴趣。

2. 政府部门

针对食品安全与监管问题，政府可通过法律法规明确临期食品分类标准，加强监管，确保商家规范销售；还可以定期对商超的食品进行抽检，严惩篡改保质期等违法行为，并公示违规案例，以儆效尤。

对于消费者认知不足，环保意识薄弱问题，一方面通过多媒体宣传，在电视机、短视频等平台科普临期食品相关知识，倡导“理性消费、减少浪费”的理念；另一方面，加强社区和校园教育，通过开展专家答疑、现场剧情表演等活动增强公众认知；并在中小学课程中融入食品节约内容，培养青少年环保意识。

参考文献

- [1] 成升魁, 高利伟, 徐增让, 等. 对中国餐饮食物浪费及其资源环境效应的思考 [J]. 中国软科学, 2012, (07): 106-114.
- [2] 中国连锁经营协会. 市场端临期食品经营现状报告 [R/OL]. <http://www.ccfa.org.cn/portal/cn/xiangxi.jsp?id=444173&type=33>, 2023-01-06.
- [3] 李如斯, 刘惠明. 社会共治背景下临期食品安全问题法律规制探究 [J]. 保鲜与加工, 2023, 23(01): 70-74.
- [4] 艾媒咨询. 2023-2024 年中国临期食品行业发展及标杆案例研究报告 [R/OL]. <https://www.iimedia.cn/c400/93052.html>, 2023-05-06.
- [5] 贾润梅. 临期食品：“捡漏”也要安全 [N]. 中国质量报, 2022-04-28(007).
- [6] 侯伟胜. 商家赢利新契机——临期食品 [J]. 商业观察, 2024, 10(12): 18-20.
- [7] 王森. 国内临期食品品牌的整合营销策略研究 [J]. 现代商业, 2023, (21): 35-38.
- [8] 孙效敏, 高桂林. 刍议“临期食品”监管 [J]. 食品工业科技, 2015, 36(01): 32-33.
- [9] 朱有刚. 山东省市场监管局发布消费提示购买临期食品理性消费勿贪多 [J]. 中国食品工业, 2021, (20): 47.
- [10] 陈冠君. 消费者对临期食品购买意愿的影响因素研究 [D]. 吉林大学, 2020.
- [11] 刘超越. 基于质量属性的消费者临期食品购买意愿干预研究 [D]. 吉林大学, 2021.
- [12] 张蓓, 招楚尧, 赖恒坚, 等. 综合质量、消费情境与临期食品购买意愿——价格敏感度的中介与食品安全素养的调节 [J]. 贵州财经大学学报, 2022, (01): 36-45.
- [13] 曹世阳. 感知风险及面子视角下消费者对次优食品的购买意愿研究 [D]. 吉林大学, 2022.

广东佛山木版年画认知现状调查与活态传承意愿分析

黄冬燕¹, 张兆芹^{2*}, 李爱萍², 冼卓毅¹

1. 广东工业大学管理学院, 广东 广州 510000

2. 广东工业大学经济学院, 广东 广州 510000

DOI:10.61369/ASDS.2025070003

摘 要 : 在全球化和现代化浪潮冲击下, 佛山木版年画作为岭南非物质文化遗产的代表, 面临代际断层与社会认知疏离的严峻挑战。本文以佛山市常住居民为调查对象, 通过分层抽样与问卷调查, 构建“认知-意愿-群体”三维分析框架, 系统探究其认知现状与传承困境。列联表分析表明, 居民对木版年画的认知存在显著代际与职业分化, 凸显文化认同断层。结构方程模型验证了文化认知、文化认同及民族自豪感对传承意愿的多层次驱动作用。基于 K-means 聚类的受众细分将群体精准划分为传承探索型、文化疏离型与浅知待塑型三类, 为差异化策略提供靶向依据。本研究为非物质文化遗产的活态传承提供了兼顾文化内核坚守与现代适应的实证路径, 对传统工艺可持续发展具有实践指导价值。

关 键 词 : 佛山木版年画; 认知差异; 传承意愿; 结构方程模型; K-means 聚类

A Survey on the Current Cognitive Status and Willingness for Living Inheritance of Guangdong Foshan Woodblock New Year Paintings

Huang Dongyan¹, Zhang Zhaoqin^{2*}, Li Aiping², Xian Zhuoyi¹

1.School of Management, Guangdong University of Technology, Guangzhou, Guangdong 510000

2.School of Economics, Guangdong University of Technology, Guangzhou, Guangdong 510000

Abstract : Under the impact of globalization and modernization, Foshan woodblock New Year paintings, a representative Lingnan intangible cultural heritage, face severe challenges including generational discontinuity and social detachment. This study targets Foshan's permanent residents, employing stratified sampling and questionnaires to construct a tripartite framework ("cognition-willingness-group") to systematically explore public awareness and transmission challenges. Contingency table analysis reveals significant disparities in awareness across generations and occupations, highlighting fractured cultural identity. Structural equation modeling validates the multi-layered driving effects of cultural cognition, identity, and national pride on transmission willingness. K-means clustering segments the audience into three distinct groups: transmission pioneers, culturally detached individuals, and the uninformed awaiting cultivation, enabling targeted strategies. The research offers an empirical pathway for the living inheritance of intangible cultural heritage—balancing core cultural preservation with modern adaptation—providing practical guidance for sustainable traditional craftsmanship development.

Keywords : Foshan Woodblock New Year Paintings; cognitive differences; inheritance willingness; structural equation model; K-means clustering

引言

在全球化与现代化浪潮的双重冲击下, 非物质文化遗产的传承与创新已成为当代文化研究的核心议题。作为岭南传统文化的璀璨明珠, 佛山木版年画起源宋代、鼎盛明清, 融合岭南民俗风格与精湛技艺, 承载传统神祇符号与祈福功能, 深植在岭南社会的集体记忆中。2006年, 佛山木版年画入选第一批国家级非物质文化遗产名录, 标志着其作为文化遗产的官方认可。然而, 随着社会变迁与消费文化转型, 这一古老技艺正面临前所未有的传承危机: 过去三十年传承人数锐减, 青年群体认知率较低。这种代际断层与社会认知的疏

基金项目: 广东省哲学社会科学规划学科共建项目“信用风险缓释工具助力广东民营企业融资的效果与对策研究”(GD23XYJ22); 广东省哲学社会科学规划学科共建项目“粤港澳大湾区先进制造业和现代服务业深度融合的双向赋能路径研究”(GD23XYJ21)。

作者简介:

黄冬燕, 广东工业大学管理学院, 本科在读, 研究方向为企业数字化转型;

李爱萍, 广东工业大学经济学院, 本科在读, 研究方向为企业数字化转型;

冼卓毅, 广东工业大学管理学院, 本科在读, 研究方向为企业金融。

通讯作者: 张兆芹, 广东工业大学经济学院, 讲师, 硕士生导师, 研究方向为公司金融与风险管理。

离，折射出传统手工艺在现代性语境下的生存困境。

在此背景下，研究佛山木版年画传承的认知差异及公众传承意愿的影响因素，对破解其传承困境、构建活态传承体系具有重要的现实意义。

一、文献综述

（一）活态传承的理论界定

活态传承的理念源自于对非物质文化遗产的保护，2018年联合国教科文组织总部文化部门非物质文化遗产科更名为活态遗产处，凸显了其动态性和活态性质^[1]。活态传承是对非物质文化遗产生成、发展的环境予以科学的保护和传承，是在人民群众生产生活中进行传承的一种方式^[2]。活态传承强调在文化原生环境中通过主体参与实现代际传递，既要求保护非遗的文化内核，又需借助生产性传承、教育传播和社区参与等方式，使其在现代生活场景中保持生命力并实现创新性发展^[3]。

（二）佛山木版年画的文化内涵与技艺特征研究

在佛山木版年画的文化内涵与技艺特征研究中，学者们聚焦其符号象征与艺术表达。徐秋枫（2010）指出佛山木版年画的艺术魅力源于红、绿、黄等对比色的大胆运用所营造的喜庆氛围，凸显岭南特色^[4]。程宜（2012）进而探讨了佛山民众的信仰体系与木版年画的互动关系，如门神画的“守护家园”寓意、神像画的“祈福纳祥”内涵，揭示了民俗信仰与年画艺术相互依存、相互促进的紧密联系^[5]。许洁玲（2016）指出佛山木版年画实质是广府民俗文化的视觉符号系统，通过象征与隐喻手法承载民众对美好生活、家族延续及神灵庇佑的集体祈愿^[6]。李晗（2019）深化此视角，指出佛山木版年画中的门神画与神像画以丹红、金银色等对比色及刚柔线条为媒介，将道教祈福、辟邪诉求转化为视觉符号，而道教神谱与广府水神崇拜的融合，反映了岭南地域文化对道教体系的本地化重构^[7]。

（三）佛山木版年画的传承与应用研发研究

佛山木版年画的传承与应用研发面临多重挑战。许洁玲（2015）指出宣传力度不足与传承方式式微导致佛山木版年画面临生存危机，强调必须在坚守核心手工技艺及原材料的基础上，通过内容、形式、传承机制、生产组织和销售模式等全方位创新实施“生产性保护”^[8]。刘钟萍（2020）结合自身实践，总结了佛山木版年画面临的三大深层困境——传统节俗淡化引发需求萎缩、手工技艺传承断层、现代印刷技术冲击非遗生产模式，并从符号学视角提出文化记忆重构路径，如将“喜神”与当代“脱单”诉求结合，赋予传统神像新文化寓意^[9]。居华倩与彭文利（2021）继而从技术维度提出数字化解决方案，主张利用三维扫描记录刻板工艺细节，结合VR/AR交互场景增强传播参与度，同时指出需解决数据标准化与长期维护的技术瓶颈^[10]。针对消费文化冲击，林观星与刘钟萍（2022）补充认为需借助媒介创新与设计赋能适应现代审美，强调“新型文化媒介人”在连接传统工艺与大众消费中的关键作用，推动“审美经济化”转型^[11]。彭梦玲与曾智琳

（2022）进一步提出需推动生活化融入与创意转化以实现活态传承，具体路径包括开发符合现代审美的文创产品，并构建“传承人+设计师+高校”协同平台，推动技艺创新与市场化闭环，形成可持续的“自我造血”机制^[12]。

（四）现有研究的不足与本研究的价值

现有研究在探讨佛山木版年画的文化内涵、技艺特征及传承应用等方面虽有所推进，但其存在明显局限。这些研究未能充分探讨公众的传承意愿及其影响因素，也缺乏对针对性创新发展路径的深入剖析。因此，现有提出的建议难以形成系统解决方案，无法有效应对传承断层和市场萎缩等核心困境。为此，本研究采用问卷调查与实证分析相结合的方法，以佛山市居民为研究对象，聚焦其认知现状、传承意愿与受众群体特征展开研究，旨在为佛山木版年画的可持续传承提供实证依据与实践对策。

二、数据采集与样本特征

（一）抽样方法与样本容量

本研究以佛山市常住居民为调查对象，采用分层抽样与三阶不等概率PPS抽样法确保样本代表性。第一阶段按人口密度将佛山市划分为人口密集区（南海区、顺德区）与稀疏区（禅城区、三水区、高明区），抽取3个区县作为一级入样单位；第二阶段从入样区县中随机抽取10个街道；第三阶段在街道范围内采用线上随机发放与线下等距拦访相结合的方式抽取受访者，线上占比70%，线下占比30%。最终的样本抽样框设计如表1所示。

表1：样本抽样框

Table 1: Sample Sampling Frame

分层	一级单位 抽样框	入样 区县	二级单位 抽样框	入样街道 / 镇	三级单位 抽样框
人口 密集区	南海区 顺德区	南海区	下辖7个街道 / 镇	桂城街道、西樵镇、里水镇、丹灶镇、大沥镇	各街道的 往来行人
		顺德区	下辖10个街道 / 镇	大良街道、勒流街道、杏坛镇、乐从镇	
人口 稀疏区	禅城区 高明区 三水区	三水区	下辖7个街道 / 镇	乐平镇	

基于预调查中了解程度 ≥ 5 的群体占比估计总体比例，关注于总体比例的样本方差，得到修正前的样本量的计算公式为（1）：

$$n_0 = \frac{t^2 p(1-p) / d^2}{1 + \frac{1}{N} [\frac{t^2 p(1-p)}{d^2} - 1]} = \frac{N t^2 p(1-p)}{(N-1) d^2 + t^2 p(1-p)} \quad (1)$$

其中为 N 总体单位的数量， P 为样本比例， d 为绝对允许误差，置信度为95%，则可确定 t 值为1.96。预调查结果表明 P

=0.58, 佛山市常住居民总数为9615400人, 当绝对误差 d 不超过4%, 则可以得出近似最佳样本量为:

$$n_0 = \frac{9615400 \times 1.96^2 \times 0.58 \times (1 - 0.58)}{(9615400 - 1) \times 0.04^2 + 1.96^2 \times 0.58 \times (1 - 0.58)} \approx 585$$

基于简单随机抽样 (95% 置信度, 误差 $\pm 4\%$), 结合预调查结果, 计算所需样本量为585份。考虑问卷发放中存在操作失误、用户弃填或数据失真等情况, 假设问卷无效率设为15%, 本次调查最终发放问卷688份。最终, 有效回收问卷590份, 有效率为85.7%。信度检验显示 Cronbach α 系数为0.840, 效度检验 KMO 值为0.812, 表明数据信效度良好, 适合后续分析。

（二）问卷设计

本次问卷调查采集了受访者的 人口统计学特征 (性别、年龄、学历、职业、收入等), 并围绕佛山木版年画的文化价值感知, 系统考察了其文化认知度、文化认同感、民族自豪感, 以及行为意向 (购买意愿、参与意愿、传承意愿) 三大核心维度。

（三）样本特征

根据第七次人口普查数据显示, 佛山市居民男女年龄比为54:46, 中青少年人口 (60岁以下) 与老年人 (60岁以上) 占比为92:8。本次调查对象在性别与代际分布上呈现均符合人口普查数据分布规律, 男女比例约为1:1, 调查者中青年人占比95%, 老年人占比5%, 能够提供较客观准确的数据。教育水平、收入水平、职业等分布广泛且均衡。

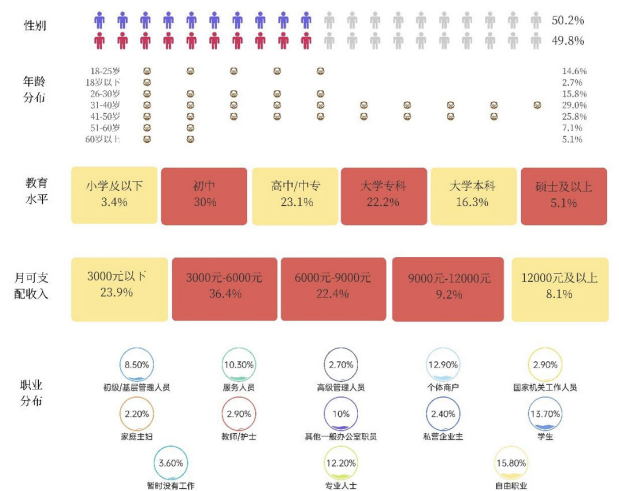


图1：样本特征分布图

Figure 1: Description of sample characteristics distribution

三、实证分析

（一）基于列联表的佛山木版年画认知现状分析

本文基于列联表分析方法探析不同特征人群对佛山木版年画的认知程度差异。表2显示了不同年龄群体对佛山木版年画了解程度的频数分布。Person 卡方独立性检验显示, 不同年龄群体对佛山木版年画的了解程度存在显著差异 ($\chi^2 = 228.769, p = 0.000$)。中老年群体因历史背景与生活经验更贴近传统文化, 而年轻人

受现代环境影响较疏离。

表2：不同年龄群体对佛山木版年画了解程度的频数分布
Table 2: Frequency distribution of knowledge level by age group

题目	名称	年龄							总计
		18岁以下	18-25岁	26-30岁	31-40岁	41-50岁	51-60岁	60岁以上	
了解程度	非常不了解	0	17	7	0	0	0	1	25
	很不了解	0	7	0	1	1	0	0	9
	较不了解	0	7	4	2	2	1	2	18
	不太了解	1	9	13	7	2	0	0	32
	一般了解	1	12	7	16	4	2	0	42
	较了解	6	17	24	21	17	8	5	98
	很了解	5	12	15	27	30	10	3	102
	非常了解	1	5	11	38	34	11	4	104
	极了了解	2	0	6	32	38	5	6	89
	完全了解	0	0	10	23	24	5	9	71
总计		16	86	97	167	152	42	30	590

表3显示不同学历群体对佛山木版年画的了解程度存在显著差异 ($\chi^2 = 228.027, p = 0.000$)。文化浸润深度与生活重心差异共同作用于非遗传承, 初高中群体因本地文化场域浸润更易接触年画, 了解程度较高; 本科群体则因学习生活目标转移而与其关联弱化, 了解程度低。

表3：不同学历群体对佛山木版年画了解程度的频数分布
Table 3: Frequency distribution of knowledge level by educational attainment

题目	名称	学历 (包含目前在读)						总计
		小学及以下	初中	高中 / 中专	大学专科	大学本科	研究生	
了解程度	非常不了解	0	2	0	2	19	2	25
	很不了解	0	0	0	0	9	0	9
	较不了解	0	0	5	1	10	2	18
	不太了解	2	4	5	6	14	1	32
	一般了解	0	5	11	7	17	2	42
	较了解	3	19	12	26	30	8	98
	很了解	2	30	32	19	15	4	102
	非常了解	1	39	29	18	9	8	104
	极了了解	4	41	24	13	6	1	89
	完全了解	8	37	18	4	2	2	71
总计		20	177	136	96	131	30	590

表4显示，不同月均可支配收入群体对佛山木版年画的了解程度存在显著差异（ $\chi^2 = 91.730, p = 0.000$ ）。经济门槛会影响对传统文化的接触机会，较高收入群体更易获取传统文化信息与活动资源，拥有更多的参与时间与渠道。

表4：不同收入群体对佛山木版年画了解程度的频数分布
Table 4: Frequency distribution of knowledge level by income group

题目	名称	月人均可支配收入（元）					总计
		3000元以下	3000-6000元	6000-9000元	9000-12000元	12000元以上	
了解程度	非常不了解	15	6	4	0	0	25
	很不了解	5	0	3	0	1	9
	较不了解	5	5	6	2	0	18
	不太了解	11	9	5	2	5	32
	一般了解	14	11	11	2	4	42
	较了解	23	29	22	14	10	98
	很了解	15	36	30	8	13	102
	非常了解	16	37	27	18	6	104
	极了解	21	41	17	6	4	89
	完全了解	16	41	7	2	5	71
总计		141	215	132	54	48	590

表5显示不同职业群体对佛山木版年画的了解程度存在显著差异（ $\chi^2 = 251.530, p = 0.000$ ）。自由职业者与专业人士深度认知占比突出；学生群体依赖课程基础认知但实践薄弱；国家机关人员及私营企业主要受限于信息渠道，整体认知不足；个体商户与家庭主妇聚焦年俗功能认知，技艺传承理解浅显。此分化本质源于文化接触场景差异与职业信息生态共构作用。

表5：不同职业群体对佛山木版年画了解程度的频数分布
Table 5: Frequency distribution of knowledge level by occupational group

题目	名称	职业分布													总计
		①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩	⑪	⑫	⑬	
了解程度	非常不了解	16	0	0	1	1	2	2	0	2	0	0	1	0	25
	很不了解	5	0	0	1	1	0	0	0	1	1	0	0	0	9
	较不了解	6	0	3	1	0	3	0	0	1	1	1	1	1	18
	不太了解	7	4	6	7	1	1	2	1	0	1	0	1	1	32
	一般了解	10	7	4	6	2	4	2	2	2	0	1	0	2	42
	较了解	20	13	16	21	2	5	4	4	3	3	3	2	2	98
	很了解	10	16	13	12	5	8	12	11	3	3	4	2	3	102
	非常了解	4	12	21	13	2	6	17	12	3	6	2	3	3	104
	极了解	2	12	19	7	1	11	12	16	1	1	2	5	0	89
	完全了解	1	12	11	3	1	10	8	15	1	1	1	5	2	71
总计		81	76	93	72	16	50	59	61	17	17	14	20	13	590

注：①学生；②个体商户；③自由职业；④专业人士；⑤高级管理人员；⑥初级 / 基层管理人员；⑦其他一般办公室职员；⑧

服务人员；⑨教师 / 护士；⑩国家机关工作人员；⑪私营企业主；⑫暂时没有工作；⑬家庭主妇

本文亦检验了受访者年龄与佛山木版年画了解程度的关联关系，结果显示不同性别对佛山木版年画了解程度没有显著性（ $\chi^2 = 8.492, p = 0.052$ ）。

（二）基于结构方程模型的佛山木版年画传承意愿影响因素分析

本研究基于“多元认知-情感整合-行为承诺”的理论框架，剖析佛山模板年画从多元认知到深度传承的转化路径。

1. 结构方程模型的理论与构建

本文选取文化认知、文化认同、参与意愿、民族自豪感、传承意愿作为结构变量，并制定研究假设。文化认知指公众对佛山木版年画本质特征的知识性掌握，通过地域特色、艺术成果、风土人文、艺术特色四维观测变量来衡量；文化认同指公众对年画文化符号产生的归宿感与价值共识，包含了历史意义和符号认同两个行为表征；参与意愿指个体介入年画实体活动及主动传播的初期行为倾向，从活动参与和主动宣传维度来衡量；民族自豪感指个体因年画衍生的民族文化价值情感共鸣，包含情绪感知和价值认可两个心理维度；传承意愿指个体将技艺传承与执事从业纳入人生轨迹的长期承诺。

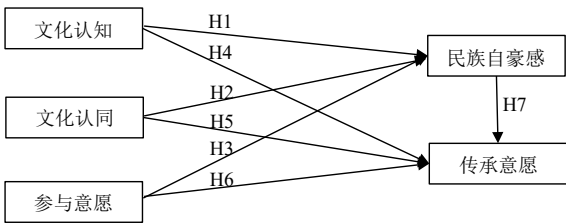


图2：理论模型图和研究假设

Figure 2: Theoretical model and research hypotheses

文化认知、文化认同、参与意愿作为多元认知的不同驱动源，既能通过强化情感枢纽催化传承意愿，又能锚定行为目标直接触达传承意愿（图2）。一方面，三者分别从知识解码、符号归属、实践接触角度切入，共同强化个体对佛山木版年画非遗价值的情绪感知与价值认可，使民族自豪感成为关键转化引擎——既催化技艺传承的习得动力，又激发执事从业的职业化承诺。与此同时，三大驱动源又能直接锚定行为目标，文化认知通过知识内化推动行为转化、文化认同借由符号归属激发使命承诺、参与意愿依托实践反馈加速行为决策，直接提升个体的传承意愿。因此，本文提出以下假设：

- H1：“文化认知”对“民族自豪感”有正向影响
- H2：“文化认同”对“民族自豪感”有正向影响
- H3：“参与意愿”对“民族自豪感”有正向影响
- H4：“文化认知”对“传承意愿”有正向影响
- H5：“文化认同”对“传承意愿”有正向影响
- H6：“参与意愿”对“传承意愿”有正向影响
- H7：“民族自豪感”对“传承意愿”有正向影响

表6为潜变量、显变量和具体测量题项。本文运用李克特5级

量表进行打分，以保障数据的准确性和稳定性，

表6：结构方程模型潜变量与显变量

Table 6: Latent and manifest variables in the structural equation model

潜变量	显变量	测量题项
文化认知	地域特色	我了解佛山木版年画与广府文化的密切联系。
	艺术成果	我认可佛山木版年画的雕刻技法与色彩运用具有很高的艺术独创性。
	风土人文	我相信佛山木版年画在婚嫁、祭祀等仪式中发挥着凝聚社群的重要作用。
	艺术特色	我能够清晰辨识佛山木版年画独特的艺术风格。
文化认同	历史意义	我认同佛山木版年画的历史见证价值。
	符号认同	我将门神、瑞兽等年画图案视为代表岭南文化身份的重要符号。
活动参与	活动参与	我愿意定期参加木版年画制作工坊或专题展览等实体活动。
	自主宣传	我会主动在社交媒体分享佛山木版年画的知识 and 鉴赏心得。
民族自豪感	情绪感知	接触佛山木版年画时，我会产生强烈的文化归属感与敬畏心。
	价值认可	我认为保护佛山木版年画对维护中华文化多样性具有战略意义。
传承意愿	技艺传承	我愿意系统学习雕版、套印等佛山木版年画核心制作技艺。
	执事从业	我考虑未来将木版年画创作或研究作为长期职业发展方向。

2. 结构方程模型的适配度检验

本文将上述题项收集到的数据导入 AMOS 软件中，对模型适配度进行检验，结果如表 7 所示。增量适配指标 IFI、CFI、TLI 均达 0.96 以上，绝对适配指标 RMSEA < 0.05，各项指标都达到最佳标准，说明模型适配度优秀，此模型对于调查问卷所得的样本数据具有较好的拟合效果。

表7：模型适配度检验

Table 7: Model fit assessment

指标	参考标准	实测结果
IFI	> 0.9 为优秀	0.975
CFI	> 0.9 为优秀	0.975
TLI	> 0.9 为优秀	0.960
RMSEA	< 0.05 为优秀	0.049

3. 研究假设检验

模型标准化系数路径分析如图 3 所示。路径分析显示，在 5% 显著性水平下，文化认知（ $p=0.048$ ）、文化认同（ $p=0.041$ ）及参与意愿（ $p=0.034$ ）显著正向影响民族自豪感；民族自豪感（ $p<0.001$ ）显著正向驱动传承意愿，验证了本文假设 H1、H2、H3 和 H7，说明文化认知、文化认同、参与意愿作为多元认知的不同驱动源，通过强化情感枢纽提升了个体的传承意愿；同时，文化认知（ $p=0.008$ ）、文化认同（ $p=0.019$ ）及参与意愿（ $p=0.034$ ）显著正向影响传承意愿，验证了本文假设 H4、H5、H6，说明文化认知、文化认同、参与意愿同时能锚定行为目标，直接提升个体的传承意愿。

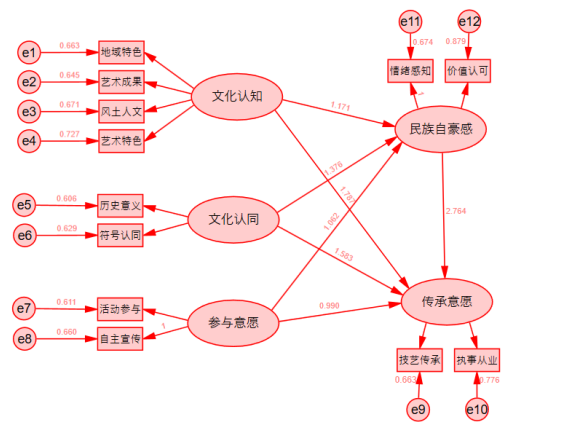


图3：标准化系数路径图

Figure 3: Standardized coefficient path diagrams

（三）基于 K-Means 聚类分析的佛山木版年画受众人群画像挖掘

为精准划分佛山木版年画受众人群特征，本研究基于认知程度、购买意愿、传承意愿三大聚类变量，通过 K-Means 算法进行群体画像挖掘。通过聚类分析，本文识别出三种不同的受众群体，最终聚类中心统计结果如表 8 所示：

表8：最终聚类中心统计表

项	最终聚类中心		
	认知程度	购买意愿	传承意愿
1	1.259	0.315	1.139
2	-0.521	0.442	0.109
3	0.304	0.819	-1.181

基于上述最终聚类中心统计结果，我们识别出三类具有鲜明特征的受众群体，其典型画像如图 4 所示：



图4：三类受众群体画像

Figure 4: Tripartite audience typology profiling

第一类人群为传承探索型人群，该群体对佛山木版年画的认知等级较高，传承意愿强烈，但其购买意愿相对理性。这个群体可能更倾向于雕刻实践等深度体验，将技艺传承视为文化责任而非单纯消费行为。对于这类受众，可强化其职业化发展通道，将技艺研习转化为可持续生计来源。

第二类人群为文化疏离型人群，该群体呈现认知浅层化与负面化、购买意愿低、传承意愿缺失。其疏离主因可能是信息触达不足与创新感知滞后。对于这类受众，需通过数字化媒介重塑年画形象，以创新载体扭转其刻板认知。

第三类人群为浅知待塑型人群，该群体购买意愿突出、但认知与传承意愿薄弱。其消费行为呈现轻量化特征——偏好成品收藏，更关注产品装饰性而非文化内涵，可转化为“文化消费客

群”但需进行引导深化认同。对于这类受众，应建立消费行为与深度体验的联动机制，通过轻量化消费引导其参与技艺实践。

四、结束语

本研究系统解析了公众对佛山木版年画的认知现状与传承意愿，并基于聚类分析实现受众群体精准画像。结果表明，大众对

佛山木版年画的认知呈现显著结构化分化，代际断层、学历隔离、收入分层与职业场景差异共同构成认知鸿沟的立体维度；文化认知、文化认同与参与意愿作为三大核心驱动力，既通过情感枢纽（民族自豪感）间接催化传承意愿，又能直接锚定行为目标触发传承行动；进一步提炼出传承探索型、文化疏离型、浅知待塑型三类受众群体，为定制差异化干预策略提供精准坐标，以系统性破解非遗传承的意愿转化瓶颈。

参考文献

- [1] 段鹏, 张丁. 互动·互嵌·互塑: 数智技术赋能中华优秀传统文化 [J]. 出版广角, 2024, (06): 25-30+37.
- [2] 谢春. 非物质文化遗产保护理念的当代变迁 [J]. 艺术传播研究, 2024, (01): 32-40.
- [3] 李亚娟, 谢双玉. 非遗的活态传承: 传承人主体性激发路径 [J]. 云南民族大学学报 (哲学社会科学版), 2023, 40(05): 49-51.
- [4] 徐秋枫. 佛山木版年画的象征研究 [J]. 艺术教育, 2010, (01): 14-15.
- [5] 程宜. 从佛山木版年画看佛山民众的信仰 [J]. 收藏家, 2012, (09): 61-66.
- [6] 许结玲. 佛山木版年画的文化内涵解读 [J]. 广西师范学院学报 (哲学社会科学版), 2016, 37(01): 37-41.
- [7] 李晗. 佛山年画中道教信仰的审美体现与影响 [J]. 北极光, 2019, (02): 143-144.
- [8] 许结玲. 国家级非遗佛山木版年画的现状、传承与保护 [J]. 文化遗产, 2015, (06): 146-152+4-5.
- [9] 刘钟萍. 佛山木版年画里的“诸神复活”——年轻传承人五年里的传承新尝试 [J]. 年画研究, 2020, (00): 217-224.
- [10] 居华倩, 彭文利. 数字化技术在非物质文化遗产保护中的应用——以佛山木版年画为例 [J]. 湖南包装, 2021, 36(04): 111-114.
- [11] 林观星, 刘钟萍. 消费文化语境下佛山木版年画工艺传承与振兴 [J]. 中国艺术, 2022, (02): 106-112.
- [12] 彭梦玲, 曾智林. 佛山木版年画的活态传承与创新研究 [J]. 工业设计, 2022, (05): 140-142.

当代广州市大学生对疗愈产品的购买意向 及影响机制研究

张桢岚¹, 张少华^{1*}, 李思妍¹, 吴永桢¹, 李宇轩², 黄校英²

1. 广东工业大学经济学院, 广东 广州 510000

2. 广东工业大学材料与能源学院, 广东 广州 510000

DOI:10.61369/ASDS.2025070004

摘 要 : 近年来, 伴随社会心理的深刻变迁, 在政策引导与市场驱动下, 疗愈经济作为融合情绪价值与符号认同的新兴消费形态快速兴起。本文聚焦广州市大学生群体, 围绕“情绪价值 + 符号价值”双重驱动下的消费行为动因识别问题展开实证研究。借助大规模问卷调查, 在样本特征识别基础上, 先运用 K-means++ 聚类划分为三类典型消费群体, 再结合随机森林与 XGBoost 模型进行购买意向预测, 并通过结构方程模型分析行为态度、主观规范、知觉行为控制与感知价值对购买意向的作用机制。研究发现, 广州市大学生自我感知压力大, 愿意为“情绪疗愈”买单, 其疗愈消费行为呈现显著的“情绪驱动 + 符号认同”特征。行为态度、主观规范、知觉行为控制和感知价值正向影响购买意向, 且感知价值在行为态度、主观规范、知觉行为控制作用购买意向的路径中具有关键中介效应。

关 键 词 : 疗愈经济; 购买意向; 计划行为理论; 感知价值

A Study on the Purchase Intention and Influencing Mechanisms of Healing Products Among Contemporary University Students in Guangzhou

Zhang Zhenlan¹, Zhang Shaohua^{1*}, Li Siyan¹, Wu Yongzhen¹, Li Yuxuan², Huang Jiaoying²

1. School of Economics, Guangdong University of Technology, Guangzhou, Guangdong 510000

2. School of Materials and Energy Economics, Guangdong University of Technology, Guangzhou, Guangdong 510000

Abstract : With a profound shift in the social psyche in recent years, under the dual influence of policy guidance and market drivers, the healing economy has rapidly emerged as a novel consumption paradigm integrating emotional value and symbolic identity. This study focuses on university students in Guangzhou and conducts an empirical research on the motivational drivers of their consumption behavior under the dual influence of "emotional value with symbolic value." Through a large-scale questionnaire survey, the study first applies K-means++ clustering to categorize respondents into three typical consumer groups based on sample characteristics. Subsequently, random forest and XGBoost models are applied to predict purchase intentions, while a structural equation model (SEM) analyzes the mechanisms through which behavioral attitudes, subjective norms, perceived behavioral control, and perceived value influence purchase intentions. The findings reveal that university students in Guangzhou experience high levels of self-perceived stress and are willing to pay for "emotional healing." Their healing consumption behavior exhibits distinct "emotion-driven + symbolic identity" characteristics. Behavioral attitudes, subjective norms, perceived behavioral control, and perceived value all positively influence purchase intentions, with perceived value playing a critical mediating role in the pathways through which behavioral attitudes, subjective norms, and perceived behavioral control affect purchase intentions.

Keywords : healing economy; purchasing intention; theory of planned behavior; perceived value

基金项目: 2024 年度广东省教育科学规划项目“新商科视域下‘赛教融合’的经济学人才培养模式探索与实践”(2024GXJK625); 2023 年度广东省社科规划项目“粤港澳大湾区先进制造业和现代服务业深度融合的双向赋能路径研究”(GD23XYJ21); 2022 年度广东工业大学“特色智库”建设引导专项课题“粤港澳大湾区生产性服务业发展指数评价研究”(2022TSZKZX007)。

作者简介:

张桢岚, 本科在读, 广东工业大学经济学院, 研究方向: 金融衍生品设计;

李思妍, 本科在读, 广东工业大学经济学院, 研究方向: 国际与贸易;

吴永桢, 本科在读, 广东工业大学经济学院, 研究方向: 金融工程证券投资;

李宇轩, 本科在读, 广东工业大学材料与能源学院, 研究方向: 能源与动力工程制冷;

黄校英, 本科在读, 广东工业大学材料与能源学院, 研究方向: 能源与动力工程制冷。

通讯作者: 张少华, 广东工业大学经济学院, 讲师, 管理学博士, 研究方向为城市服务经济。

引言

随着社会心理需求的变化，传统功能性消费已难以满足年轻人对情感价值的深层次诉求，从而催生融合情绪价值与符号认同的“疗愈经济”新型消费形态。在此背景下，本文以广州市大学生为研究对象，基于642份有效调查问卷数据，整合K-means++聚类、机器学习预测模型（随机森林与XGBoost）和结构方程模型等分析方法，系统考察疗愈消费的行为特征与决策机制。本文通过构建包含行为态度、主观规范、知觉行为控制和感知价值的分析框架，尝试揭示“情绪-符号”双驱动的新型消费逻辑，拓展了TPB理论在情绪型消费领域的适用性。同时实证分析结果可以为企业疗愈产品的开发与营销策略制定提供重要参考。

一、文献综述

随着全球心理健康问题加剧，情绪性消费逐渐成为主流趋势，推动“疗愈经济”作为新兴消费形态迅速崛起。从体验消费理论（Holbrook & Hirschman, 1982）来看，商品的情绪价值与象征意义正逐步取代其功能属性，成为驱动消费行为的重要动因^[1]。这一观点在近年来研究中得到了印证。王为久等（2022）指出，盲盒消费中的“情绪不确定性”显著激发了消费者的情绪参与感^[2]；陆瑶（2017）则发现，迪士尼IP产品中的“情感资本”能够增强用户忠诚度并引发情绪共鸣^[3]。

在此背景下，学界日益关注符号消费机制在疗愈型产品中的嵌入作用。鲍德里亚（2001）指出，社会消费行为逐步从物质需求向符号价值需求转变，当商品承载特定符号认同时，消费者愿意为“情绪溢价”支付更高成本^[4]。商品的符号价值存在自我认同和社会认同两个维度（贾中海和李娜，2021）^[5]，消费者选购商品后，商品的符号价值可以让消费者产生身份认同，满足社会认同的需要（由军强，2023）^[6]。马睿（2023）发现，Z世代人群在“电子木鱼”“丧文化周边”等消费行为中，体现出对身份认同与情绪表达的双重追求^[7]。然而“情绪营销”过于泛化可能带来“符号通胀”问题，即符号被资本异化后失去其原有情感意义，从而削弱消费者的共鸣基础（田芯等，2018）^[8]。

从行为解释来看，计划行为TPB理论（Ajzen, 1991）为情绪消费提供了理论工具^[9]，后续拓展研究重点强调主观规范与知觉行为控制可通过“情绪价值”间接影响行为意向（Rindfleisch et al, 2009）^[10]。但目前多数TPB研究仍聚焦于理性消费领域，对大学生情绪型购买行为中“感知价值”与“符号属性”的整合分析不足。

在研究方法层面，疗愈经济领域的定量研究正逐步从线性路径模型向数据驱动模型演化。比如Blei et al.（2003）运用LDA主题模型挖掘社交评论中的亚文化表达，如利用“电子功德+1”等语义标签反映焦虑戏谑^[11]。杨千帆等（2024）则利用随机森林模型对消费行为进行预测^[12]。

综上，当前研究尚存在不足：其一，过度聚焦于功能型疗愈产品（如香薰、按摩仪等），忽视符号类疗愈产品的情绪价值识别与量化；其二，TPB模型与机器学习的融合尚不成熟，定量研究相对不足；其三，对特定地域的消费群体分层结构与心理机制路径的联合探讨较少。因此，本文在TPB理论上引入感知

价值变量，以广州市大学生群体为研究对象，通过大规模问卷调查，整合K-means++聚类、机器学习预测模型和结构方程模型等方法构建大学生疗愈消费“识别-预测-解释”的全链路实证框架，系统考察大学生疗愈产品消费的行为特征与决策机制。

二、数据来源与探索性分析

本文在问卷设计与抽样方法上进行了严谨考量，以确保数据的可靠性和有效性。在问卷设计方面，研究团队基于前期对“疗愈经济”相关文献的系统梳理和网络舆情分析，通过多轮研讨修订，并经过线上预调查发放回收118份问卷，结合预调查信效度情况和受访者反馈修订问卷，最终形成了涵盖基本人口信息、疗愈产品认知、购买意向、行为态度、主观规范、知觉行为控制与感知价值维度的调查问卷。问卷语言简洁明了，问题设置科学客观且易于理解，具有良好的可答性和测量针对性。

（一）基本数据信息

在抽样方法上，考虑到广州市高校的多样性和代表性，研究采用分层抽样结合配额抽样的策略。根据广东省教育厅公布的广州市高校名录，从全市38所本科院校中选取12所不同类型（包括综合类、理工类、师范类等）和规模的院校作为样本来源。

最佳样本量的计算公式为

$$n = \frac{(Z_{\alpha/2})^2 \sigma^2}{\Delta^2} \quad (1)$$

其中置信度为95%， Δ 取值为5%，z值为1.96， σ 取值为0.64，可以计算出所需的样本容量为629。考虑到可能存在无效问卷和未回收问卷，我们将预调查的问卷有效率作为参考，得到发放问卷数量为800份。问卷通过线上线下渠道发放（线上占比为2/3），经严格筛选剔除无效问卷后最终获得642份有效问卷，有效率为80.25%。所有数据均经过多次核对校验，确保录入准确性。如表1所示，正式调查问卷的整体信度系数为0.920，行为态度、主观规范等二级维度的信度系数均大于0.7，信度良好。而在效度上，总体KMO数值达到0.868，Bartlett球形检验的显著性为0.000。信效度检验结果表明，所获数据具有较高的可靠性和有效性。

表1：问卷整体和分维度信度情况
Table 1: Reliability of the overall questionnaire and sub-dimensions

题项	项数	Cronbach's α 系数
整体	18	0.932
BA 行为态度	3	0.830
SN 主观规范	4	0.785
PBC 知觉行为控制	4	0.804
PV 感知价值	4	0.790
BI 购买意向	3	0.839

如表2所示，受访者男女比例均衡，受教育水平相对均衡，大部分人表示接触过疗愈类产品，且对“疗愈经济”“情绪疗愈”等概念不陌生。值得注意的是，不同性别、专业背景、消费能力和压力感知的个体在消费频率与偏好上存在显著差异。但总体而言，当代青年自我感知压力大，愿意为“情绪疗愈”买单，疗愈需求已深度嵌入广州大学生的日常生活。

表2：受访者基本信息情况
Table 2: Basic information of respondents

维度	类别	占比	类别	占比
性别	男	49%	女	51%
专业类别	文科类	27%	理科类	26%
	工科类	24%	其他	23%
自我压力感知	轻度	15%	中度	44%
	重度	32%	极度	5%
年级	本科（每个年级平均约占）	20%	研究生及以上	17%
可支配收入	1000元 -1500元	33%	1501元 -2000元	30%
	2001元 -3000元	19%	3000元以上	11%

（二）基于 K-means++ 方法的人群细分

为进一步挖掘大学生疗愈消费群体的内部结构，本文运用 K-means++ 聚类方法对样本进行了群体划分。

1. 理论基础与聚类方法

本文基于计划行为理论，假设不同消费群体在价值感知、行为态度等维度存在显著差异特征。为系统考察这一假设，采用 K-means++ 聚类方法，选取学历、每月可支配收入、消费频率、压力感知和溢价认同感五个关键指标作为聚类变量，旨在揭示疗愈产品消费群体的差异化行为特征和市场细分结构。

2. 方法挖掘过程

采用标准化处理消除变量量纲差异，并运用手肘法确定最优聚类数 k。该方法通过计算不同 k 值对应的组内平方和 SSE 式 2），观察 SSE 下降速率的拐点来确定最佳聚类数。经分析，最终确定 k=3 为最优聚类数。

$$SSE = \sum_{i=1}^k \sum_{x \in C_i} x - \mu_i^2 \quad (2)$$

3. 聚类结果与分析

基于 k=3 的聚类分析结果显示，广州市大学生疗愈产品消费者可划分为三类典型群体（表3）。（1）情绪敏感型群体表现出较高的压力感知和情感价值认同，疗愈消费已经成为其压力缓解的重要手段，其消费行为具有明显的情绪驱动特征，这一结果反映了当代大学生群体对情绪疗愈的强烈需求。这类群体对价格敏感，易受短期促销活动影响。（2）高消费核心型群体在可支配

收入和品牌溢价认同方面表现突出，显示出较强的消费能力和对产品精神价值的重视，愿意为品牌情感价值付费，更注重产品品质，是疗愈市场的核心客群。（3）低频保守型群体则呈现出理性消费特征，消费行为受经济能力主导，在疗愈产品选择时关注性价比，不愿为品牌溢价买单。

表3：购买疗愈产品消费者聚类人群细分表
Table3: Consumer segmentation of healing products based on clustering

维度	全局均值	低频保守型群体	高消费核心群体	情绪敏感型群体
学历	3.11	2.75	4.40	2.44
每月可支配收入	3.01	2.62	3.71	2.76
购买频率	3.12	1.27	3.14	4.12
压力感知	3.14	2.96	3.24	3.17
溢价认同感	3.21	2.89	3.51	3.18

值得注意的是，我们发现疗愈产品的消费行为具有显著的符号化特征，特别是在高消费核心型和情绪敏感型群体中，产品的情感价值认同超越了其功能性价值。这一结论与马斯洛需求层次理论契合，揭示了在基本的物质需求得到满足后，当代大学生消费者对心理慰藉和情感满足的追求日益增长。

三、疗愈消费的购买意向预测与决策分析

（一）基于机器学习的购买意向预测分析

在建模过程中，我们从人口统计、经济行为和情绪驱动三个维度选取了8个初始变量。为优化模型性能，先通过随机森林评估变量重要性，再结合 SHAP 值分析验证特征贡献度，最终筛选出5个核心预测变量。数据经过标准化和独热编码处理后，采用交叉验证进行模型训练与参数优化。

随机森林是一种基于集成学习的机器学习算法，通过构建多棵决策树并聚合其预测结果，能够有效处理高维数据和非线性关系。其核心思想是通过“投票法”或“平均法”降低单棵树的过拟合风险，提升模型泛化能力。

$$\hat{y} = \frac{1}{N} \sum_{i=1}^N T_i(x) \quad (3)$$

其中， T_i 为第 i 棵决策树的预测结果， N 为树的总数。

XGBoost 是一种提升树（Boosting Trees）模型，其训练过程的核心在于不断最小化目标函数（Objective Function），并通过加法模型和正则化手段实现高效拟合与防止过拟合。

$$Obj^{(t)} = \sum_{i=1}^n l\left(y_i, \hat{y}_i^{(t)}\right) + \sum_{j=1}^T \Omega(f_j) \quad (4)$$

$$Obj_{min}^{(t)} = -\frac{1}{2} \sum_{j=1}^T \frac{G_j^2}{H_j + \lambda} + \gamma T \quad (5)$$

从 XGBoost 回归下变量检验数据结果（表4）来看，大学生

疗愈产品消费行为呈现以下关键特征：①消费习惯是最核心的预测指标，购买频率的重要性得分（0.4000）显著高于其它变量，表明既往消费行为能有效预测未来购买意向，消费习惯养成一定程度上会产生对疗愈产品的依赖性。②独生子女身份（0.1925）和可支配收入（0.1628）作为社会经济特征，对消费决策的影响排名第二和第三。独生子女可能因家庭经济支持更充裕或心理需求不同，更倾向于购买疗愈产品。较高的可支配收入可能提高购买意向。③压力感知（0.1206）虽影响相对较小，但证实了情绪因素在消费行为中的作用，压力高的人更可能购买疗愈产品。同时 ROC 曲线分析显示机器学习模型的 AUC 值达0.8209，表示模型具有较强的区分能力和良好的预测效能，特别在识别高频用户和独生子女群体方面表现突出。

表4：XGBoost 回归下变量检验数据结果

Table 4: Variable test results under XGBoost regression

独热编码	Feature	Importance
0	半年内，购买“疗愈”产品的频率	0.4000
4	是否为独生子女	0.1925
1	每个月的可支配收入为	0.1628
3	三个月内，购买“疗愈”产品的金额	0.1241
2	压力感知程度	0.1206

综合来看，机器学习模型分析结果说明大学生疗愈产品消费呈现出“习惯主导、经济调节、情绪补充”的复合驱动模式，其中消费习惯的持续性特征最为突出，而社会经济因素和情绪需求则在不同层面调节着消费决策。

（二）基于 TPB 理论的行为机制分析

为深入解析疗愈产品消费决策的心理机制，接下来本文基于计划行为理论（TPB）与感知价值理论的整合框架，我们通过结构方程模型深入解析疗愈产品消费决策的机制。从理论层面来看，一方面行为态度、知觉行为控制及主观规范、感知价值可以直接影响消费者的购买意向，另一方面行为态度、知觉行为控制及主观规范又通过影响感知价值进而作用于消费者的购买意向（图1）。

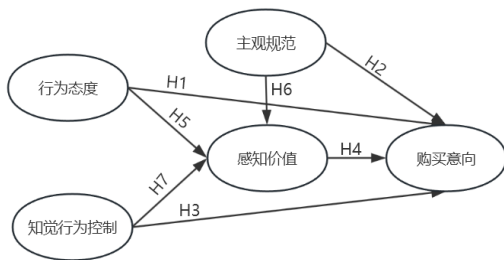


图1：理论模型图

Figure1: Theoretical model

基于此，本文提出以下研究假设：

H1：行为态度正向影响购买意向

H2：主观规范正向影响购买意向

H3：知觉行为控制正向影响购买意向

H4：感知价值正向影响购买意向

H5：行为态度正向影响感知价值

H6：主观规范正向影响感知价值

H7：知觉行为控制正向影响感知价值

本研究采用测量模型（式6和式7）和结构模型（式8）相结合的分析方法，重点考察了行为态度、主观规范和知觉行为控制三大核心构念对购买意向的影响路径。

$$X = \Lambda_x \xi + \delta \quad (6)$$

$$Y = \Lambda_y \eta + \varepsilon \quad (7)$$

$$\eta = b\eta + \Gamma \xi + \zeta \quad (8)$$

$$z = \frac{\hat{a}\hat{b}}{\sqrt{\hat{b}^2 s_a^2 + \hat{a}^2 s_b^2}} \quad (9)$$

利用 Amos28.0对模型进行适配度分析，结果如表5所示。可以看到，CFI值、GFI值、AGFI值、NFI值和TLI值均大于0.9，RMSEA 值为0.059，也小于0.08，支持理论模型的有效性，说明该模型能较好地解释变量之间的关系，与数据拟合良好。

表5：模型适配度结果

Table 5: Model fit results

指标	绝对适配度指标			增值适配度指标		
	AGFI	CFI	RMSEA	NFI	GFI	TLI
适配标准	>0.9	>0.9	<0.05 良好； <0.08 合理4	>0.9	>0.9	>0.9
实际拟合值	0.909	0.945	0.059	0.923	0.923	0.935

结构方程模型的标准化路径分析结果如表6所示，可以看到该模型的路径系数全部显著，主观规范对购买意向的影响路径通过了5%的显著性水平检验，其它路径则均通过了1%的显著性水平检验。行为态度、主观规范、知觉行为控制和感知价值对购买意向产生了显著的正向影响，假设 H1 至 H4 成立；行为态度、主观规范和知觉行为控制对感知价值也产生显著的正向影响，假设 H5 至 H7 成立

表6：标准化路径分析结果

Table 6: Standardized Path Analysis Results

路径		估计系数	标准差	Z 值	P 值	验证结果
行为态度	→ 购买意向	0.153	0.061	2.508	***	H1 成立
主观规范	→ 购买意向	0.276	0.056	4.292	**	H2 成立
知觉行为控制	→ 购买意向	0.272	0.058	4.69	***	H3 成立
感知价值	→ 购买意向	1.055	0.073	14.488	***	H4 成立
行为态度	→ 感知价值	0.383	0.059	6.494	***	H5 成立
主观规范	→ 感知价值	0.644	0.051	12.739	***	H6 成立
知觉行为控制	→ 感知价值	0.564	0.062	9.058	***	H7 成立

注：*、** 和 *** 分别表示在 10%、5% 和 1% 水平上显著。

运用 Amos 进行中介效应分析，感知价值在行为态度、主观规范和知觉行为控制三者对购买意向的影响中均发挥着部分中介

作用（表7），三者在 Sobel 检验中的 Z 值均远大于 1.96，P 值则远小于 1%。这说明在疗愈消费领域，传统 TPB 要素对购买意向的影响呈现出以感知价值为核心的中介传导机制，揭示了情感性消费决策中价值认知的关键中介作用。

表7：感知价值的中介效应检验结果
Table:7: Standardized Path Analysis Results

中介	中介效应	总效应	间接效应占比
行为态度→感知价值→ 购买意向	0.379***	0.651***	部分中介作用 (58.2%)
主观规范→感知价值→ 购买意向	0.392***	0.545***	部分中介作用 (71.9%)
知觉行为控制→感知价值→ 购买意向	0.426***	0.701***	部分中介作用 (60.7%)

注：总效应 = 直接效应 + 中介效应；*、** 和 *** 分别表示在 10%、5% 和 1% 水平上显著。

四、结束语

本文通过对广州市大学生疗愈消费行为的预测和决策分析，揭示了当代青年疗愈消费的新特征。研究发现，不同性别、专业

背景、消费能力和压力感知的个体在疗愈产品消费频率与偏好上存在显著差异。总体来看，广州市大学生自我感知压力大，愿意为“情绪疗愈”买单，他们对疗愈产品的消费行为具有显著的符号化特征。基于聚类分析可将其划分为情绪敏感型、高消费核心型和低频保守型三类群体。行为态度、主观规范、知觉行为控制和感知价值均显著正向影响购买意向，从作用机制来看行为态度、主观规范、知觉行为控制都通过感知价值影响购买意向，即感知价值是该传导机制的重要中介。

从实践角度看，研究结果为企业疗愈产品的开发与营销策略制定提供了重要参考。企业要重视产品的情感价值创造，关注消费群体的文化认同需求，并结合三类典型群体的特点构建“功能导向—情绪赋能—符号认同”的三级产品矩阵，实行精准营销，破解消费决策障碍。

参考文献

[1]Holbrook M B, Hirschman E C. The experiential aspects of consumption: Consumer fantasies, feelings, and fun[J].Journal of Consumer Research, 1982,9(2):132-140.

[2] 王为久, 孟元琦, 徐敏亚. 社会支持与情感动机对盲盒消费的驱动机制分析 [J]. 商业经济研究, 2022,(13):62 - 67.

[3] 陆瑶. 上海儿童对迪士尼品牌忠诚的影响因素研究 [J]. 科技经济市场, 2017,(9): 187-188.

[4] 让·鲍德里亚著（法），刘成富，全志钢译. 消费社会 [M]. 南京：南京大学出版社，2001.

[5] 贾中海, 李娜. 消费社会的符号价值与后现代的主体性丧失 [J]. 社会科学战线, 2021,(5):68-72.

[6] 由军强. 感知符号价值影响消费者购买意向的机理——基于感知风险的中介效应 [J]. 商业经济研究, 2023,(3):84-87.

[7] 马睿. 从“丧”到“躺平”：后疫情时代青年亚文化的转向 [J]. 传媒论坛, 2023,(13):91-93.

[8] 田芯, 徐绍元, 徐桂娣. 从抽象的异化到异化的抽象：符号消费背后的异化逻辑及其本质 [J]. 东北大学学报（社会科学版）, 2018,(20):7-12.

[9]Ajzen I.The theory of planned behavior[J].Organizational Behavior and Human Decision Processes, 1991, 50(2):179-211.

[10]Rindfleisch A, Burroughs J E, Wong N. The safety of objects: Materialism, existential insecurity, and brand connection[J].Journal of Consumer Research, 2009, 36(1):1-16.

[11]Blei D M, Ng A Y, Jordan M I. Latent Dirichlet allocation[J].Journal of Machine Learning Research,2003,(3): 993-1022.

[12] 杨千帆, 李涛, 贾志强. 基于参数优化随机森林模型的消费行为预测算法 [J]. 计算机与数字工程, 2024,(7):1959-1965.

关于“国潮风”下传统中医药零食消费者购买意向调查

李凤娇¹, 曾雪怡², 张俊超³, 吴芯怡⁴, 李雯琴², 黄家良^{2*}

1. 广州大学 数学与信息科学学院, 广东 广州 511400

2. 广州大学 管理学院, 广东 广州 511400

3. 广州大学 计算机科学与网络工程学院, 广东 广州 511400

4. 广州大学 人文学院, 广东 广州 511400

DOI:10.61369/ASDS.2025070005

摘要： 本研究基于国潮风背景，系统探讨传统中医药零食消费者购买意向的影响机制。结果显示，中等偏上经济能力的年轻女性为主要消费群体，其购买决策显著受产品口感、健康功效及品牌知名度驱动。不同群体存在明显差异：女性更偏好美容养颜与缓解疲劳功能，男性更倾向增强免疫力产品；年龄结构影响口感偏好。收入水平调节价格敏感度，中等收入群体消费意愿最强，高收入群体对高价产品接受度更高。结构方程模型验证产品属性对购买意愿的直接效应最强，国潮元素通过美学吸引力与文化认同显著提升购买率。针对行业存在的功效质疑与标准缺失，建议企业实施分层策略，政府推动质量标准与科普教育，并在未来加强口感研发与功效可视化，推动产业可持续发展。

关键词： 消费者偏好；群体差异；国潮元素；价格敏感度；信任危机；分层策略

A Survey on the Purchase Intentions of Traditional Chinese Medicine Snacks Among Consumers under the "National Trend"

Li Fengjiao¹, Zeng Xueyi², Zhang Junchao³, Wu Xinyi⁴, Li Wenqin², Huang Jialiang^{2*}

1.School of Mathematics And Information Science, GU, Guangzhou, Guangdong 511400

2.School of Management, GU, Guangzhou, Guangdong 511400

3.School of Computer Science and Cyber Engineering, GU, Guangzhou, Guangdong 511400

4.School of humanities, GU, Guangzhou, Guangdong 511400

Abstract : This study examines how "National Trend" (Guochao) shapes consumer purchase intentions for traditional Chinese medicine (TCM) snacks. Economically capable young women form the core consumer group, prioritizing taste, health efficacy, and brand awareness. Key group differences emerged: women favor beauty-enhancing and fatigue-relieving products, men prefer immune boosters, and age influences taste preferences. Income moderates price sensitivity—middle-income groups show the highest purchase intent, while high-income consumers accept premium pricing. Structural modeling confirmed product attributes directly drive purchase intention, whereas national trend elements boost purchases through aesthetics and cultural identity. To address efficacy skepticism and quality standard gaps, the study proposes stratified marketing, government-led quality standards and public education, alongside future efforts to optimize taste and efficacy visualization for sustainable industry growth.

Keywords : consumer preferences; group differences; national tide elements; price sensitivity; crisis of trust; layering strategy

引言

本研究聚焦“健康中国2030”背景与后疫情健康意识提升下快速增长的药食同源产业，其年轻化趋势显著，2026年市场规模预计

基金项目：2020年度国家自然科学基金“电子商务系统中游戏设计元素与基本心理需要满足：‘需要-可供性-功能特征’视角”；2024年度广州市基础与应用基础研究专题“零工平台工作游戏化中的动机拥挤效应与激励机制优化研究”。

作者简介：

李凤娇，广州大学数学与信息科学学院，本科，研究方向为数学与应用数学；

曾雪怡，广州大学管理学院，本科，研究方向为会计学；

张俊超，广州大学计算机科学与网络工程学院，本科，研究方向为网络工程；

吴芯怡，广州大学人文学院，本科，研究方向为汉语言文学；

李雯琴，广州大学管理学院，本科，研究方向为电子商务。

通讯作者：黄家良，广州大学管理学院，副教授，博士研究生，研究方向为数字商务。

超7514亿，Z世代消费意愿强。针对当前三大不足——决策变量（口感/功效/国潮）权重缺乏量化验证、人口分层（性别/收入/年龄）调节效应实证缺失、行业信任危机（70.1%消费者存疑），本研究以计划行为理论为框架，结合问卷与访谈，运用T检验、逻辑回归与结构方程等模型，系统量化消费群体分层特征与关键驱动因素，提出分层市场策略与监管建议，为产业可持续增长提供支撑。

一、绪论

（一）研究背景

传统中医药零食源于古代“药食同源”的食疗理念，药材融入食材，造就兼具美味与保健功效的小吃。进入二十一世纪后，健康意识提升与市场扩张推动多样化新品的推出，逐渐吸引大量年轻群体；在传承传统工艺的基础上，现代品牌通过改良配方满足当代口味，借助互联网传播文化并融入国潮元素进行包装设计，使产品成为承载东方韵味与文化价值的载体^[1]。

（二）研究目的

本研究面向各年龄段消费者，旨在通过评估市场潜力并分析当前发展状况，识别影响其购买传统中医药零食的关键因素，并验证这些因素与购买意向的相关性，为传统中医药零食行业提供发展策略与建议。调查核心目标如下：

表1 调查目标

核心消费者画像	决策偏好维度	Purchase Intent 量化	创新与突破路径
年龄、性别、收入；购买行为：购买频率、品类偏好	口感、价格、包装、健康功效	购买意向、市场规模、增长潜力	产品竞争力、品牌影响力、可持续经营

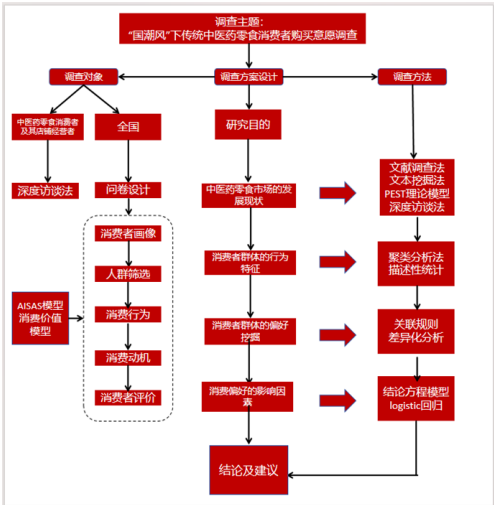


图1 研究框架流程图

二、中医药零食的消费者偏好分析

传统中医药零食消费群体呈现青年化主导的核心趋势：消费者对传统中医药零食的了解程度呈现“广覆盖但浅认知”的倒U型趋势：最显著的特征是大众认知集中在中低水平，即“有一定了解”“一般”及“不太了解”占比最高，而真正深入认知者比例明显偏低。这表明当前消费者对产品的了解多停留在概念层面，

市场亟需深化健康价值与中医药原理的普及教育^[2]。

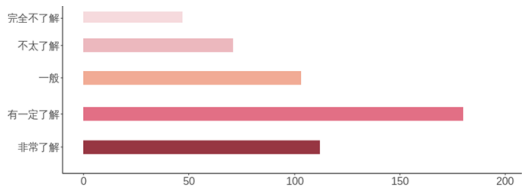


图2 消费者对传统中医药零食的认知状况

三、中医药零食消费偏好的差异化

（一）性别对产品功能偏好的影响

本分析运用独立样本T检验判断性别对产品功能偏好是否存在显著差异，通过效应量衡量差异的实际影响程度。

T检验逻辑：假设“男女对产品功能偏好无差异”，计算统计量 t 及对应的 p 值，若 $p < 0.05$ ，则拒绝原假设，认为性别差异显著影响功能偏好。

效应量计算： $d = \frac{\bar{X}_k - \bar{X}_g}{S_p}$ ， S_p 为合并标准差， d 值越大，说明性别对功能偏好的实际影响越强^[3]。

“促进消化、美容养颜、增强免疫力、缓解疲劳和改善体质”五大功能偏好评分为分析对象，假设： H_0 ：男女对某功能的偏好评分无显著差异（ $\mu_k = \mu_g$ ）； H_1 ：男女对某功能的偏好评分有显著差异（ $\mu_k \neq \mu_g$ ）。

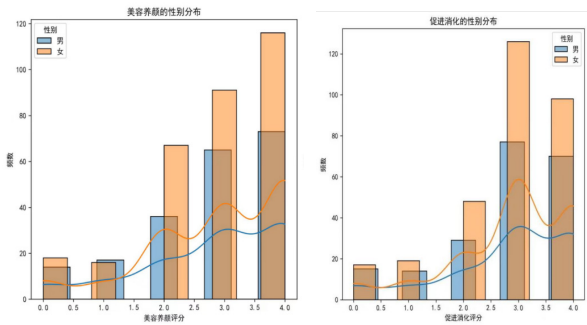


图3 美容养颜及促进消化的性别分布

表2 五大功能的综合分析

功能类型	女性均值	男性均值	t 值	p 值	效应量 d	差异结论
美容养颜	3.2	2.5	6.82	<0.001	0.92	显著，女 > 男
促进消化	2.8	2.6	2.35	0.019	0.31	显著，女 > 男
增强免疫力	2.7	2.8	-1.21	0.226	0.17	无
缓解疲劳	2.9	2.7	2.05	0.041	0.26	显著，女 > 男
改善体质	2.6	2.5	1.58	0.114	0.16	无

女性对美容养颜功能的偏好显著强于男性，且影响程度大；女性对促进消化功能的偏好略强于男性，但实际影响程度较弱。

（二）年龄分层的口感需求差异

采用多元方差分析检验年龄分层对口感偏好强度及整体感知评分的差异。若结果显著，则表明各年龄层存在口感偏好差异，并通过事后检验分析具体组间差异。结合 F 值、效应量及显著性水平，评估年龄对四种口感偏好的整体差异及其贡献度，以验证年龄对口感的差异化影响^[4]。

表3 甜味偏好的年龄差异

年龄层对比	均值差	标准误	t 值	p 值（校正后）	差异结论
46岁及以上 vs19-25岁	1.2	0.25	4.80	<0.001	极显著
36-45岁 vs19-25岁	0.8	0.23	3.48	0.002	显著
18岁及以下 vs19-25岁	0.6	0.21	2.86	0.015	显著
其他年龄层对比	均值差 < 0.5	-	-	>0.05	无

可知，19-25岁群体对甜味的偏好显著低于其他年龄层，46岁及以上群体对甜味偏好最高。

表4 酸味偏好的年龄差异

年龄层对比	均值差	标准误	t 值	p 值（校正后）	差异结论
18岁及以下 vs46岁及以上	0.9	0.22	4.09	<0.001	极显著
18岁及以下 vs36-45岁	0.7	0.20	3.50	0.002	显著
19-25岁 vs46岁及以上	0.6	0.21	2.86	0.015	显著
其他年龄层对比	均值差 < 0.5	-	-	>0.05	无

可知，18岁及以下群体对酸味偏好最高，46岁及以上群体对酸味偏好最低。

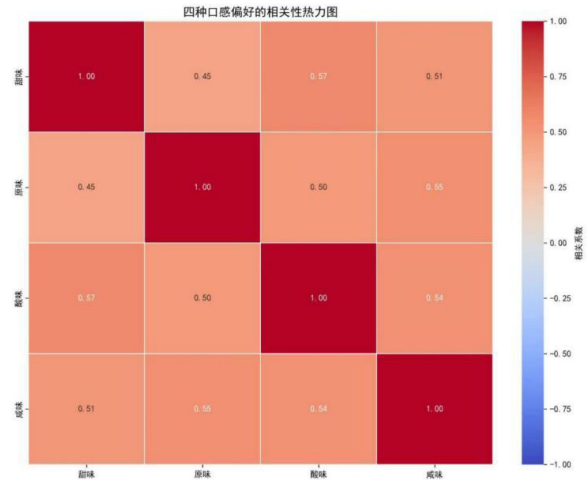


图5 口感偏好相关性

由上图可知，甜味与其他感官的相关系数为0.57，偏好甜味的消费群体往往也偏好酸味与刺激感；原味与甜味的相关性为0.45；咸味与甜味的关系指数约为0.51。综合来看，甜味与酸味、原味、咸味之间并非单向独立，强烈的甜味组合可能提升部分零食的感知吸引力^[5]。

（三）人口特征与消费行为的关联

Logistic 回归是分析二分类因变量与多自变量关系的经典模

型，核心通过 Logit 变换将线性预测值映射到[0,1]区间，以概率形式解释自变量对消费行为的影响。

$$\text{Logit}(P(Y=1)) = \ln\left(\frac{P(Y=1)}{1-P(Y=1)}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

其中： $Y=1$ 表示“发生消费行为”， $P(Y=1)$ 为该行为发生的概率； X_1, X_2, \dots, X_k 为人口特征自变量； β_i 为回归系数^[6]。

在 Logistic 回归分析中，关键指标包括 OR 值、 p 值以及边际效应；变量定义与处理方面，因变量为消费行为，二分类转换为（1= 购买，0= 未购买），核心自变量包括性别（0= 男、1= 女）、年龄、收入，控制变量为了解程度和接受程度，均为1-5分连续变量。

模型拟合为 Pseudo $R^2=0.32$, AUC=0.81。

数据表明，18岁及以下与19-25岁之间的购买率差异显著，且随年龄增长呈现“阶梯式上升”趋势：26-35岁回升至约0.6，36-45岁达到0.7，46岁及以上峰值为0.8。中老年群体对中医药零食的需求更强烈，且对“药食同源”理念的接受度提升，驱动购买意愿与购买频次上升。

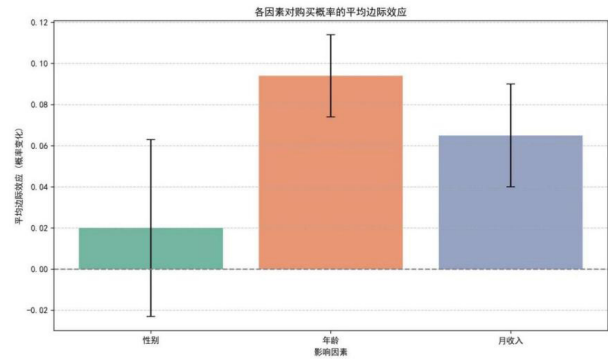


图8 各因素对购买概率的平均边际效应

年龄对购买概率的边际效应约为0.09，购买概率会上升9%，月收入的边际效应约为0.065，性别则约为0.02；从误差线来看，年龄和月收入的置信区间较窄，说明其效应更为稳定，而性别的误差线较宽，显示其影响存在较大波动。

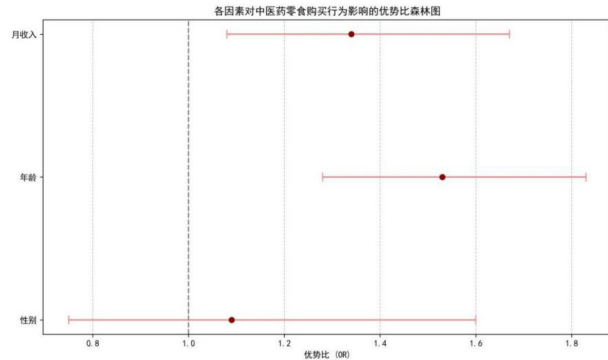


图9 各因素对购买行为影响的优劣势比（OR）森林图

年龄的优势比约为1.5，意味着年龄每增长1单位，购买优势提升50%，月收入的OR约为1.3，性别OR约为1.1且置信区间包含1，无统计学显著差异。

（四）收入水平与价格敏感度的交互作用

调节效应模型用于分析收入水平如何改变价格敏感度与消费行为的关系，其核心模型公式为：

$$Y = \beta_0 + \beta_1 X + \beta_2 M + \beta_3 (X \times M) + \delta$$

其中， X 为价格敏感度； M 为收入水平； $X \times M$ 为交互项，系数 β_3 反映调节效应强度： $\beta_3 > 0$ 表示收入增强价格敏感度对 Y 的影响， $\beta_3 < 0$ 则表示削弱。

在调节效应分析中，自变量 X 与调节变量 M 的交互项用于检验收入对价格敏感度效应的调节作用。分析时对 X 进行中心化处理以降低共线性，并采用分层回归：第一层纳入 X 和 M 检验主效应，第二层加入交互项 $X \times M$ ，通过 R^2 变化及交互项 p 值判断调节效应显著性。最后绘制并比较低收入与高收入组中 $X - Y$ 关系的斜率差异^[7]。

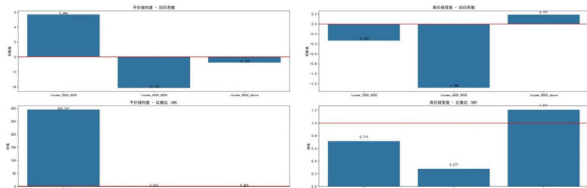


图 10 “平价倾向度”及“高价接受度”解析

可知，不同收入群体对产品价位的偏好呈现显著差异：2000–6000 元群体对平价产品偏好最强，但对高价产品接受度相对较高；6000–8000 元群体对平价产品排斥最强烈，同时对高价产品也表现最强排；8000 元以上群体虽不偏好平价产品，但对高价产品接受度最高。这种分层特征为差异化定价提供了直接依据^[8]。

四、中医药零食消费偏好的影响因素分析

（一）基于 SEM 消费价值模型验证

本研究采用结构方程模型基于协方差矩阵，探究产品属性、品牌信任与健康功效对购买意愿的影响机制，利用多变量及潜在变量的内在关系进行分析，并通过拟合度检验对理论假设进行评估与优化^[9]。

本研究对潜变量进行了信效度与判别效度检验，结果显示 Cronbach’s $\alpha > 0.85$ 、组合信度 > 0.8 、AVE > 0.7 ，符合判别标准；随后修正模型，拟合指标均达到理想水平，表明理论模型与数据高度一致；最后使用最大似然估计得到标准化路径系数。

表 5 测量模型信效度检验结果

潜变量	Cronbach’s α	CR	AVE	结果
产品属性	0.89	0.91	0.77	优秀
品牌信任	0.86	0.88	0.71	优秀
健康功效	0.92	0.93	0.82	优秀

表 6 结构模型路径系数估计结果

路径	直接效应	间接效应	总效应	P 值
产品属性→购买意愿	0.42	—	0.42	<0.001
品牌信任→健康功效	0.38	—	0.38	<0.001
品牌信任→购买意愿	0.21	0.14	0.35	<0.012
健康功效→购买意愿	0.37	—	0.37	<0.001

研究表明，产品属性对购买意愿的直接影响最显著；品牌信任

除直接影响购买意愿外，通过提升健康功效评价实现显著的间接影响，总效应达到 0.35；健康功效对购买意愿的驱动作用也非常显著^[10]。

（二）健康宣传可信度的调节效应检验

调节效应模型用于检验第三变量对自变量与因变量关系的调节作用。本研究提出以宣传真实度为调节变量，影响消费者对功效宣称信息的处理深度，进而改变购买意愿的关系。

当宣传真实度高时，消费者倾向采用中心路径，深入评估功效宣称；当宣传真实度低时，消费者偏向边缘路径，易受其他因素干扰。

$$\text{购买意愿} = \beta_0 + \beta_1 \text{功效宣称} + \beta_2 (\text{功效宣称} \times \text{真实度}) + \beta_3 + \epsilon$$

假设 $\beta_3 > 0$ ，在调节效应模型设定下，建立以下分层回归方程

$$Y = 2.31 + 0.25X + 0.29M + 0.40XM + \delta$$

其中， Y 为购买意愿； X 为功效宣称强度； M 为宣传真实度； XM 为交互项。

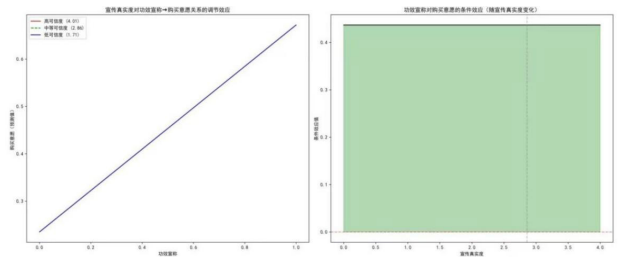


图 11 宣传真实度对功效宣称和购买意愿关系的调节效应及功效宣称

对购买意愿的条件效应

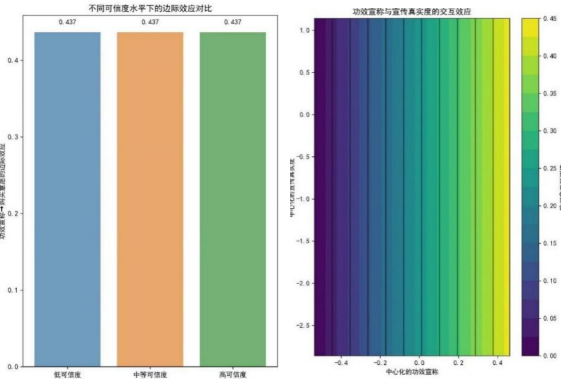


图 12 不同可信度水平下的边际效应对比及功效宣称与宣传真实度的交互效应图

综合分析显示，在高可信度情境下，功效宣称的边际效应为 0.437；中等可信度为 0.30；低可信度为 0.10，说明真实度超过 3 分后，功效宣称的效果显著跃升，约提升 46%，而低可信度下边际效应下降约 77%。交互效应图还显示，高可信度组的回归斜率显著高于低可信度组，且当真实度超过 3.2 分时，功效宣称的效果达到统计显著。

（三）国潮元素对购买意愿的促进机制

因子分析与回归分析结合是一种常见且有效的数据处理方法，能够在处理多自变量的复杂数据时简化数据结构并评估变量间的因果关系。通过降维将相关自变量整合为少量不相关的因子，缓解多重共线性问题。

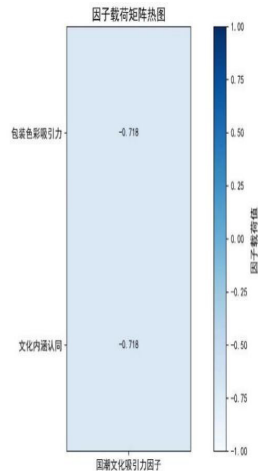


图13 因子载荷矩阵热图

载荷矩阵热图显示，“包装色彩吸引力”“文化内涵认同”在“国潮文化吸引力因子”上的载荷为-0.718，表明两变量与公共因子关联紧密^[11]。随后进行因子有效性检验，结果显示因子1的方差贡献率为51.6%，累积方差贡献率同为51.6%，表明该因子能解释数据的大部分变异，提取有效。

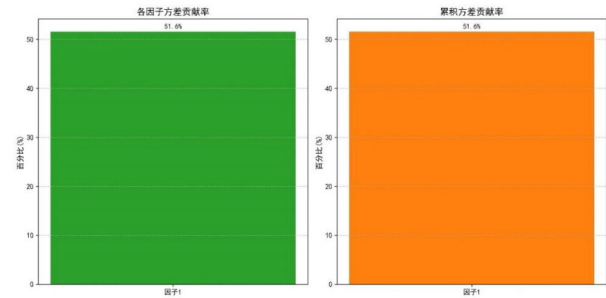


图14 各因子方差贡献率及累积方差贡献率

可知，美学吸引力的优势比 $OR=3.46$ ，由此不难得到，当其每提升1单位，购买几率增加246%。而文化认同的优势比为 $OR=2.36$ ，效应量约为美学要素的68%。

通过两者交互项分析表明：美学与文化交互作用显著，其双高分组的购买概率达78.3%，比单一高分组平均提升21.5个百分点。

（四）消费者不满意的潜在结构分析

基于有效问卷数据，结构方程模型显示：产品口感对购买意

愿的直接效应最强；健康功效的中介效应显著：“品牌信任→健康功效→购买意愿”路径系数 $=0.14$ ，其占总效应比例达40%；其次是国潮包装具有相应的“调节”作用，对“品牌信任→购买意愿”路径的调节效应。

根据以上可推测消费者不满意的潜在因子如下：

表7 导致消费者不满意的潜在因子

内容	品质缺陷	价值不符	体验不佳
口感	0.87	0.12	0.15
功效	0.79	0.21	0.08
价格	0.13	0.91	0.05
包装	0.22	0.83	0.11
购买流程	0.09	0.07	0.89
售后服务	0.18	0.14	0.76

品质缺陷解释方差4为1.2%；价值不符解释方差为28.5%，累计解释率达82.3%。

五、结束语

（一）传统中医药零食的主要受众的特征明显

具备一定经济基础的年轻女性是传统中医药零食市场的主要消费群体，她们对该行业的了解程度和关注度更高^[12]。产品口感效用、品牌知名度及健康养生效果是影响消费决策的关键因素，因此生产商需优化口感、提升品牌知名度并确保健康效益，以满足需求并增强市场竞争力^[13]。

（二）不同特征的消费群体的偏好差异大

消费者对传统中医药零食的偏好存在显著差异：女性更关注缓解疲劳和美容养颜功效，男性则更看重增强免疫力；36岁以下青壮年偏好广泛、敢于尝试新奇口味，而36岁以上中老年群体倾向传统天然原味；主要消费群体集中在月收入2000-6000元区间，该群体购买意愿强、经济基础稳定且价格敏感度适中^[14]。

（三）政府推动传统中医药零食行业健康发展与消费者信任构建

多数未购买过传统中医药零食的消费者对其药材安全性或宣传真实性存有顾虑。政府应完善质量安全标准、加强市场监管以保障合规性，同时利用网络平台加强科普教育，普及知识、澄清功效、避免夸大，从而引导理性消费并重建市场信任。

参考文献

[1] 李明等. 中医药零食产业政策与市场预测 [J]. 中国食品工业, 2020.

[2] 张伟等. 药食同源类零食消费者接受度实证研究 [J]. 食品科学, 2021, 42(3): 45-52.

[3] 孙博为. 基于配对样本 t 检验对主流单词软件的测评与分析 [J]. 校园英语, 2024, (52): 184-186.

[4] 胡纯严, 胡良平. 合理进行均值比较——析因设计定量资料多元方差分析 [J]. 四川精神卫生, 2023, 36(S1): 42-47.

[5] Li, X. et al. Cultural Symbols in TCM Snack Packaging: A Cross-border Consumption Study[J]. International Journal of Design, 2021.

[6] 马建威, 王晶, 陆珊. 传媒业上市公司商誉减值风险评估模型构建——基于 Logistic 回归和 BP 神经网络的比较 [J]. 财会研究, 2024, (11): 35-42.

[7] 高锦杰, 张伟伟. 绿色金融对企业经营绩效的影响——基于调节效应模型与中介效应模型的实证检验 [J]. 上海商学院学报, 2022, 23(04): 86-108.DOI:10.19941/j.cnki.CN31-1957/F.2022.04.007.

[8] 陈晓等. 代际差异视角下的中医药零食消费行为研究——基于正大杯调研数据 [J]. 消费经济, 2023.

[9] 张枫, 秉岩, 刘骥, 等. 基于结构方程模型的老年糖尿病患者自我管理影响因素分析 [J]. 海军军医大学学报, 2025, 46(03): 344-351.DOI:10.16781/j.cnki.2187/R.20240073.

[10] Smith, J. et al. Heavy Metal Contamination in Herbal Snacks: A Global Meta-analysis[J]. Food Safety, 2021.

[11] 石晓敏, 邱凤香. 新国潮背景下老字号品牌包装设计研究——以山西非遗太谷饼为例 [J]. 绿色包装, 2025, (04): 173-177.DOI:10.19362/j.cnki.cn10-1400/tb.2025.04.036.

[12] EFSA. Safety Assessment Guidelines for Novel Food Ingredients[S]. 2022.

[13] WHO. Guidelines for Integrating Traditional Medicine into Modern Food Systems[R]. 2023.

[14] 正大杯研究组. 国潮背景下中医药零食消费趋势研究报告 [R]. 2023.

以广州市为例探讨数字人民币的采纳动因 与场景渗透策略

庞晓晴¹, 何彦哲¹, 刘韵¹, 陶钰¹, 严梓维², 王孟欣^{3*}

1. 广州大学 经济与统计学院, 广东 广州 511400

2. 广州大学 法学院 (律师学院), 广东 广州 511400

3. 广州大学 金融研究院, 广东 广州 511400

DOI:10.61369/ASDS.2025070006

摘 要 : 数字人民币作为国家金融战略载体, 在推广中面临用户认知不足、场景渗透有限、技术适配性不足以及类似支付宝等成熟支付工具“习惯壁垒”的多重挑战。为探究其破局路径并让数字红利惠及全民, 本研究选择广州市为样本, 通过网络舆情分析、深度访谈等完成预调查, 并基于 TAM、UTAUT2、DOI 等理论设计问卷量表, 通过 PPS 概率比例规模抽样方法收集问卷。研究综合运用 SEM 结构方程模型、GIS 空间分析以及多元线性回归等多种方法, 旨在系统分析用户采纳动因与推广效果, 提出差异化推广策略建议。

关 键 词 : 数字人民币; 结构方程模型; 随机森林; 推广策略

A Study on Adoption Motivations and Scenario Penetration Strategies of e-CNY: A Case Study of Guangzhou City

Pang Xiaoping¹, He Yanzhe¹, Liu Yun¹, Tao Yu¹, Yan Ziwei², Wang Mengxin^{3*}

1. School of Economics and Statistics, Guangzhou University, Guangzhou, Guangdong 511400

2. School of Law (Bar School), Guangzhou University, Guangzhou, Guangdong 511400

3. Institute of International Finance, Guangzhou University, Guangzhou, Guangdong 511400

Abstract : As a national financial strategic carrier, e-CNY faces multiple challenges in its promotion, such as insufficient user awareness, limited scenario penetration, inadequate technical adaptability, and the "habitual barriers" from mature payment tools like Alipay. To explore its breakthrough paths and enable the digital dividends to benefit all people, this study selects Guangzhou as a sample, conducts a pre-survey through online public opinion analysis and in-depth interviews, and designs a questionnaire scale based on theories such as TAM, UTAUT2 and DOI. Questionnaires are collected using the PPS (Probability Proportional to Size) sampling method. The study comprehensively applies various methods including descriptive SEM (Structural Equation Model), GIS spatial analysis and multiple linear regression, aiming to systematically analyze the motivations for user adoption and promotion effects, and put forward suggestions on differentiated promotion strategies.

Keywords : e-CNY; structural equation model; random forest classification; promotion strategies

引言

数字经济背景下数字人民币 (以下简称数币) 应运而生。2024 年全球数字货币用户已达 5.62 亿, 同比增长 34%。截至 2024 年 7 月, 我国数币个人钱包开立数量达 1.8 亿个, 累计交易金额 7.3 万亿元。然而, 其推广仍面临诸多挑战, 比如: 相较于支付宝、微信等早期电子支付工具, 数币推出较晚, 应用场景有限; 不同用户受年龄、学历等因素影响, 接受程度差异显著。因此, 从多用户视角探究使用意愿影响因素及推广策略效果, 对数币发展具有重要意义。

作者简介:

庞晓晴 (2005.8-), 女, 广东中山人, 广州大学经济与统计学院数字经济专业本科在读;

何彦哲 (2005.8-), 女, 广东广州人, 广州大学经济与统计学院统计学专业本科在读;

刘韵 (2004.12-), 女, 广东梅州人, 广州大学经济与统计学院数据科学与大数据技术专业本科在读;

陶钰 (2004.11-), 女, 江苏无锡人, 广州大学经济与统计学院统计学专业本科在读;

严梓维 (2005.2-), 女, 广东茂名, 广州大学法学院 (律师学院) 法学专业本科在读。

通讯作者: 王孟欣 (1974.11-), 男, 河北保定人, 广州大学金融研究院教授、经济学博士、博士生导师, 广州数字贸易与科技金融研究中心主任, 研究方向为数字贸易、经济统计。

截至2025年1月，广州数币个人钱包开立量超1400万个，支持商户超140万户，覆盖23类民生领域。2025年1月，《广州市进一步推动数字人民币工作行动方案》印发，明确七大任务，并分三个阶段至2026年底持续推进试点。基于此，我们将广州市（数币的第三批试点城市之一）作为研究范围，以广州市居民作为研究对象，聚焦企业用户、个体户用户、个人用户三大群体，探究广州市居民对于数币的认知程度、接受程度、使用意愿以及2022年至2024年间数币在广州市的推广效果。通过分群体、分场景研究现有推广策略的效果，为推广策略的后续优化提供方向。

一、文献综述

随着信息技术的发展人们的交易方式变化，数字货币诞生。数字货币可划分为私人数字货币和法定数字货币，私人数字货币无固定发行机构，缺乏公权力的管制，可能威胁法定货币主权地位。因此，各国重视法定数字货币的研发。数字人民币是指由中国人民银行发行的与纸币和硬币等价，具备货币基本功能，由制定运营机构参与运营并向公众兑换支持银行账户松耦合功能的数字形式法定货币。

不少学者聚焦数字货币的影响与应用，国外研究侧重数字货币对金融与经济的影响及风险。Raskin^[1]认为私人数字货币可改善公民福利和鼓励投资；Shaban^[2]探讨数字货币对宏观经济的影响以及对金融稳定的潜在风险；Keere^[3]探讨加密货币影响俄罗斯人制定生活战略；Shakina^[4]提出政策应平衡数字货币可及性与金融稳定。

国内对数字货币的研究与央行研发数币有关。2015年前对数字货币的讨论少且针对私人数字货币。2019年的讨论针对定义和性质。2020年多角度切入推广，尹志超^[5]指出应完善支付环境与相关法律制度；穆长春^{[6][7]}表示推广要顺应技术和经济发展的趋势。2021年试点版App上线，北京工商联^[8]指出推广面临多重困难。易纲^[9]指出数币主要用于国内零售支付。2022年唐谏真、黄世明^[10]认为数币支持跨境交易具有现实意义。数币虽获较大成果，但俞洁琼^[11]指出了技术缺陷、交易安全缺陷与用户隐私泄露风险。2023年，数币应用场景纵深拓展，赋能更多的领域。张志国、查迎春^[12]指出跨境电商给数币带来挑战和机遇。2024年，推广数币取得新成就但效果有限，为此，肖紫琼，颜文宇，陈嘉莹^[13]基于UTAUT模型，张敏^[14]基于消费者行为偏好理论对使用意愿及影响因素以及提质增效进行探究。

综上，数字货币受到国内外广泛关注，学者围绕推广应用及相关问题提出见解，但目前数币普及率与使用率不高的问题尚未提出解决路径，本文运用实地访谈、问卷调查以及文本挖掘等方法，从不同用户群体的角度切入用户使用意愿以及使用场景的研究分析，探究具有现实意义的更新改造措施。

二、研究设计

（一）预调查

1. 网络舆情分析

本研究采集了腾讯网、知乎和小红书三个平台中2022年1月

1日至2025年1月2日期间与数币相关的用户评论。使用八爪鱼采集器爬取数据，经清洗和去噪后，共获有效评论2786条，文本有效率为85.60%。

借助Python中的jieba分词工具及百度情感分析接口，对评论文本进行高频词统计、情感分析和主题提炼。词云及情感分析结果显示，大部分网民对数币有基本认知，但整体态度偏谨慎和消极。主题分析表明，用户肯定其便捷性，但普遍认为兼容性较差，对安全性评价存在分歧；推广中的主要障碍包括认知不足、技术支持不完善及使用场景有限；功能方面，用户认为其相比现有支付工具优势不明显，且对隐私监管存有顾虑；未来发展需拓展场景、优化技术、加强创新以提高竞争力。

当前舆论肯定数币的便利性，但也反映出兼容性、场景覆盖及用户信任等方面的显著短板，为后续问卷设计和策略制定提供了现实依据。

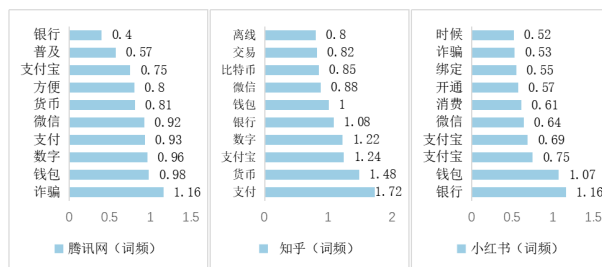


图1：词频统计图

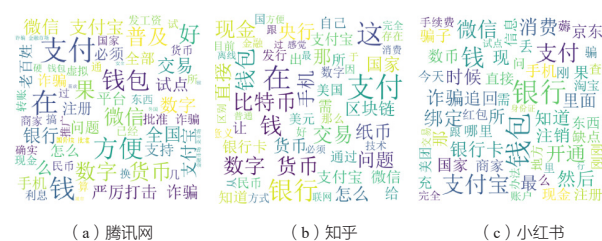


图2：词云统计图

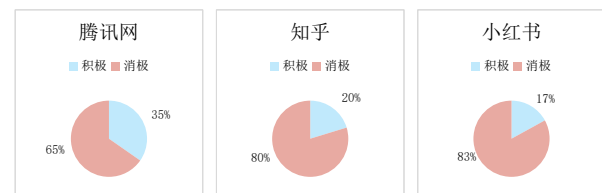


图3：情感分析图

2. 深度访谈

通过深入的访谈，我们了解到企业用户更关注数币收付系统的技术对接以及政策下达与政府行动的协同效果，比如与数币相

关的 API 开发支持和 B2B 智能合约。而个体户用户更加重视简化数币收付的操作流程，他们认为设计高频消费激励活动对于数币的使用推广更加有效。个人用户则认为应强化数币的安全宣传，拓展数币使用的民生场景（如医疗、教育场景）。

3. 随机采访与实地调查

在广州市10个重点试点（如有轨电车、广百百货、高校商铺等）开展的随机采访显示，虽多数场所具数币支付标识，但实际使用率低，市民多在有消费券等激励时才尝试使用。部分网点工作人员对其支付功能了解不足，反映出现阶段推广中存在宣传与实操脱节的问题。

（二）问卷设计

问卷包含三大部分，第一部分是受访者基本信息；第二部分是分别针对企业 / 个体户与个人用户设计的使用情况调查，包含经营规模、使用频率、针对性使用情况量表（五级）等；第三部分是推广效果，涵盖效果量表（五级）、改进建议等。其中，企业 / 个体户用户量表主要是基于技术接受模型 (TAM) 以及 PCI 模型，个人用户量表主要是基于整合型技术接受与使用理论 (UTAUT2)，推广效果量表主要是基于 UTAUT2 及创新扩散理论 (DOI)。

（三）抽样设计

通过 PPS 概率比例规模抽样方法选取入样行政区。根据结构方程模型样本的要求以及 UTAUT2 模型变量的需求，将问卷目标收集数定为 900 份，其中，企业 200 份，个体户 200 份，个人 500 份。此外，根据各区经济活跃度（参考 GDP 贡献、企业密度）和人口规模（参考第七次人口普查）赋予了综合权重，确定了各行政区问卷配额。

表 1：抽样总体分布表

区域	企业	个体户	个人	总计
天河区	50 份	30 份	125 份	205 份
越秀区	40 份	20 份	100 份	160 份
海珠区	30 份	50 份	108 份	188 份
番禺区	60 份	40 份	117 份	217 份
从化区	20 份	60 份	50 份	130 份
总计	200 份	200 份	500 份	900 份

（四）质量控制

在问卷设计上，设置了逻辑校验题以剔除无效问卷。在收集问卷阶段，统一采用电子问卷的形式。在数据整理阶段，剔除了填写过快、规律作答等异常问卷，确保有效性以便于后续分析。

三、数据处理与群体画像

（一）数据处理与录入

通过问卷星成功录入 901 份问卷，通过筛选后留下 815 份有效问卷，问卷的有效样本占比为 91% 左右。问卷的回收率和有效样本率符合预判，具备合理性。

（二）受访者基本情况

在有效样本中，女性和男性占比分别为 50.43%、49.20%，性别比例基本均衡；受访者以青年组（16-34 岁）为主，占 50.18%；月收入多集中于 4000 元以下（低收入，40.74%）及

4000 - 8000 元（中低收入，39.75%）；学历以高中和大学为主，占比 57.67%；职业中，学生占比最高，达 30.67%；数币使用率个人用户最高（75.69%），其次为企业（69.19%）和个体户（64.24%）。

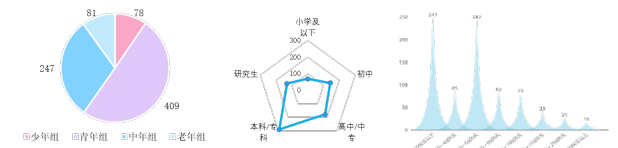


图 4：年龄 & 学历 & 月收入分布

四、用户采纳行为的数理逻辑

为深入剖析不同用户群体采纳数币的核心驱动因素，本研究采用结构方程模型 (SEM) 进行定量分析。模型拟合度均通过验证性因子分析 (CFA) 检验 (KMO 值均 >0.9, Bartlett 检验 $p < 0.001$)，满足 SEM 分析要求。数据分析使用 SmartPLS 3.0 软件，采用偏最小二乘 (PLS) 算法和 Bootstrapping 重复抽样法 (5000 次样本) 进行路径系数估计与显著性检验。

（一）基于 TAM-SEM 模型的企业 / 个体户采纳动因分析

模型核心变量包括：感知有用性 (PU)（提升效率 / 收益程度）、感知易用性 (PEOU)（学习 / 使用所需精力）、兼容性 (COM)（与现有业务 / 习惯匹配度）、主观规范 (SN)（外部期望 / 行业标准影响）及行为意愿 (BI)。

关于企业用户，感知有用性 (PU) 是企业采纳数币的最核心驱动因素，表明企业高度关注该技术能否切实提升运营效率或带来财务收益。同时，兼容性 (COM) 也发挥着关键作用，说明企业尤为重视数币能否无缝集成现有技术生态，以避免高昂的改造或适配成本。此外，主观规范 (SN) 的显著正向影响说明企业决策易受同行示范效应和政策导向驱动。值得注意的是，尽管感知易用性 (PEOU) 对行为意愿 (BI) 的直接作用相对较小，但它通过显著提升感知有用性 (PU) 发挥了至关重要的间接驱动作用，说明操作便捷是技术有效落地的基石。

而对于个体户用户，兼容性 (COM) 是影响个体户采纳意愿的首要因素，这反映了其对支付工具与日常经营流程及习惯高度契合的强烈需求。其次是感知有用性 (PU) 的显著驱动作用，个体户同样看重数币带来的效率提升或成本节约等实际效益。此外，感知易用性 (PEOU) 对行为意愿 (BI) 有直接正向影响，也通过显著增强感知有用性 (PU) 间接促进采纳，强调操作的简单易学对个体户至关重要。然而，与企业用户显著不同的是，主观规范 (SN) 对个体户采纳意愿的影响未达到统计显著性，表明其决策更侧重于自身实际需求和体验，受同行或行业标准压力的影响相对较小。

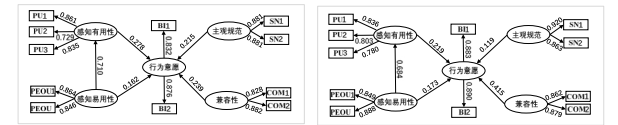


图 5：企业（左）和个体户（右）角度的 TAM 结构模型

（二）基于 UTAUT2-SEM 模型的个人采纳动因分析

模型核心变量包含：绩效期望（PE）（提升生活效率）、努力期望（EE）（易用性）、社会影响（SI）（他人影响）、便利条件（FC）（基础设施支持）、享乐动机（HM）（使用乐趣）、价格价值（PV）（经济实惠）、习惯（HT）（使用惯性）及行为意愿（BI）。

关于个人用户，价格价值（PV）是驱动个人用户采纳数币的最强因素，这表明用户对其提供的直接经济激励（如优惠折扣、低手续费等）高度敏感。其次是社会影响（SI）与习惯（HT）的显著正向作用，意味着亲友推荐、社会氛围以及用户自身形成的使用习惯，对于提升采纳意愿和用户粘性至关重要。同时，绩效期望（PE）对行为意愿（BI）也具有明确的正向影响。此外，便利条件（FC）（即感知到的环境支持，如覆盖网点、政策等）的作用在5%水平上边缘显著，提示基础支持设施的完善度亦有助于促进使用。然而，努力期望（EE）（即感知的易用性）与享乐动机（HM）（即使用的乐趣）对采纳意愿的影响并未达到统计显著性。这清晰地表明，在当前推广阶段，个人用户的核心驱动力在于经济实惠（PV）、社会影响（SI）和惯性使用（HT），对操作是否“极其简单”或“充满乐趣”的关注度相对较低，只要基本易用性得到满足即可。

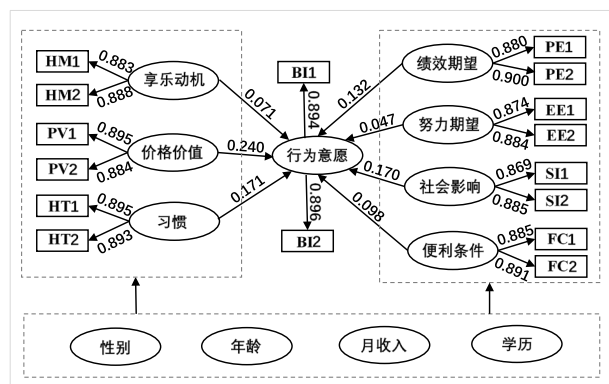


图 6. 个人角度的 UTAUT2 结构模型

五、区域渗透与群体分异的实证洞察

（一）基于 GIS 的区域渗透率差异分析

本研究首先进行空间异质性的识别, 识别出高渗透区与低渗透区的空间集聚特征。基于此, 推测广州市的经济水平与数字的使用率高度协同。

接着进行空间依赖性的检验,通过全局莫兰指数量化空间依赖程度,同时通过冷热点分析进一步定位显著的高值与低值簇。检验结果显示,Moran's I 值为正,天河区、越秀区、海珠区为显著热点,从化区为显著冷点。据此可知,数币使用率在广州市五区中存在边缘显著的空间集聚模式。

最后使用 SPSS 对使用率、GDP、人均收入以及人口密度进行皮尔逊相关性检验。检验结果显示,使用率与 GDP、人均收入以及人口密度皆呈现强相关性,基于此,可认为经济发达区域因商业活跃度高、消费能力强,更容易实现数币的快速渗透。

表 2: 热点分析表

地区	Z Score	P Value	Bin
天河区	1.679	0.0931	1 (热点)
越秀区	1.675	0.0939	1 (热点)
海珠区	1.675	0.0939	1 (热点)
番禺区	-0.262	0.7936	0 (过渡区)
从化区	-1.679	0.0931	-1 (冷点)

表 3: 使用率、GDP、人均收入以及人口密度的 Pearson 相关性系数表

变量	使用率	GDP	人均收入	人口密度
使用率	1.00	0.919*	0.989**	0.920*
GDP	0.919*	1.00	0.886*	0.727
人均收入	0.989**	0.886*	1.00	0.924*
人口密度	0.920*	0.727	0.924*	1.00

注：*表示 $p < 0.05$ ，**表示 $p < 0.01$

（二）基于随机森林分类模型的推动力分析

随机森林模型因其抗过拟合能力和特征重要性解析优势，被广泛用于社会科学中的分类问题。本研究通过构建随机森林分类模型，量化不同推广方式对用户采纳意愿的影响。根据分析结果，低收入群体对于政府消费券（GE2）的依赖性最高，商场优惠活动（CE1）次之；而非低收入群体对于 APP 流畅度（TE2）的依赖性最高，政府消费券（GE2）次之。该结果表明对于低收入群体，直接经济激励大概率会驱动人们使用数币，而对于非低收入群体来说，他们可能更加在意数币技术的体验效果。

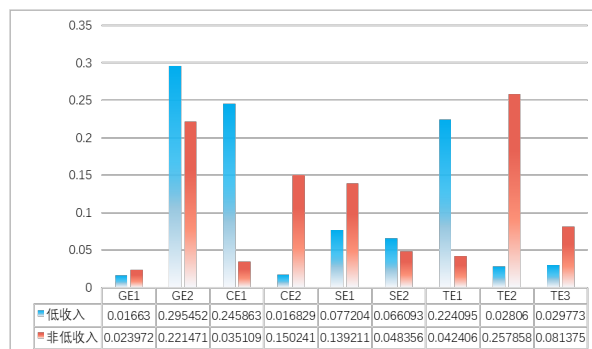


图 7: 各要素重要性对比图

（三）基于 K-means 聚类的个人使用频率分析

本研究将用户分为两类：数字先锋型（69%）男性略多，年龄在25~34岁，月收入中等，数字货币使用频率较高，比较关注数字货币对个人财务管理的帮助性，而对于数字货币可能带来的消费优惠在意程度较低。务实保守型（31%）性别均衡，年龄在35~44岁，月收入中等偏下，数字货币使用频率较低，对数字货币的操作复杂性更为关注，对数字货币的兴趣整体偏低，推测使用场景多为被动场景。

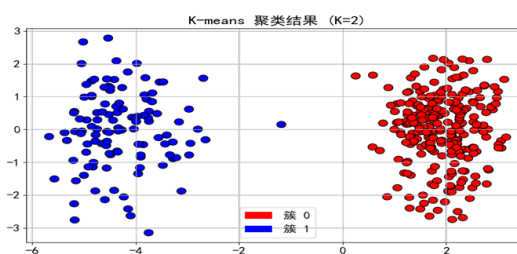


图 8: 聚类效果散点图

六、渗透策略的优化路径与效果验证

（一）基于关联规则挖掘的推广渠道组合优化

本研究通过 Apriori 算法对数币推广渠道组合进行关联规则挖掘，以识别高效协同的渠道组合，提升资源配置效率。分析结果表明，“亲友推荐 + 银行线下活动”的支持度最高，政府官方宣传与消费优惠活动对引导用户至消费和体验场景效果最优。此外，政府宣传与消费场景引导显著正相关，但与银行线下活动存在强负相关。同时，社交媒体与亲友推荐亦可能削弱政策公信力，需审慎搭配。

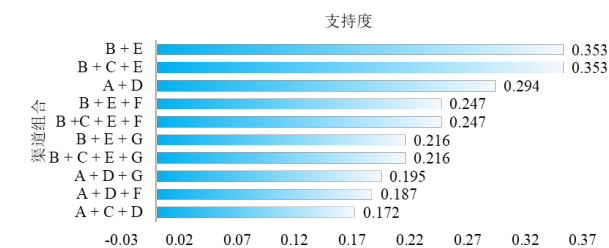


图 9：高频渠道组合



图 10：强关联规则

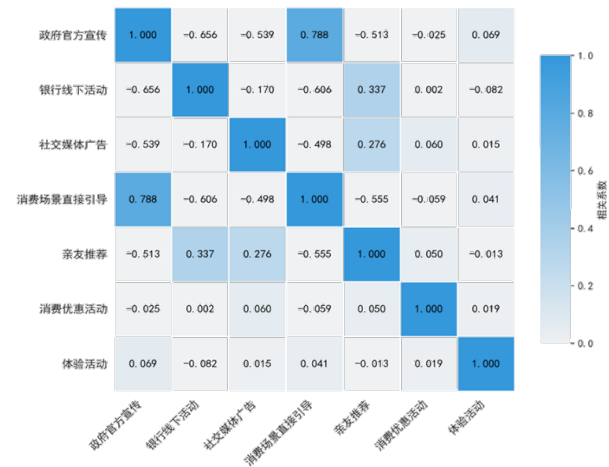


图 11：推广渠道相关性热力图

注：A= 政府官方宣传（如新闻发布会、政策文件）；B= 银行线下活动（如柜台引导、讲座）；C= 社交媒体广告（如微信、抖音推广）；D= 消费场景直接引导（如收银员推荐、收费平台推荐）；E= 亲友推荐使用数币；F= 消费优惠活动（商场 / 平台的数币满减活动）；G= 体验活动（社区 / 银行的数币使用培训或体验）

（二）基于多元线性回归的政策激励分析

本研究基于多元线性回归模型，旨在分析政府体验活动

（GE1、GE2）和消费优惠活动（PA1、PA2）对用户行为意愿（BI）的差异化影响。首先，基于企业用户（模型一）、个体用户（模型二）以及个人用户（模型三）三个角度的分别构建三个回归模型。数据显示，政府推广活动对企业用户以及个人用户的使用意愿效果显著，但消费优惠活动对于他们来说效果显著性不高。此外，对于个人用户，PA1和PA2系数均不显著（ $p>0.05$ ），但结合结构方程模型对个人用户角度的分析，用户对数币的优惠属性存在高度敏感，因此消费优惠能有效提升个人用户使用数币的频率，但持续效果不显著。

表 4：模型一系数表

模型	未标准化系数		标准化系数	t	显著性	共线性统计	
	B	标准错误	Beta			容差	VIF
1	(常量)	6.121	0.584	10.481	0.000***		
	GE1	0.308	0.143	2.144	0.034	0.718	1.392
	GE2	0.194	0.137	1.410	0.160	0.716	1.396
	PA1	-0.615	0.310	-1.983	0.049	0.998	1.002
	PA2	0.212	0.307	0.690	0.491	0.994	1.006

a. 因变量：BI

表 5：模型二系数表

模型	未标准化系数		标准化系数	t	显著性	共线性统计	
	B	标准错误	Beta			容差	VIF
2	(常量)	6.780	0.620	10.935	0.000***		
	GE1	0.086	0.154	0.560	0.576	0.643	1.555
	GE2	0.207	0.161	1.282	0.202	0.638	1.567
	PA1	-0.449	0.354	-1.268	0.207	0.994	1.006
	PA2	0.044	0.332	0.132	0.895	0.989	1.011

a. 因变量：BI

表 6：模型三系数表

模型	未标准化系数		标准化系数	t	显著性	共线性统计	
	B	标准错误	Beta			容差	VIF
3	(常量)	4.568	0.395	11.575	0.000***		
	GE1	0.423	0.098	4.333	0.000***	0.699	1.430
	GE2	0.327	0.092	3.544	0.000***	0.697	1.434
	PA1	-0.004	0.206	-0.017	0.986	0.995	1.005
	PA2	-0.173	0.201	-0.037	0.389	0.994	1.006

a. 因变量：BI

七、调研结论与建议

（一）调研结论

综合以上调研结果，本文从社会认知现状、用户采纳动因、使用率差异性、推广渠道效果总结出以下结论。

在社会认知现状方面，市民对数币认知模糊。大多数青年群体对数币认知仅停留在国家出台的数字化货币，而部分中老年群体甚至不知道数币的存在。此外，部分用户因数币支付操作门槛

高和资金流向监管问题对其存在抵触情绪。

在用户采纳动因方面，不同用户群体的核心驱动因素存在显著差异。企业用户重视其对降本增效的实际价值以及与现有支付系统的衔接；个体户则重视收付款的操作简便与流程适配；个人用户则对优惠激励和亲友推荐更为敏感。

在使用率差异性方面，使用率与地区 GDP、人均收入及人口密度均呈显著正相关。低收入群体对政府消费券等直接经济激励更为关注，而非低收入群体更重视系统流畅度等体验属性。此外，青年中等收入群体使用数币频率较高且关注理财功能，而中年中低收入群体使用频率较低且对操作复杂性更为敏感。

在推广渠道效果方面，调研显示政府宣传与消费优惠组合推广效果最优。然而，政府宣传与银行线下活动之间存在强负相关，提示资源投放存在重复或冲突。同时，网络社交媒体可能在一定程度上稀释政策公信力。此外，政府推广活动对企业与个人用户意愿提升显著，但商场满减优惠的持续效果不佳。

（二）关于推广策略的建议

为破解数币推广中的群体、区域与渠道难题，本文提出如下政策建议：

其一，实施分群体精准激励策略。针对企业用户，政府应联

合金融机构提供系统兼容改造补贴，推动 B2B 智能合约应用；针对个体户，需推广低成本“一码多付”终端，降低技术门槛；针对个人用户，须设计梯度优惠机制，强化消费券与民生场景绑定，并依托社区开展社交推荐活动，增强传播效应。

其二，推进分区分类场景覆盖。在热点区域（如天河、越秀）试点跨境支付、智能分账等创新应用；在冷点区域（如从化）设立数币服务点，重点拓展公交、政务、农贸等高频场景，提供定向受理补贴和操作辅助，弥补基础设施与认知短板。

其三，优化渠道协同与资源布局。明确政府、银行与商户功能分工：政府侧重政策赋能与跨区域协调，银行负责系统支撑与用户培训，商圈承担场景运营与促销活动。政府应定期评估渠道效能，避免资源重复投入。

其四，构建长效推广机制与生态体系。推动支付流程优化，引入生物识别技术提升便捷性。同时，联合社区开展公众教育，强化隐私保护宣传，消除安全疑虑。建立动态监测机制，定期追踪群体使用反馈与区域渗透进展，形成“政策—市场—用户”良性互动的可持续发展生态。

参考文献

[1]Raskin, M. How do private digital currencies affect government policy? [J]. JOURNAL OF FINANCIAL STABILITY,2024, (73).

[2]Shaban, OS. Digital Currencies: Its Features and Macroeconomic Implications [A]. Tsounis, N; Vlachvei, A. Springer Proceedings in Business and Economics [C]. Bio-By-word Scientific Publishing,2020.

[3]Keere, KD. A crypto way out: cryptocurrency, techno-economic imaginaries, and crisis in Russia[J]. JOURNAL OF CULTURAL ECONOMY,2024,(12).

[4]Shakina, E. Central Bank Digital Currencies: Experimental Evidence of Deposit Conversion [J]. BE JOURNAL OF ECONOMIC ANALYSIS & POLICY,2024, (11).

[5] 陈婧 . 完善支付环境和法律保障 为数币交易提供软硬件支撑 [N]. 中国经济时报 ,2020-08-19(002).

[6] 姚进 . 使用数字人民币安全吗 [N]. 经济日报 ,2020-11-26(007).

[7] 穆长春 . 顺应技术演进和经济发展趋势 积极推进以我为主的法定数字货币 [J]. 旗帜 ,2020,(11):65-66.

[8] 加快推进数字人民币应用场景建设 [J]. 北京观察 ,2021,(09):18-19.

[9] 范子萌 . 人民银行行长易纲: 将有针对性地完善数字人民币设计和使用 [N]. 上海证券报 ,2021-11-10(002).

[10] 唐谏珍 , 黄世明 . 数字人民币支持跨境贸易发展的相关研究 [J]. 区域金融研究 ,2022,(11):78-85.

[11] 俞洁琼 . 推进数字人民币发展的现实挑战与优化研究 [J]. 商讯 ,2022,(20):92-95.

[12] 张志国 , 查迎春 . 跨境电商背景下数字人民币的推广与应用 [J]. 商展经济 ,2023,(14):71-74.

[13] 肖紫琼 , 颜文字 , 陈嘉莹 , 刘伟珍 . 居民视角下数字人民币应用的影响因素 [J]. 现代营销 (下旬刊) ,2024,(08):10-13.

[14] 张敏 . 消费者数字人民币使用行为的影响因素研究——基于修正 UTAUT2 模型 [J]. 商业经济研究 ,2024,(15):125-128.

中老年“两病”患者基层医疗机构首诊意愿影响因素分析——对广东省中山市639位中老年“两病”患者的调查

邱伊汶^{*}, 张玥, 杨美婵

广东药科大学 医药商学院, 广东 中山 528458

DOI:10.61369/ASDS.2025070007

摘 要 : 基于安德森模型, 研究中老年“两病”患者基层医疗机构首诊意愿及影响因素, 为优化“两病”门诊政策提供参考。方法: 采用便利抽样法在广东省中山市抽取639位中老年“两病”患者进行问卷调查, 将改进后的安德森模型各维度变量纳入统计分析, 探讨中老年“两病”患者基层医疗机构首诊行为的影响因素。二元 logistic 回归显示, 报销比例认知情况、支付限额认知情况、门诊慢性病种认知情况、药品费用满意程度以及治疗效果满意程度对中老年“两病”患者基层首诊意愿有显著影响。结论: 中山市中老年“两病”患者基层医疗机构首诊意愿较高。建议有关部门继续加强政策宣传, 提高基层医疗机构服务质量, 优化医保政策, 促进分级诊疗政策的实施。

关 键 词 : 安德森模型; 糖尿病与高血压; 基层首诊; 中老年患者

Analysis of Influencing Factors on the Willingness of Middle-Aged and Elderly Patients with "Two Diseases" to Seek Initial Treatment at Primary Medical Institutions—A Survey of 639 Middle-Aged and Elderly Patients with "Two Diseases" in Zhongshan City, Guangdong Province

Qiu Yiwen^{*}, Zhang Yue, Yang Meichan

School of Pharmaceutical Business, Guangdong Pharmaceutical University, Zhongshan, Guangdong 528458

Abstract : Based on the Andersen Model, this study explores the willingness of middle-aged and elderly patients with "Two Diseases" to seek initial treatment at primary medical institutions and its influencing factors, so as to provide references for optimizing the outpatient policies for "Two Diseases". Methods: A questionnaire survey was conducted among 639 middle-aged and elderly patients with "Two Diseases" in Zhongshan City, Guangdong Province using the convenient sampling method. Variables from each dimension of the modified Andersen Model were included in the statistical analysis to investigate the factors influencing the willingness of these patients to choose primary medical institutions for initial treatment. Binary logistic regression analysis showed that the following factors had a significant impact on the patients' willingness: awareness of reimbursement ratio, awareness of payment limit, awareness of outpatient chronic disease types, satisfaction with drug costs, and satisfaction with treatment effects. Conclusion: The willingness of middle-aged and elderly patients with "Two Diseases" in Zhongshan City to seek initial treatment at primary medical institutions is relatively high. It is suggested that relevant departments should continue to strengthen policy publicity, improve the service quality of primary medical institutions, optimize medical insurance policies, and promote the implementation of the hierarchical. diagnosis and treatment policy.

Keywords : Anderson model; diabetes mellitus and hypertension; primary care first consultation; middle-aged and elderly patients

基金项目: 基金项目: 2024年度省级大学生创新创业训练计划项目“基于中山市居民医保“两病”门诊统筹的老年病患基层首诊行为现状及影响因素研究”(S202410573034)。
作者简介:
张玥, 广东药科大学医药商学院, 本科生;
杨美婵, 广东药科大学医药商学院, 本科生。
通讯作者: 邱伊汶, 广东药科大学医药商学院, 本科生。

引言

目前中国患有高血压和糖尿病的人数已经高达数亿^[1]，以高血压和糖尿病（以下简称“两病”）为主的慢性非传染性疾病，在我国患病率高，并发症较多，“两病”患者的患病状况给社会、家庭以及个人带来了重大的负担。为此国家出台《关于做好2021年基本公共卫生服务项目工作的通知》及《国家基本公共卫生服务规范（第三版）》要求要以重点人群为切入点提升基本公共卫生服务质量，切实提升慢病规范管理质量^[2]。广东省中山市已通过出台《城乡居民高血压糖尿病门诊用药保障机制》，建立门诊慢性病种（两病专病联盟）及优化慢性病种患者医保统筹待遇等措施，引导“两病”患者的基层首诊行为。研究表明^[3-7]基层医疗机构对于慢性病的管理具有关键作用，而“两病”患者个体的经济负担、对医疗保障的满意度在形成门诊统筹和基层首诊的良性循环中发挥着重要的作用。目前中老年群体构成了“两病”的主要患者群体^[8]。基于此背景，本研究基于安德森模型对中老年“两病”患者基层医疗机构首诊现状及影响因素进行研究，助力中山市及类似地区优化“两病”门诊政策，加快实现分级诊疗新格局，促进医疗资源合理配置。

一、资料来源与方法

（一）研究对象

以年龄大于45周岁（基于2018年中国健康与养老追踪调查数据标准，选取45岁及以上中老年人作为研究对象^[9]），且参加了中山市居民医保的中老年“两病”患者为研究对象，采用便利抽样法，于2023年10月到12月到社区卫生服务中心及周边地方连续进行走访调查，以了解中山市中老年“两病”患者基层医疗机构首诊行为现状。调查员对调查对象发放问卷700份，回收有效样本639份，有效率为91%。

（二）研究方法

本研究采用便利抽样法与分层抽样法。在中山市东区、西区、南区等五大区域的社区卫生服务中心、乡镇卫生院及周邊社区，直接选取符合纳入标准（年龄 ≥ 45 岁、已确诊高血压/糖尿病、参与中山市居民医保）的中老年患者作为初始样本，为确保区域代表性，在完成方便抽样后，再根据中山市各镇区人口比例进行分层补充抽样。调研团队在医疗机构门诊大厅、慢性病管理科候诊区等患者集中区域，现场邀请符合条件的患者填写问卷。同时与中山市各大医院及社区卫生服务中心合作，在医院的“两病”相关科室等区域张贴专属二维码海报或摆放宣传展架。安排专人加入中山市本地的“两病”患者线上交流群、论坛等网络社交平台，发布问卷，鼓励群成员积极参与问卷填写。

1. 模型设计与变量选择

安德森模型（Andersen Model）由美国学者罗纳·安德森于1968年提出^[10]，是分析医疗卫生服务利用行为的经典框架，并广泛应用于健康研究领域。该模型的核心在于探讨影响个体或家庭使用医疗服务的多重因素。2013版安德森模型对原有框架进行了更新，强调环境因素、人口特征、医疗行为和医疗结果四大要素的相互作用，认为这些因素共同决定了居民在特定环境下的医疗需求和行为。本研究通过改良安德森模型得到中老年两病患者首诊意愿影响因素模型，见图1，并以此设计问卷，因变量为中老年“两病”患者首诊意愿，通过问卷问题“您在进行‘两病’就诊时会选择基层首诊吗？”确定患者是否选择基层首诊；自变量包括人口特征维度、环境维度、医疗结果维度三个层面。个人

特征维度中的人口倾向特征选择年龄、性别、职业以及受教育程度，使能资源选择收入、社区医保政策宣传频率，取药渠道，需求选择自感健康状况。环境维度选择两病门诊统筹待遇政策认知情况具体包括（报销比例认知情况、支付限额认知情况）、基层医疗体系建设认知情况包括（分级诊疗政策认知情况、门诊慢性病种认知情况）。医疗结果维度选择药品费用满意程度、治疗效果满意程度。

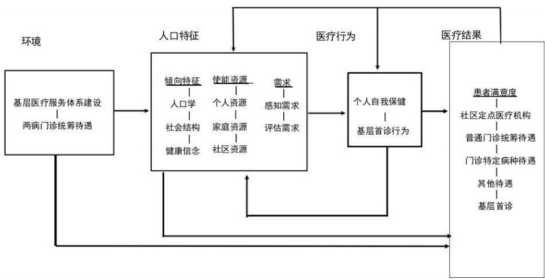


图1 基于安德森模型的中老年两病患者首诊意愿影响因素模型

Figure 1: Model of Influencing Factors on the Willingness of Middle-Aged and Elderly Patients with “Two Diseases” to Seek Initial Treatment Based on the Andersen Model

2. 问卷信效度检验

本研究基于改进后的安德森模型设计了问卷，并进行了整体信效度检验。问卷的 Cronbach’s Alpha 系数为0.829，效度值为0.756，信效度检验结果表明该问卷在测量目的上具有可靠性和有效性，因此认为数据适合进行后续分析。

（三）统计学方法

使用 IBM SPSS 26.0 分析数据，利用卡方检验进行单因素分析，通过二元 logistics 回归模型进行多因素分析。

二、结果

（一）调查对象基本情况

在被调查的639位研究对象中，在性别上，男性占总人数的38%，女性占62%。在年龄上，60岁及以上的占总人数的25.8%，45岁到60岁占74.2%。在文化程度上，初中及以下的占总人数的41.3%，高中或中专的占总人数的31.0%，大学本科或专科的

占总数的25.8%，研究生及以上的占总数的1.9%。在月收入上，1000元及以下的两病患者占总人数的13.6%，1000~2000元占27.2%，2000~3000元占24.40%，3000元以上患者34.70%。无固定职业占总人数27.2%，企事业单位职工占30.0%，个体工商户占24.9%，农民占17.8%。

（二）中老年“两病”患者基层首诊意愿的单因素分析

采用卡方检验比较组间差异，结果见表1，对中山市居民医保“两病”门诊统筹保障待遇的不同报销比例认知情况、支付限额认知情况、门诊慢性病种认知情况、药品费用满意程度以及治疗效果满意程度的组别居民的中医药服务首诊意愿有统计学差异（ $P<0.05$ ）；不同年龄、性别、文化程度、个人收入、社区医保政策宣传频率、自感健康状况、取药渠道、分级诊疗政策的认知情况的居民中医药服务首诊意愿不存在统计学差异（ $P>0.05$ ）见表1。

表1 中老年“两病”患者基层首诊意愿卡方检验表（n=639）

Table 1: Chi-Square Test Table of Willingness of Middle-Aged and Elderly Patients with “Two Diseases” to Seek Initial Treatment at Primary Level (n=639)

变量	组别	愿意 [n(%)]	不愿意 [n(%)]	χ^2	P
年龄				0.645	0.694
	60 岁以下	432 (91.14)	42 (8.86)		
	60 岁及以上	156 (94.55)	9 (5.45)		
性别				0.059	0.808
	男	225 (92.59)	18 (7.41)		
	女	363 (91.67)	33 (8.33)		
文化程度				2.868	0.413
	初中及以下	249 (94.32)	15 (5.68)		
	高中 / 中专	177 (89.39)	21 (10.61)		
	大学本科 / 专科	153 (92.72)	12 (7.28)		
	硕士研究生及以上	9 (75.00)	3 (25.00)		
职业				9.377	0.025
	无固定职业	144 (82.76)	30 (17.24)		
	企事业单位职工	183 (95.31)	9 (4.69)		
	个体工商户	153 (96.23)	6 (3.77)		
	农民	108 (94.74)	6 (5.26)		
月收入（元）				3.976	0.264
	≤ 1000	72 (82.76)	15 (17.24)		
	1001~2000	162 (93.10)	12 (6.90)		
	2001~3000	147 (94.23)	9 (5.77)		
	>3000	207 (93.24)	15 (6.76)		
医保宣传频率				0.337	0.845
	不经常	117 (90.70)	12 (9.30)		
	一般	249 (93.26)	18 (6.74)		
	经常	222 (91.36)	21 (8.64)		
自感健康状况				0.545	0.761
	差	42 (93.33)	3 (6.67)		
	一般	291 (90.65)	30 (9.35)		
	好	255 (93.40)	18 (6.60)		
取药渠道				5.658	0.129

变量	组别	愿意 [n(%)]	不愿意 [n(%)]	χ^2	P
	无需取药	69 (85.19)	12 (14.81)		
	线上购药	69 (85.19)	12 (14.81)		
	一级及以下医院	243 (96.43)	9 (3.57)		
	二三级医院	207 (92.00)	18 (8.00)		
报销比例认知				34.264	<0.001
	不了解	69 (74.19)	24 (25.81)		
	一般	288 (94.12)	18 (5.88)		
	了解	231 (96.25)	9 (3.75)		
支付限额认知				41.512	<0.001
	不了解	84 (71.79)	33 (28.21)		
	一般	288 (94.12)	18 (5.88)		
	了解	216 (97.30)	6 (2.70)		
分级诊疗认知				9.458	0.009
	不了解	120 (85.11)	21 (14.89)		
	一般	297 (93.40)	21 (6.60)		
	了解	171 (95.00)	9 (5.00)		
门诊慢病认知				52.098	<0.001
	不了解	117 (76.47)	36 (23.53)		
	一般	270 (96.77)	9 (3.23)		
	了解	201 (97.10)	6 (2.90)		
药品费用满意度				71.539	<0.001
	不满意	21 (53.85)	18 (46.15)		
	一般	282 (91.26)	27 (8.74)		
	满意	285 (97.94)	6 (2.06)		
治疗效果满意度				56.421	<0.001
	不满意	12 (50.00)	12 (50.00)		
	一般	264 (89.80)	30 (10.20)		
	满意	312 (97.20)	9 (2.80)		

（三）中老年“两病”患者基层首诊意愿的二元 logistic 回归分析

将各维度纳入二元 logistic 模型中，结果见表2所示：

1. 在人口特征维度上，倾向特征中，二元 logistic 回归结果显示相比起无固定职业或其它职业的“两病”患者，企事业单位职工的“两病”患者首诊意愿更低（ $P<0.05$ ， $OR=0.050$ ）。

2. 在需求维度上，习惯在二级或三级医院门诊取药的“两病”患者相比于无需服药的患者基层首诊意愿高（ $P<0.05$ ， $OR=20.264$ ）；

3. 在环境维度上，对“两病”门诊统筹保障待遇报销比例（ $P<0.05$ ， $OR=12.173$ ）、门诊慢性病种（ $P<0.05$ ， $OR=8.031$ ）了解度高的“两病”患者基层首诊意愿高，而对支付限额了解度高的患者基层首诊意愿低（ $p < 0.05$ ， $OR=0.098$ ）；

4. 在患者满意度维度上，对药品费用（ $P<0.05$ ， $OR=6.155$ ）、治疗效果（ $P<0.05$ ， $OR=7.070$ ）满意度高的“两病”患者基层首诊意愿高；

综上，习惯在二级或三级医院门诊取药，对“两病”门诊统

筹保障待遇报销比例和门诊慢性病种认知度高的，以及对药品费用和治疗效果满意度高的“两病”患者基层首诊意愿高。

表2 中老年“两病”患者基层首诊意愿的二元 Logistic 回归分析 Table 2: Binary Logistic Regression Analysis of Willingness of Middle-Aged and Elderly Patients with “Two Diseases” to Seek Initial Treatment at Primary Level						
变量组	变量及参照组	B	Wald X ²	P	OR	95% CI
倾向特征	职业（参照：无固定职业）					
	企事业单位职工	-2.995	4.683	0.030	0.050	(0.003, 0.754)
	个体工商户	-2.454	2.672	0.102	0.086	(0.005, 1.630)
	农民	-0.741	0.220	0.639	0.476	(0.021, 10.569)
用药需求	（参照：无需服药）					
	医保平台线上购药	0.342	0.098	0.754	1.408	(0.165, 12.009)
	一级及以下医院门诊取药	0.376	0.105	0.746	1.457	(0.149, 14.228)
保障待遇与认知	二三级医院门诊取药环境	3.009	4.383	0.036	20.264	(1.212, 338.944)
	报销比例认知	2.499	7.015	0.008	12.173	(1.915, 77.380)
	支付限额认知	-2.323	5.734	0.017	0.098	(0.015, 0.656)
	门诊慢性病种认知	2.083	7.055	0.008	8.031	(1.726, 37.361)
患者满意度	药品费用满意度	1.817	4.092	0.043	6.155	(1.058, 35.795)
	治疗效果满意度	1.956	4.833	0.028	7.070	(1.236, 40.429)

三、结论

（一）倾向特征与需求对中老年“两病”患者基层首诊行为的影响倾向

倾向特征方面，以无固定职业及其它为参照，企事业单位职工在基层医疗机构首诊行为相对较低与吕凯等学者^[11]结论相似。原因可能在于企事业单位职工经济收入较稳定，更偏好治疗效果与服务水平高的医疗机构。而基层医疗机构可能存在服务能力不足的问题，包括医疗设施、医务人员的技术水平等，许多患者对基层医疗机构信任度降低，尤其是在病情较重或需要专业治疗时，他们更愿意选择三级医院^[12, 13]。而在需求层面对于三级医院门诊取药患者首诊行为相对较高的现状，可能与中山市三级医院与基层医疗机构用药待遇基本一致相关。

（二）环境与患者满意度是中老年“两病”患者基层医疗机构首诊行为的影响因素

进一步研究发现，环境和患者满意度是中老年“两病”患者基层医疗机构首诊行为的影响因素。具体而言，环境因素主要聚焦于“两病”门诊统筹保障待遇，涵盖报销比例、支付限额、门诊慢性病种，以及特定门诊慢性病种（例如两病专病联盟）的知晓程度。而患者满意度则侧重于中老年“两病”患者对基层医疗机构药品费用和治疗效果的满意程度。在环境因素方面，对报销比例和门诊慢性病种了解度高的患者更倾向于选择基层首诊。然而，支付限额的了解程度却与基层首诊意愿呈负相关，这可能是因为患者认为支付限额过低，无法满足其医疗需求，从而更倾向于选择高级别的医疗机构。针对患者满意度而言，本研究发现居民基层就诊效果满意度越高，则基层医疗机构首诊意愿越高，与吕凯等^[11]的研究结果类似。即患者过往的医疗经历会深刻影响他们的医疗行为选择。当患者对基层医疗机构提供的药品费用和治疗效果感到满意时，他们就会对基层医疗服务产生更高的信任感和依赖感，从而更倾向于在未来的就医过程中选择基层首诊。

四、建议

（一）多渠道宣传门诊统筹政策，提高参保人员政策认知度参与度

有关机构应该加强门诊统筹政策的宣传力度，提高两病门诊统筹的群众认知度，医疗机构可以在其网站公众号等新媒体平台发布门诊统筹相关的政策信息，方便患者进行查阅和了解。同时，医疗机构也可以组织医务人员在机构内开展宣传活动，对过往的就诊人员进行门诊统筹政策的宣传与普及。社区可以通过制作相应的宣传海报、宣传单等资料，向社区内居民进行发放，同时也可以组织相应的宣传活动，提高“两病”患者前往基层医疗机构就诊的意识，引导患者前往基层医疗机构就诊，逐渐形成“小病在基层，大病到医院，康复回基层”的合理就医格局。

（二）强基层、惠民生，提升基层医疗机构服务质量与患者满意度

提高慢性病门诊医疗服务质量，增强两病对基层医疗机构的满意度。首先，加大对基层医务人员的培训力度，提高其专业能力和服务水平，增强患者对基层医疗机构的信任度。建议一方面通过培训、轮岗等渠道提升现有人员技术水平，另一方面通过建立医供体、医联体让高素质医生在基层机构执业。其次，在医疗资源下沉至基层的基础上，开展多样化的医疗保健服务，推动家庭医生签约服务，密切关注群众身体健康，使群众能享受优质快捷的健康卫生服务，并加强对“两病”患者的病情回访，提高基层医疗机构的吸引力。

（三）继续优化医保政策，降低患者就医成本与报销难度

一方面继续提高基层医疗机构的医保报销比例，扩大报销范围，以鼓励患者优先选择基层医疗机构进行首诊。

另一方面简化报销流程，减少患者的报销难度和时间成本，提高医疗服务的可及性和效率。根据实地调查，中老年两病患者要完成门诊统筹的报销需要提供较为复杂的证明材料，同时需要往返多个办事窗口，这都体现出门诊统筹报销手续的繁琐。因

此，建议通过医保信息系统的升级改造，实现门诊费用的在线直接结算。参保人员在就医购药时，仅需支付个人自付部分，而门诊统筹基金支付的部分则由定点医疗机构与医保经办机构依据协议规定进行结算，减轻患者的经济负担和时间成本。

参考文献

- [1] 中国心血管健康与疾病报告编写组. (2022). 中国心血管健康与疾病报告2021 概要. 心脑血管病防治, 22(04), 20-36+40.
- [2] 张俊丽, 张毅华, 苏文, 等. (2023). 横琴粤澳深度合作区“两病”患者用药减免政策实施效果调查. 医学与哲学, 44(09), 31-34.
- [3] Kringos DS, Boerma WG, Hutchinson A, et al. The breadth of primary care: a systematic literature review of its core dimensions[J]. BMC health services research, 2010, 10(1): 65.
- [4] Madden, H., Harris, J., Blickem, C. et al. “Always paracetamol, they give them paracetamol for everything”: a qualitative study examining Eastern European migrants’ experiences of the UK health service. BMC Health Serv Res 17, 604 (2017). <https://doi.org/10.1186/s12913-017-2526-3>.
- [5] The Models of Primary Care Organisation in Europe, Canada, Australia and New-Zealand (nivel.nl).
- [6] LIPTON B J. Association between Health Insurance and Health among Adults with Diabetes: Evidence from Medicare [J]. Journal of the American Geriatrics Society, 2020, 68(2): 388-94.
- [7] YOU C H, CHOI J H, KANG S, et al. Association between supplementary private health insurance and visits to physician offices versus hospital outpatient departments among adults with diabetes in the universal public insurance system [J]. Plos One, 2018, 13(4).
- [8] 中国居民营养与慢性病状况报告编写组. (2020). 中国居民营养与慢性病状况报告（2020年）. 营养学报, 42(06), 521.
- [9] 刘贝贝, 田庆丰, 郭金玲. (2022). 我国中老年人群慢性病患病现状及共病模式分析. 医学与社会, 35(08), 58-61+66. DOI: 10.13723/j.yxysh.2022.08.011.
- [10] Andersen, R. M. (1968). FAMILIES' USE OF HEALTH SERVICES: A BEHAVIORAL MODEL OF PREDISPOSING, ENABLING, AND NEED COMPONENTS. Purdue University.
- [11] 吕凯, 周娜, 吕巧红, 等. (2023). 医共体模式下海岛居民基层首诊意愿及行为的影响因素分析. 护理与康复, 22(09), 63-67.
- [12] 吕艳霞, 胡盟, & 徐爱军. (2021). 江苏省门诊患者利用中医药服务行为的影响因素探讨. 中国药房, 32(21), 2668-2672.
- [13] 王薇, 李力桢, & 任毅. (2018). 广州市基层首诊政策执行困境及对策. 现代医院管理, 16(02), 17-19+23.

ESG 绩效对农林牧渔业企业融资约束的影响分析

——以企业声誉中介变量

庄丽凤¹, 陈晓彤¹, 黄洁洁², 裴祺阳³, 陈永宁¹, 陈立宇³, 周燕^{1*}

1. 华南农业大学 数学与信息学院, 广东 广州 510642

2. 华南农业大学 经济管理学院, 广东 广州 510642

3. 华南农业大学 软件学院, 广东 广州 510642

DOI:10.61369/ASDS.2025070008

摘 要 : 在高质量发展的背景下, ESG 绩效已成为衡量企业综合实力与长期发展潜力的关键指标。本文聚焦农林牧渔业上市公司, 基于信号传递、治理优化与声誉资本理论, 构建“ESG—声誉—融资”分析框架, 系统探讨 ESG 绩效对企业融资约束的影响路径, 并重点考察企业声誉的中介作用。研究选取 2019—2023 年申银万国行业分类中 102 家农林牧渔类上市公司为样本, 结合华政 ESG 评级数据与 CSMAR 数据库中的财务数据, 运用因子分析法构建企业声誉指标, 采用固定效应模型与中介效应检验方法进行实证分析。结果表明 ESG 绩效显著缓解企业融资约束, 企业声誉在二者之间发挥部分中介作用, 且该缓解效应在小规模企业中更为显著。本文创新性地将 ESG 表现、企业声誉与融资约束纳入统一分析框架, 填补了农林牧渔行业相关研究的空白, 为投资者决策和企业融资实践提供了理论支持与实践参考。

关 键 词 : ESG 绩效; 农林牧渔企业; 融资约束; 企业声誉; 中介效应

Analysis of the Impact of ESG Performance on the Financing Constraints of Agriculture, Forestry, Animal Husbandry and Fishery Enterprises — Taking Corporate Reputation as a Mediating Variable

Zhuang Lifeng¹, Chen Xiaotong¹, Huang Jiejie², Pei Qiyang³, Chen Yongning¹, Chen Liyu³, Zhou Yan^{1*}

1. College of Mathematics and Information, South China Agricultural University, Guangzhou, Guangdong 510642

2. College of Economics and Management, South China Agricultural University, Guangzhou, Guangdong 510642

3. School of Software, South China Agricultural University, Guangzhou, Guangdong 510642

Abstract : In the context of high-quality development, ESG performance has become a key indicator for measuring the comprehensive strength and long-term development potential of enterprises. This article focuses on listed companies in agriculture, forestry, animal husbandry, and fisheries. Based on signal transmission, governance optimization, and reputation capital theory, a "ESG reputation financing" analysis framework is constructed to systematically explore the impact path of ESG performance on corporate financing constraints, with a particular emphasis on the mediating role of corporate reputation. The study selected 102 listed companies in the agriculture, forestry, animal husbandry, and fishery industries in the Shenyin Wanguo industry classification from 2019 to 2023 as samples, combined with Huazheng ESG rating data and financial data in the CSMAR database. Factor analysis was used to construct corporate reputation indicators, and fixed effects models and mediation effect testing methods were used for empirical analysis. The results indicate that ESG performance significantly alleviates corporate financing constraints, and corporate reputation plays a partial mediating role between the two, with this alleviation effect being more significant in small-scale enterprises. This article innovatively incorporates ESG performance, corporate reputation, and financing

基金项目: 华南农业大学数学与信息学院、软件学院 2024 年国家级大学生创新创业训练项目; 2025 年全国统计建模大赛省级三等奖; 广东省高等教育教学改革项目“基于创新能力培养的统计学专业数据分析类实验课程改革探索与实践”(粤教高函[2024]9号516)。

作者简介:

庄丽凤(2004—), 女, 学士, 研究方向为应用统计;

陈晓彤(2004—), 女, 学士, 研究方向为应用统计;

黄洁洁(2003—), 女, 学士, 研究方向为金融;

裴祺阳(2005—), 男, 学士, 研究方向为软件工程;

陈永宁(2004—), 女, 学士, 研究方向为应用统计;

陈立宇(2005—), 男, 学士, 研究方向为软件工程。

通讯作者: 周燕, 华南农业大学数学与信息学院, 讲师, 研究方向为数据挖掘与智能决策。

constraints into a unified analysis framework, filling the gap in research related to the agriculture, forestry, animal husbandry, and fishery industries, and providing theoretical support and practical reference for investor decision-making and corporate financing practices

Keywords : ESG performance; agriculture forestry animal husbandry and fishery enterprises; financing constraints; corporate reputation; intermediary effect

在高质量发展的背景下，企业 ESG 绩效成为衡量其综合实力和长期发展潜力的关键指标。本研究聚焦于申银万国行业农林牧渔类上市公司，依托信号传递、治理优化及声誉资本理论，构建“ESG-声誉-融资”框架，旨在深入探讨 ESG 绩效对企业融资约束的影响及其通过企业声誉发挥的中介作用。运用文献分析法和实证研究法，填补现有研究中关于 ESG 与企业声誉关联的空白，为理解融资约束提供新视角，研究结果对提升 ESG 绩效和强化企业声誉以获取融资优势具有重要的理论和实践价值。

研究创新体现在三个方面，一是将 ESG 绩效与企业融资约束的联系通过企业声誉这一中介变量进行系统性分析，为现有文献提供新的理论视角；二是聚焦农林牧渔行业，首次实证分析 ESG 绩效、企业声誉与融资约束的相互作用，填补研究空白；三是为投资者决策提供 ESG 因素和企业声誉的参考，具有实践指导意义。

一、文献综述

（一）ESG 绩效的相关研究

ESG 研究自 2011 年起逐步发展，早期多聚焦于 ESG 与企业财务绩效的关联^[1]，朱紫妮（2024）发现 ESG 通过提升信息披露质量和降低融资成本正向影响企业绩效^[2]。随着可持续发展议题的推进，2019 年后研究重点转向 ESG 的治理机制与外部驱动因素。Khaw 等（2024）指出 ESG 绩效受企业规模、行业特征及监管环境的显著影响^[3]。近年来学者开始关注 ESG 对企业韧性及融资能力的直接作用，发现 ESG 表现能通过降低经营风险、优化资源配置提升企业的抗风险能力^[4]。

（二）农林牧渔业企业融资约束研究

农林牧渔企业普遍面临融资渠道单一、抵押物不足和信用评级偏低等问题。受行业特性影响，生产周期长、自然风险高导致金融机构对涉农企业信贷持审慎态度，中小型畜牧企业因缺乏有效抵押物，贷款利率较基准上浮 20%-50%（孙宇等，2015）^[5]。此外，政策支持对融资约束也有缓解效果，税收优惠和财政补贴的调节作用更显著（朱柏洋，2023）^[6]。综上农林牧渔企业融资难题需结合行业特性和政策环境解决。

（三）ESG 表现对农林牧渔企业融资约束的影响

研究表明 ESG 表现通过信号传递、治理优化及声誉积累三条路径缓解融资约束。张亚连（2023）发现高 ESG 评分企业融资约束降低 5%，且小企业效应更显著^[7]。在治理效应上 ESG 实践优化内部治理结构，降低代理成本。刘树良等（2024）指出 ESG 表现通过减少管理层短视行为，提升投资效率，进而缓解融资压力^[8]。在声誉中介作用上，企业声誉是 ESG 影响融资约束的关键中介变量。邵鹏等（2024）发现 ESG 表现的同群效应可增强资本市场信任，降低融资成本^[9]。

（四）ESG 表现对企业声誉的影响机制研究

企业声誉是利益相关者对企业的综合评价，反映企业满足利益相关者利益诉求的程度。只有企业切实履行对利益相关者的承诺，企业才能拥有良好声誉，实现可持续发展^[10]。ESG 信息披露

有利于缓解企业与利益相关者之间的信息不对称，提高企业的信息透明度，积累声誉资本^[11]。企业践行 ESG 并进行信息的披露释放出积极信号，促进与利益相关者间的良好关系，带来良好声誉，从而增强企业竞争力^[12]。

综上所述，现有研究虽揭示了 ESG 绩效缓解融资约束的路径及农林牧渔行业特殊性，但多聚焦单一 ESG 指标和传统行业，缺乏对该行业自然风险与政策敏感性的针对性分析，且未量化声誉中介贡献度及动态机制。为此，本研究以农林牧渔上市公司为对象，基于信号传递理论，构建动态中介模型，重点探究行业融资约束特征、ESG 绩效的影响、企业声誉的中介作用和企业规模的调节效应。通过对比 E/S/G 分维度贡献差异，提出针对性融资方案，兼具理论价值与实践意义。

二、研究设计

（一）分析农林牧渔业企业的基本特征与融资现状

利用企业样本数据进行描述性统计与可视化分析，揭示农林牧渔业企业的基本特征及其融资约束机制。

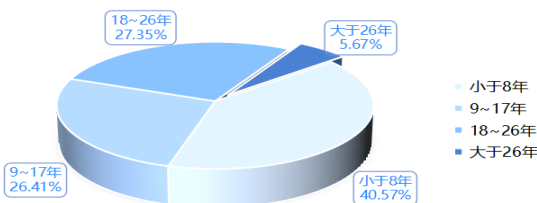


图1 年龄分布

根据饼图可知农林牧渔业企业中年轻企业占比较大，成立时长小于8年的企业占比最高，达到40.57%，9至17年的企业，占比26.41%，说明多数企业处于初创阶段，有较高的成长潜力和融资需求，但缺乏稳定的现金流与多样的融资渠道。

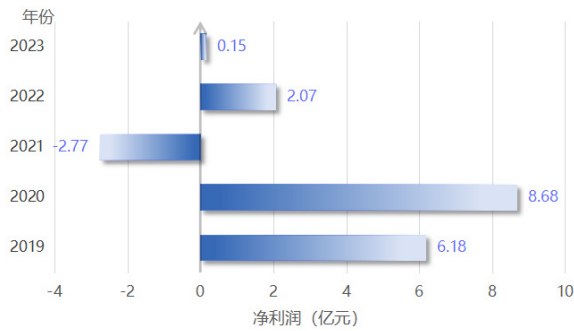


图2 不同年度的净利润

由图2知农林牧渔企业从2019年到2023年，企业净利润经历了显著波动，2020年达到最高点8.68亿元，而2021年净利润为-2.77亿元。净利润的波动显示其财务状况不稳定，融资风险增加，特别是在亏损年份，易影响投资者和贷款机构的信心。

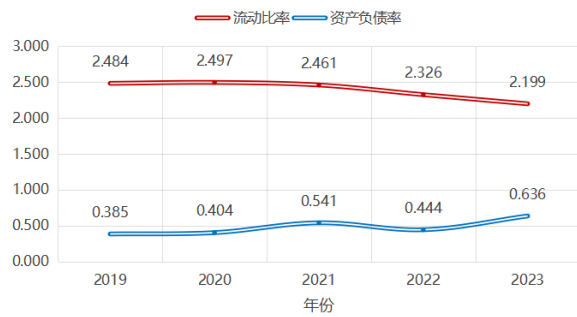


图3 流动比率与资产负债率变化

根据农林牧渔业企业的流动比率和资产负债率的变化趋势可以得到，流动比率整体上呈现出下降趋势，而资产负债率整体上呈现出上升趋势，表明企业的短期偿债能力减弱，对外部资金的需求增加，同时在融资结构上更依赖于长期债务。

（二）变量的定义

1. 被解释变量

参考顾雷等（2020）的研究^[13]，将logit模型回归拟合企业融资约束发生的概率定义为融资约束指数FC。同理本文采用FC指数度量农林牧渔企业的融资约束，随着FC增大，企业的融资约束程度也越大。

2. 解释变量

借鉴多名学者的研究，同时考虑到ESG数据的易获得性、完整性和可靠性，本研究利用华政ESG评分数据度量ESG表现，其分为C、CC、CCC、B、BB、BBB、A、AA、AAA共9个等级，从1至9分别赋值。

3. 中介变量

基于理论本文选取企业声誉和网络舆情影响农林牧渔企业融资的中介变量，参考管考磊等（2019）的度量方法，构建包含消费者和社会角度对企业资产、收入、净利润和价值在行业内排名的评估，债权人角度对资产负债率、流动比率、长期负债比的关注，股东角度对每股收益、每股股利以及是否由国际四大会计师事务所审计的重视，以及企业角度对可持续增长率和独立董事比例的关注共12个评价指标的体系。对这些指标进行因子分析，将主成分分析得到的4个因子进行旋转，得到的方差累计贡献率作为权重，计算企业声誉的综合得分，再从低到高分10组，分别赋

值1~10^[14]。

（1）KMO和Bartlett检验

表1 KMO和Bartlett检验结果

KMO和Bartlett检验		
KMO检验		0.796
Bartlett检验	近似卡方	19283.901
	自由度	55
	显著性	0.000*

确定指标并利用SPSS软件进行因子分析，当KMO>0.8，非常适合主成分分析，0.7<KMO<0.8，适合主成分分析，0.6<KMO<0.7，比较适合，KMO<0.6则不适合。结果显示KMO值为0.796，适合主成分分析。Bartlett检验在5%水平下显著性为0.000*，说明度量企业声誉的变量具有相关性且显著，适合提取公因子。

（2）计算特征根的方差贡献率和旋转成分矩阵

表2 总方差解释

因子	特征值	方差	方差百分比	累积贡献率
Factor1	2.630	2.462	0.2239	0.2239
Factor2	2.063	2.082	0.1893	0.4131
Factor3	1.341	1.464	0.1331	0.5463
Factor4	1.156	1.181	0.1074	0.6537

由表2可知基于特征值大于1的标准，使用主成分分析法提取了4个公因子，利用最大方差法得Factor1的特征值为2.630，Factor2的特征值为2.063，这4个的方差贡献率为65.37%，说明原有指标丢失的信息较少，分析效果理想。

由碎石图得越高点其势能越大，重要性越强。第1个因子的特征值最高，对解释原有指标的贡献最大，第5个以后的特征值都小于1，对解释原有指标的贡献小，可忽略，因此提取4个因子是最适合的。

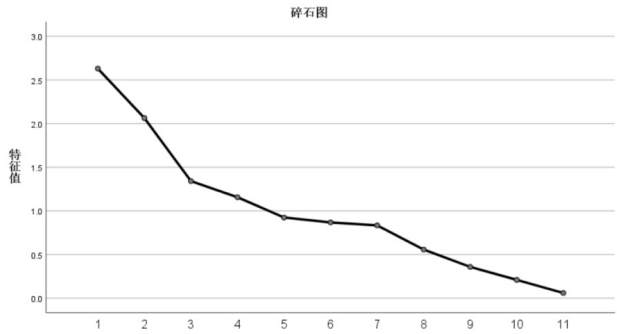


图4 碎石图

主成分分析共提取了4个因子，成分得分系数矩阵如表3。

表3 成分得分系数矩阵

	Factor1	Factor2	Factor3	Factor4
x1	0.010	0.456	0.036	0.047
x2	-0.002	0.455	0.034	0.054
x3	0.327	0.037	-0.074	-0.018
x4	-0.169	0.151	-0.322	-0.190
x5	0.012	-0.040	0.546	0.011
x6	-0.156	0.195	0.479	-0.160
x7	0.377	-0.051	-0.023	-0.046
x8	0.286	0.028	0.018	0.046
x9	0.007	-0.094	-0.114	-0.569
x10	0.027	-0.068	-0.223	0.570
x11	-0.180	0.069	0.055	0.362

4. 控制变量

为排除其他因素影响，借鉴现有文献，本研究对企业年龄、企业规模、资本结构、总资产收益率、固定资产比率、股权集中度3、独立董事占比、两职合一进行控制，增强实证分析结果的可信度。

表4变量解释

变量类型	变量名称	变量符号	变量定义
被解释变量	融资约束	FC	融资约束 FC 指数
解释变量	ESG 表现	ESG	华政 ESG 评级，赋值 1~9
中介变量	企业声誉	REP	因子分析得分后排序赋值 1~10
	企业年龄	Age	企业已生存年数
	企业规模	Size	企业总资产取自然对数
	资本结构	Lev	资产负债率
控制变量	总资产收益率	ROA	净利润 / 平均资产总额
	固定资产比率	Fix	固定资产 / 资产总额
	股权集中度 3	Top3	公司股东持股比例
	独立董事占比	Indep	独立董事数量 / 董事会人数
	两职合一	Dual	董事长与经理兼任取 1，否则为 0

5. 理论模型

基于理论分析和变量定义，提出研究理论模型，ESG 绩效直接影响农林牧渔企业融资约束与 ESG 绩效通过企业声誉影响农林牧渔企业融资约束。路径如图 5 所示：

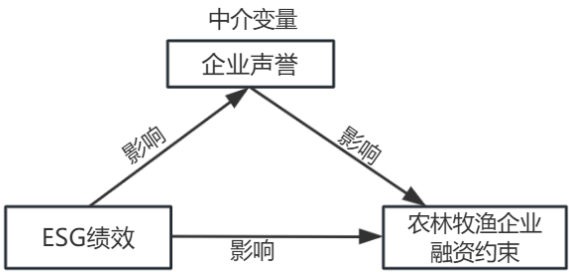


图5理论路径图

(三) 研究假设

针对理论路径的研究分析，提出如下研究假说：
H1：ESG 绩效与农林牧渔企业融资约束呈负相关关系。
H2：企业声誉在 ESG 绩效与农林牧渔企业融资约束之间起中介作用。

- H21：ESG 绩效对企业声誉有显著影响。
- H22：企业声誉对农林牧渔企业融资约束有显著影响。

(四) 样本选取与数据处理

本文以申银万国行业的农林牧渔类上市公司为研究对象，从 CSMAR 数据库中获取了 2019–2023 年的 113 家企业财务数据，包括企业规模、资产负债率、净利润等，总共 495 条面板数据。ESG 评级和得分数据来自 Wind 数据库。为保证数据分析的完整性和可信度，利用 Excel 对数据进行预处理：（1）剔除 ST、*ST 及 PT 类上市公司样本；（2）剔除主要变量缺失的样本，少量缺失数据采用插值法填补；（3）对连续变量进行上下 1% 的缩尾处理。最终得到 102 个企业的 458 条有效面板数据。

(五) 模型设计

1. ESG 绩效与农林牧渔企业融资约束的影响模型

$$FC_{i,t}=\alpha_0+\alpha_1ESG_{i,t}+\sum\alpha_2Controls+\sum Year+\epsilon$$
 公式（1）

2. 企业声誉在 ESG 绩效与农林牧渔企业融资约束的中介效应模型

$$REP_{i,t}=\beta_0+\beta_1ESG_{i,t}+\sum\beta_2Controls+\sum Year+\mu$$
 公式（2）

$$FC_{i,t}=\gamma_0+\gamma_1ESG_{i,t}+\gamma_2REP_{i,t}+\sum\gamma_3Controls+\sum Year+\upsilon$$
 公式（3）

三、实证分析

(一) 描述性统计与变量相关性分析

表5描述性统计

Variable	Obs	Mean	Std. Dev.	Min	Max
id	457	49.403	28.591	1	102
year	457	2021.131	1.405	2019	2023
fc	457	0.483	0.265	0.00016	0.994
esg	457	3.836	0.94	1	7
rep	457	5.473	2.865	1	10
age	457	12.586	8.343	0	31
size	457	22.194	1.211	19.288	26.198
lev	457	0.431	0.211	0.031	1.484
roa	457	-0.532	11.022	-235.096	3.627
fix	457	0.267	0.141	0.012	0.704
top3	457	50.366	15.185	14.074	91.506
indep	457	0.387	0.062	0.222	0.6
dual	457	0.304	0.461	0	1

根据 Hausman 检验结果 $P<0.05$ ，选择固定效应模型。对变量进行描述统计，结果如表 5。融资约束指标 FC 的均值为 0.483，标准差为 0.265，最大值为 0.994，最小值为 0.00016，说明融资约束较为严重且不同企业间融资约束差异显著。ESG 评级的均值为 3.836，标准差为 0.94，最大值为 7，最小值为 1，说明各企业间 ESG 表现存在较大差异，总体上企业的 ESG 水平较差，有较大提升空间。

企业的资产负债率均值为 43.1%，标准差为 0.211，最大值为 148.4%，最小值为 3.1%，表明农林牧渔企业资产负债率差异较大，行业整体负债水平适中，但部分企业杠杆率极高或极低，风险分化明显。多重共线性检验显示 VIF 小于 5，说明数据合理。

对变量做相关性分析，结果显示融资约束 FC 与 ESG 评级呈较强的负相关性，相关系数为 -0.19，显著性为 **，表明 FC 与 ESG 具有一定的相关关系，初步验证了假设。ESG 表现与企业声誉的相关系数为 0.29，ESG 表现良好可能提升企业声誉，缓解融资约束，进一步的实证分析将验证相关性。

(二) 回归结果

为探究 ESG 表现对农林牧渔企业融资约束的影响，以年份固定效应，借助 stata17 对样本进行多元回归分析，输出主效应结果如表 6。数据显示 ESG 评级与融资约束 FC 指数系数为 -0.007，在 5% 水平显著为负，表明农林牧渔企业 ESG 评级越高，其面临的融资约束越低，ESG 每提升一个等级，融资约束 FC 将下降 0.007，说明良好的 ESG 能够有效缓解该行业面临的融资限制，

通过降低环境风险、提升合规性和增强投资者信心等渠道发挥作用，故假设 H1 成立。

企业年龄 (Age) 与融资约束 FC 的系数为 -0.003 ，表明年龄越大的企业，融资约束越低。相较于年轻企业，较老的经营时间长，抗风险能力强且在市场上积累了土地、技术、品牌和政策资源等优势，具有较强的竞争力和较好的信用和声誉，更易取得投资者的信任，赢得更多外部资金的投入和支持。

固定资产比率 (Fix) 的系数为 0.191 ，影响显著为正，揭示了农林牧渔企业的土地、农机、养殖场等固定资产占比越高，融资约束 (FC) 越严重，可能因行业资产流动性差、抵押变现难，导致外部融资成本上升。

表6主效应分析结果

	fc	fc
age	-0.003^* (-0.676)	-0.003^{***} (-0.679)
size	-0.140^{***} (-5.367)	-0.143^{***} (-5.440)
lev	-0.509^{***} (-8.149)	-0.507^{***} (-8.113)
roa	-0.000 (-0.044)	0.000 (0.024)
fix	0.193^{**} (1.982)	0.195^{**} (2.002)
esg		-0.007^{***} (0.942)
controls	YES	YES
Year	YES	YES
_cons	3.786^{***} (6.594)	3.823^{***} (6.642)
N	457	457
R ²	0.674	0.775
F	25.891	23.106

*** $p < 0.01$ ”，** $p < 0.05$ ”，* $p < 0.10$

(三) 中介效应

在基准回归已验证良好的 ESG 评级能够有效缓解农林牧渔企业融资约束的基础上，采用分层回归法验证农林牧渔企业能否通过提高企业声誉缓解融资约束程度，结果如表7，考虑到篇幅表格省略了控制变量的结果。分别为 0.714 ， 0.806 ，表明模型拟合效果较好。

表7中介效应分析结果

	(1)	(2)
	rep	fc
esg	0.03^{***} (0.035)	-0.006^{***} (0.942)
rep		-0.002^{**} (-0.361)
controls	YES	YES
Year	YES	YES
_cons	-15.843^{**} (-2.404)	3.796^{***} (6.532)
N	457	457
R ²	0.714	0.806
F	27.167	20.756

*** $p < 0.01$ ”，** $p < 0.05$ ”，* $p < 0.10$

模型 (1) 控制其他变量不变，以 ESG 为自变量，以企业声誉作为因变量，得回归系数为 0.03 ，回归显著，表明良好的 ESG 评级能够提升企业声誉，企业积极履行环境、社会和治理责任能够显著增强公众形象和市场认可度。

模型 (2) 在主效应基础上加入企业声誉 REP 进行多元回归，ESG 和 REP 对 FC 影响显著为负，系数分别为 -0.006 和 -0.002 。相较于 ESG 与 FC 的主效应模型，中介模型系数有所下降，采用 Sobel 检验中介效应的稳健性，结果显示 Z 统计量为 1.285 ，在 1% 水平下显著，验证了中介效应的存在，企业声誉起部分效应作用，ESG 和声誉的协同作用进一步强化市场信任，增强投资者信心，减少信息不对称，从而降低风险溢价，使得农林牧渔企业在资本市场上获得更有利的融资条件。

(四) 内生性和稳健性检验

为验证结论的稳健性并排除潜在的内生性问题，进行内生性与稳健性检验。通过工具变量法和替换关键变量测度方式，有效控制了模型可能存在的估计偏差，确保实证结果具有统计可靠性和经济意义。

表8内生性检验结果

	(1)	(2)	(3)
	esg	Fc	Fc
lesg	0.158^{***} (3.030)	-0.004^{***} (0.244)	-0.003^{***} (0.301)
rep			-0.005 (-0.812)
Controls	YES	YES	YES
year	YES	YES	YES
_cons	3.260^{***} (16.177)	3.547^{***} (5.182)	3.492^{***} (5.073)
N	355	355	355
R ²	0.735	0.809	0.811
F	9.182	14.078	12.719

*** $p < 0.01$ ”，** $p < 0.05$ ”，* $p < 0.10$

采用工具变量法检验，选取滞后一期 ESG 评级作为工具变量 (Lesg)，再次进行多元回归，如表8。ESG 与 Lesg 的系数显著为正，认为 Lesg 可作为内生变量。模型 (2) 中 Lesg 对 Fc 的回归系数为 -0.004 ，在 1% 水平下显著为负，说明良好的 ESG 评级确实能够缓解农林牧渔企业的融资约束，与主效应结果一致，假设 H1 结果稳健。模型 (3) 中企业声誉仍在 ESG 评级对融资约束的影响中起部分中介作用。

表9稳健性检验结果

	fc	fc
esgl		-0.009^{***} (1.234)
Controls	YES	YES
year	YES	YES
_cons	3.786^{***} (6.594)	3.799^{***} (6.620)
N	457	457
R ²	0.374	0.377
F	25.891	23.219

*** $p < 0.01$ ”，** $p < 0.05$ ”，* $p < 0.10$

借鉴刘梦昕等 (2024) 更换企业 ESG 表现衡量的方式^[15]，研

究对 ESG 评级重新赋值, AAA~A 为 3, BBB~B 为 2, CCC~C 为 1, 定义为变量 ESG1, 重新进行回归分析。结果如表 9, ESG1 与 FC 指数的系数仍显著为负, 说明实证分析结果可靠。

（五）异质性检验

根据农林牧渔企业规模的划分标准, 针对企业规模 (Size) 指标将样本企业分为大规模、中规模、小规模企业, 并分别进行回归。结果如表 10。

表 10 异质性检验结果			
	Fc(大规模)	Fc(中规模)	Fc(小规模)
esg	-0.003*** (1.920)	-0.009*** (0.647)	-0.028** (-0.333)
Controls	YES	YES	YES
year	YES	YES	YES
_cons	2.604** (1.522)	4.950*** (2.836)	1.843* (1.948)
N	168	192	97
R ²	0.695	0.683	0.611
F	7.103	4.216	3.064

表 10 显示大规模企业、中规模企业和小规模企业的 ESG 评级对农林牧渔企业融资约束影响的系数分别为 -0.003、-0.009 和 -0.028, 显著为负, 表明 ESG 表现对小规模农林牧渔企业的融资约束缓解效应最为明显, 而对大规模企业的影响较小。小规模企业普遍存在更严重的信息不对称问题, ESG 评级作为重要的非财务信号, 能更有效地降低投资者与企业之间的信息壁垒; 相比之下, 大规模企业凭借其固有的规模优势和成熟的融资渠道, 对 ESG 评级的边际依赖程度相对较低。

（六）实证结果

以上实证分析验证了三个假设, 认为 ESG 绩效与农林牧渔企业融资约束呈负相关关系。企业声誉在 ESG 绩效与农林牧渔企业融资约束之间起中介作用。良好的 ESG 会提升企业声誉, 进而缓解融资约束。同时认为 ESG 对小规模企业融资约束的缓解效应相较于大规模企业更大, 因此小规模农林牧渔企业应重视 ESG 体系建设, 通过加强环境管理、履行社会责任和完善公司治理来提升市场认可度, 同时积极维护企业声誉, 以获取更优惠的融资条件, 实现可持续发展。

参考文献

[1]Cai, Y., Idris, N., & Fuzi, N. M. (2024). An Overview of Literature on ESG Performance Using Bibliometric Analysis. *International Journal of Academic Research in Business and Social Sciences*, 14(12), 1526 - 1540.

[2] 朱紫妮. ESG 表现、融资约束与企业绩效 [J]. *运筹与模糊学*, 2024, 14(6): 180-192.

[3]Khaw Y T , Amran A , Teoh P A . Factors influencing ESG performance: A bibliometric analysis, systematic literature review, and future research directions[J].*Journal of Cleaner Production*, 2024, 448141430-.

[4]Xu, S., Li, H., Chen, J., Huo, J., & Kuang, X. (2024). Sustainable competitiveness through ESG performance: An empirical study on corporate resilience. *Journal of Competitiveness*, 16(3), 53-72.

[5] 孙宇, 辛立秋. 供应链金融对中小型畜牧企业融资约束的缓解效应研究 [J]. *黑龙江畜牧兽医*, 2023, (04): 7-15+20.

[6] 朱柏洋. ESG 表现对农业上市公司绩效的影响研究 [J]. *西南大学学报 (自然科学版)*, 2023, 45(01): 110-122.

[7] 张亚连, 苏昌祥. 企业 ESG 表现对融资约束的影响研究——基于信号不对称与委托代理视角 [J]. *新疆财经*, 2023, (02): 48-57.

[8] 刘树良, 闫明丽. ESG 表现、代理成本与融资约束——基于企业生命周期的调节作用 [J]. *工业技术经济*, 2024, 43(12): 62-72.

[9] 邵鹏, 董小琦, 马冰. ESG 表现同群效应与企业融资约束缓解 [J]. *中国环境管理*, 2024, 16(04): 44-51.

[10]Andrea Bonime-Blanc. 论声誉风险, ESG 议题的环环相扣 [J]. *WTO 经济导刊*, 2015 (1): 17-17.

[11]Geng L, Cui X, Nazir R, et al. How Do CSR and Perceived Ethics Enhance Corporate Reputation and Product Innovativeness? [J]. *Economic Research-Ekonomika Is-trazivanja*, 2022, 35(1): 5131-5149.

[12]Broadstock D C, Matousek R, Meyer M, et al. Does Corporate Social Responsibility Impact Firms' Innovation Capacity? The Indirect Link Between Environmental & Social Governance Implementation and Innovation Performance [J]. *Journal of Business Research*, 2020, 119(7): 99-110.

[13] 顾雷雷, 郭建鸾, 王鸿宇. 企业社会责任、融资约束与企业金融化 [J]. *金融研究*, 2020, (02): 109-127.

[14] 管考磊, 张蕊. 企业声誉与盈余管理: 有效契约观还是寻租观 [J]. *会计研究*, 2019, (01): 59-64.

[15] 刘梦昕, 杜苇. 企业 ESG 表现对企业融资约束的影响研究 [J]. *中北大学学报 (社会科学版)*, 2025, 41(01): 136-143.

四、结论与建议

（一）结论

1. 我国农林牧渔业企业中年轻企业占比较大, 具有较高融资需求, 但财务状况不稳, 缺乏稳定现金流, 多数企业面临融资约束。企业对外部融资的依赖性较大, 缺乏多样的融资渠道; 不稳定的财务状况增加了企业的融资风险; 在融资结构上更依赖于长期债务, 对外部资金的需求不断增加。

2. ESG 绩效与农林牧渔企业融资约束间存在显著的负相关关系。提升环境管理、注重社会管理和公司规范治理能够增强企业的融资实力, 减小融资约束, 提升信誉。

3. 企业声誉在 ESG 绩效与农林牧渔企业融资约束之间起到了显著的中介作用。良好的 ESG 管理降低了经营风险, 增强了金融机构和投资者的信任, 维护了企业声誉, 融资约束较小; ESG 问题导致信用崩塌, 损害了企业声誉, 融资成本上升。

4. ESG 对小规模企业融资约束的缓解效应相较于大规模企业更大。小规模企业普遍存在更严重的信息不对称问题, ESG 评级作为重要的非财务信号, 能更有效地降低投资者与企业之间的信息壁垒; 相比之下, 大规模企业凭借其固有的规模优势和成熟的融资渠道, 对 ESG 评级的边际依赖程度相对较低。

（二）建议

从企业层面看, 强化 ESG 体系建设, 提升企业可持续竞争力。采用绿色生产技术, 减少环境污染, 提升 ESG 评级, 提高投资者信心; 加强员工福利保障, 提高供应链透明度, 提升品牌公信力; 优化治理结构, 提高财务透明度, 降低违规风险, 增强投资者信任。定期发布 ESG 报告, 参与权威评级; 建立声誉风险管理机制, 对负面事件快速响应, 制定整改方案, 维护企业信用。

从政府部门层面看, 建立 ESG 信用评价体系, 将企业 ESG 表现纳入征信系统, 对 ESG 评级高的企业给予政策激励, 对存在重大 ESG 违规的企业实施融资限制或黑名单管理。重点扶持小规模企业, 缩小 ESG 发展差距。

从金融机构层面看, 对 ESG 评级较高的企业提供更优惠的利率和更灵活的还款方式。组织面向小规模企业的 ESG 培训, 分享 ESG 融资成功经验, 为企业提供 ESG 提升咨询和融资方案设计。

化解运营困境：广州市地下综合管廊 PPP 项目 资金缺口成因与实战对策研究

陈晓坚

广州市城市建设事务中心，广东 广州 510000

DOI:10.61369/ASDS.20250700095

摘要： 本文以广州市属地下综合管廊 PPP 项目为研究对象，梳理其运营阶段出现的资金缺口现象，围绕建设期延长、管线入廊率偏低、政府补助与偿债时点错位以及部分社会资本金未足额到位等核心诱因展开剖析。基于广州市城市建设事务中心提供的运行数据与处置经验，本文提出从资金监控强化、运营效率提升、财政保障机制优化及社会资本履约督促四个维度切入可操作性对策，以期实现项目稳健运行与长期可持续发展，并为同类型基础设施 PPP 案例提供可复制、可推广的实践范式。

关键词： 地下综合管廊；PPP 模式；资金缺口；对策研究

Resolving Operational Challenges: A Study on the Causes and Practical Countermeasures for Funding Shortfalls in Guangzhou's Urban Utility Tunnel PPP Project

Chen Xiaojian

Guangzhou Urban Construction Affairs Center, Guangzhou, Guangdong 510000

Abstract： This paper examines the funding shortfalls encountered during the operational phase of Guangzhou's urban utility tunnel PPP project. It analyzes key contributing factors, including prolonged construction periods, low pipeline occupancy rates, misalignment between government subsidies and debt repayment schedules, and insufficient capital contributions from some private investors. Drawing on operational data and management experience from the Guangzhou Urban Construction Affairs Center, the study proposes actionable countermeasures across four dimensions: strengthening financial oversight, improving operational efficiency, optimizing fiscal safeguards, and enforcing private capital commitments. The aim is to ensure stable project operation and long-term sustainability while providing replicable and scalable solutions for similar infrastructure PPP projects.

Keywords： urban utility tunnel; PPP model; funding shortfall; countermeasure study

引言

地下综合管廊，被誉为城市“看不见的脊梁”，在增强城市韧性、集约开发地下空间、降低管线事故率等方面扮演着关键角色。自广州入选国家第二批地下综合管廊试点以来，便积极探索政府与社会资本合作（PPP）模式，逐步搭起“政府授权+特许经营+可行性缺口补助”的三层架构。然而，当建设期与运营期握手之际，资金缺口悄然出现，解决这个项目运转堵点，既关乎羊城此类项目的顺利推进，也能为全国同行提供镜鉴。

一、广州市属地下综合管廊 PPP 项目资金财务状况剖析

（一）项目模式架构

广州市的四个市属地下综合管廊项目均采用政府和社会资本

合作（PPP）模式进行运作。具体为，政府先通过公开招标的方式选定合适社会资本方，然后政府授权的项目实施机构会与政府资本代表按照 34%：66% 的比例共同组建专门的 PPP 项目公司。在这个架构中，资本金占项目总投资的 20%，而剩余的 80% 资金需求则由项目公司向银行等金融机构进行融资来满足。

作者简介：陈晓坚，广州市城市建设事务中心，高级会计师，研究方向：成本核算与控制、资金统筹与调配、预算管理与执行、费用征收管理（如城市基础设施配套费）、投资评审与造价审核、财务合规与审计、资金使用效率优化、征收补偿资金管理、项目融资与资金保障、财务信息化建设。

政府会授予项目公司特定年限的特许经营权，在此项目运营期间，倘若项目公司的运营收入无法足够涵盖项目建设成本（其中包含合理收益）以及运营成本（其中包含合理利润），那么政府将会依据绩效评价所得出的结论，采用向项目公司支付“可行性缺口补助”的形式来达成平衡，这一补助机制目的在于保证项目可在特许经营期结束后完整且无偿地移交给政府，以此保障项目能够顺利推进并实现公共利益^[1]。

（二）资金投入与使用现状

截至2024年12月31日，四个市属地下综合管廊项目已累计投入资金达1,742,541.36万元，占总投资估算2,164,953.00万元的80.49%，其中政府方出资123,976.20万元，社会资本方出资232,971.66万元，融资提款1,286,470.3万元，调增建设用地费62,355.29万元。不过，环城管廊、智慧城管廊和琶洲管廊社会资本方未完成全部资本金出资，分别欠缴3,537.53万元、3,875.30万元、33,52.74元。

当前资金到位的实际情况表明，项目在资金筹集以及后续管理环节存在着诸多挑战。其一，部分社会资本没有依照协议规定一次性将资本金足额注入，如此便直接使得项目公司现金流所拥有的安全垫遭受削弱，进而导致财务杠杆有所抬升。其二，这也给各方敲响了警钟，即在PPP整个周期当中，务必要对资金的流向加以严格监管，并且要保证出资承诺能够刚性地得以兑现。

（三）管线入廊收费情形

对地下综合管廊PPP项目而言，入廊费是运营现金流的核心“水龙头”。截至2024年12月31日，四条管廊的收款进度呈现显著差异：

- 环城管廊已到账4,670.98万元，尚待追缴9,464.17万元；
- 广花管廊已入账3,051.10万元，仍有5,640.35万元悬而未决；
- 智慧城管廊仅收回156.89万元，余款487.05万元未落实；
- 琶洲管廊进账50.43万元，尾款88.61万元仍待结清。

综合来看，入廊率与收费完成度均处于低位，与“管廊收费需覆盖建设成本30%”的初始目标相去甚远。缺口之下，项目公司在运营初期即背负沉重现金流压力，风险敞口随之放大。收费滞后的成因多元：管线单位对入廊成本敏感、信息沟通链路过长、协议签订及付款流程繁复等皆有影响。下一步，项目公司应建立“一对一”沟通机制，简化入廊协议及付款节点；政府端可同步推出减免、补贴等柔性激励，引导管线单位尽快入廊，以平滑资金曲线，保障项目稳健运营。

二、广州市属地下综合管廊PPP项目资金缺口成因深度探究

（一）建设期延长的连锁反应

受施工场地条件、疫情等多因素影响，各项目均出现不同程度的建设期延期。由于贷款协议是基于原定建设期签订的，项目实际未竣工而进入约定运营期时间，导致自2023年起开始按还款计划还本，且自2024年第二季度起，贷款利息不再符合资本化条

件，付息不能再使用贷款支付。广花管廊和智慧城管廊因工期延长最多，面临最大的资金压力，尽管已通过协调国开行申请贷款展期来缓解还本资金压力，但在2025 ~ 2026年的运营初期，还款压力仍将对项目公司产生持续影响。

（二）管线入廊率低与收费管理滞后

地下综合管廊的修建所需费用颇高，管线单位在入廊之后所要缴纳的各项费用，可能会超出其自行开展设计、修建以及运营管理等工作时所产生的费用，管线单位入廊的积极性便难以被调动起来。并且，项目公司没有依照PPP合同中所作的约定，在规定的时间内及时与管线单位签订入廊协议，进而收取相应的入廊费用，这更加剧了收费不足的严重程度^[2]。目前，四个市属地下综合管廊所收到的入廊费以及日常维护费，仅仅只占到了项目总建设成本的0.36%（入廊费及日常维护费合计占项目总建设成本的比例），这个比例远远低于原先预期的30%管廊收费占比。正是由于入廊率偏低且收费管理工作滞后，这才致使项目公司在运营初始阶段，就面临资金缺乏的困境，难以满足还本付息方面所需的资金需求。

（三）政府补助与贷款还本付息时间错配

PPP合同写明：管廊须在竣工验收并平稳运营满一年后，且绩效达标，才能申请首笔“可行性缺口补助”。这类公益性项目天生造血偏弱，偿债现金几乎全靠财政“输血”。可工期一延误，贷款提前进入还本付息节奏，而补助拨付仍按原节点“慢动作”，收入与债务顿时错位：银行已到收款日，财政补助却仍在路上，时间差把项目公司的流动性瞬间压到红线。

（四）部分社会资本方资本金未足额到位及其根源

环城管廊社会资本方中铁平安投资有限公司未按合同约定注入资本金3,537.53万元，智慧城管廊社会资本方因概算未全部批复，未签订PPP补充协议，未能注入资本金3,875.30万元。社会资本方作为项目投融资主体，本应负责项目资金筹措、运营管理等全过程，但个别社会资本方未能履行出资义务。环城管廊社会资本方资本金注入延迟超1年，智慧城管廊社会资本方则纠结于调整总投资的PPP补充协议未签订而忽视自身投资责任和融资风险^[3]。

表1：资金缺口成因总结表

项目	建设期延长（月）	资本金缺口（万元）	入廊费收取率（%）	政府补助错配程度（高/中/低）
环城管廊	+18	3,537.53	33.0%	高
广花管廊	+24	0	35.1%	高
智慧城管廊	+22	3,875.30	24.3%	中
琶洲管廊	+12	33.5274	36.3%	低
行业预期标准	0	0	≥70%	无错配

注：表中“入廊费收取率”的定义是指已收/应收入廊费的比率

三、广州市属地下综合管廊 PPP 项目资金缺口问题的实战性解决方案

（一）强化资金监管与风险防控体系

1. 项目实施机构应当着力构建起完备的资金监管机制，针对项目公司资金使用的实际状况予以严密的监控把控。一方面要建立起资金预警机制，实现对资金流入与流出情况的实时动态监控。具体来讲，项目公司需要每天给项目实施机构呈报资金收支报表，把预警指标设定成当资金余额低于某个特定阈值（比如未来三个月的最低资金需求额度）的时候，能够自动触发预警信号，以此来提前识别出潜在的风险所在，像是没有及时收到可行性缺口补助或者入廊费用这类情况，从而可以采取相应的应对办法^[4]。

2. 项目公司应培养专业的运营管理队伍，强化运营管理方面的各项工作，以此来对运营成本加以把控。项目实施机构需要定期针对项目公司的运营成本预算展开审查，务必要保证其具备合理性，同时还要对其成本控制措施的实际落实状况予以监督。

3. 项目实施机构需组织专家团队针对项目的各类风险展开全面评估，如市场风险、政策风险、技术风险等，并且应制定出具体的应对举措。如面对市场风险，可以制定灵活多变的入廊费定价策略，同时还要构建多元化的运营收益模式；而针对政策风险，则要和政府部门维持紧密的沟通联系，实时掌握政策方面的动态情况，进而依据这些动态去调整项目的运营策略。

（二）提升运营效率与收益水平的实战策略

实施机构需积极与主管部门做好对接工作，争取把“强制入廊”这一规定纳入到具有刚性约束力的制度当中，并且要配套制定更为细化的罚则，以此来保证相关文件一旦落地便能够切实执行。如要求所有的管线都必须进入廊道，给首批签约的单位给予“运营首年入廊费减免”以及“维护费补贴”等相关优惠举措，对于那些不履行的单位，限制其在后续的城市基础设施建设方面的投标资格。这样既能高举起重惩戒的“戒尺”，又能伸出合作的“橄榄枝”。

项目公司需搭建起“一对一”的沟通桥梁，细致摸排管线单位所存在的顾虑情况，并且为其量身打造相应的入廊方案^[5]。比如，针对那些在一次性缴费方面感到压力颇大的单位，可以推出分期付款的选项，以此来缓解其现金流较为紧张的状况。公司方面还需组建起专职的“入廊服务队”，专门负责协议的谈判事宜、签订工作以及后续的催收任务，从而确保相关款项能够及时且足额到账。

除此之外，项目公司也可尝试多元化的增收途径。如在管廊的出入口、通风井等这些较为显眼且人流量相对较大的黄金点位设置广告屏，进而向广告商收取租金。又如把闲置的舱室租借给通信运营商或者物流企业，供其存放设备或者是相关物资，还可为管线单位提供诸如监测、维保等收费增值服务。不仅如此，项目公司还可携手科研机构共同开展技术试验，既能获取科研经费，又能够在后续成果成熟之后将其转让出去。通过采取多种不同的增收举措，从而把原本单一的过路费收入成功升级成为更为

丰富的综合现金流。

（三）优化财政资金保障机制的创新举措

现行机制下，项目公司需等待“运营满一年+绩效考核”两道闸门全部开启后，方可启动可行性缺口补助申请；再叠加财政审批周期，资金往往滞后6-12个月才能到账，与银行贷款“按月还本付息”的刚性时点形成尖锐错位^[6]。

为彻底扭转这一错位，政府可在政策允许范围内，将补助拨付节奏与还贷日历“对表”：

1. 事前会商：财政、贷款银行、SPV 三方依据年度还贷计划，提前锁定补助“预付表”。

2. 季度预付：每季度首月，财政按当年应补总额的80%先行拨付，用于当期还本付息，做到“钱到点、账不拖”。

3. 年末清算：次年一季度，根据全年绩效考核结果一次性结算剩余20%差额，多退少补。

4. 资金通道：市级财政单列“管廊专项”，实行线上申报、并联审批、限时办结，所有流水实时晒单，接受社会与审计双重监督。

通过“先预拨、后清算”的闭环设计，既保障项目公司现金流不断链，又让财政资金安全可控，形成错位纠偏、风险可控、效率倍增的多赢格局。

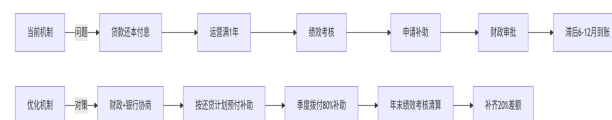


图1：财政补助拨付流程优化对比图

（四）督促社会资本方履行投融资责任的有效措施

当社会资本方未足额注入资本金时，项目实施机构需采取措施督促整改。如：组建约谈小组与其沟通，了解未注资原因及困难，要求其制定明确整改方案，包括注资时间节点、金额等，并提交书面承诺书，先行追加注资用于还本付息，该部分资金后续可转为资本金，以满足项目资金需求。同时，强化合同约定，明确社会资本方违约责任及处罚机制，如支付违约金、承担融资成本、丧失分红权等，并严格按合同执行处罚^[7]。加强对其履约情况考核，与未来项目合作机会、信用评级等挂钩，对良好履约方给予招标加分或优先考虑，对不良履约方纳入失信名单，限制其参与其他政府投资项目，从而推动社会资本方积极履行投融资责任。

四、结论与展望

地下综合管廊 PPP 模式对完善城市“里子工程”意义深远，但其运营初期暴露的资金风险不容忽视。本文以广州样本为切口，系统梳理了资金缺口的深层成因——工期拉长、入廊率低迷、补助与还贷节奏错位以及部分社会资本金不到位等，并据此提出“监管提质—效率提速—财政提效—履约提级”的闭环对策，力图在降低资金风险的同时，保障项目稳健与可持续^[8]。

面向未来，城镇化仍在加速，管廊需求将持续释放。政府与

社会资本需携手探索更灵动的合作范式：一方面可尝试夹层融资、绿色 REITS 等多元工具，撬动更多长期资金入场；另一方面，持续完善使用者付费机制，提升项目“自我造血”功能，并借助资产证券化盘活存量收益权，为后续建设输送多元“弹药”。此外，建议设立跨部门研究专班，动态跟踪政策效果，及时迭代制度设计，为地下综合管廊 PPP 项目行稳致远再添制度护栏。

参考资料

[1] 张强，梁芳. 建设项目资金缺口分析与控制 [C]//《施工技术》杂志社. 2024年全国土木工程施工技术交流会论文集（下册）. 昆明经百实业有限公司；, 2024:1151-1152. DOI:10.26914/c.cnkihy.2024.075643.

[2] 谢晓晓. PPP 模式如何填补农村生活垃圾处理资金缺口——基于两个国家级示范项目的分析 [J]. 武汉交通职业学院学报, 2020, 22(02):58-63.

[3] 张华志，黄全祥. 乡村产业振兴的信贷资金缺口分析 [J]. 财经科学, 2022, (11):58-72.

[4] 徐新. 地方政府债务风险识别、分档化解与资金缺口平衡——基于 KMV 模型对贵州数据的实证分析 [J]. 生产力研究, 2021, (11):14-23+161. DOI:10.19374/j.cnki.14-1145/f.2021.11.003.

[5] 方佳敏. 吉林省城镇化建设融资模式与融资效率研究 [M]. 中国轻工业出版社 :201706.177.[6] 龚澧. 湖南省公路建设项目融资方式及对策研究 [D]. 中南大学, 2012.

[7] 潘睿. 基于系统动力学的高速公路 PPP 项目资金缺口政府补助测算研究 [D]. 广西大学, 2023. DOI: 10.27034/d.cnki.ggxiu.2023.001803.

[8] 戴皇. 基于 PPP 模式的高速公路项目平台公司投资效益分析 [J]. 财经界, 2025, (11):24-26. DOI:10.19887/j.cnki.cn11-4098/f.2025.11.036.

加强基层辅助调查员管理的几点思考

曹秀平

河南省洛阳市城乡一体化示范区统计调查服务中心，河南 洛阳 471000

DOI:10.61369/ASDS.2025070010

摘 要： 基层辅助调查员是统计调查工作落地见效的前沿执行者和重要力量，其履职能力与工作质量直接影响调查数据的真实性、准确性和时效性。当前，基层辅助调查员队伍在人员稳定性、业务技能、法制意识、激励机制等方面仍然存在着流动性大、业务水平不高、依法统计意识淡薄、工作积极性不高等问题，制约了统计调查工作的高质量开展。本文结合工作实践，从优化选聘机制、创新培训方式、完善激励机制、强化法律意识方面，探讨加强基层辅助调查员规范管理的方法路径，旨在提升队伍整体素质，夯实基层基础力量，为高质量统计数据提供有力保障。

关 键 词： 基层辅助调查员；管理优化；队伍建设；工作效能；制度完善

Reflections on Strengthening the Management of Grassroots Auxiliary Investigators

Cao Xiuping

Statistical Survey Service Center of Luoyang Urban-Rural Integration Demonstration Zone, Luoyang, Henan 471000

Abstract： Grassroots auxiliary investigators are the frontline implementers and crucial forces for the effective implementation of statistical survey work. Their ability to perform duties and work quality directly affect the authenticity, accuracy and timeliness of survey data. At present, the team of grassroots auxiliary investigators still faces problems such as high mobility, low professional proficiency, weak awareness of statistics in accordance with the law, and low work enthusiasm in aspects of personnel stability, professional skills, legal awareness and incentive mechanisms. These issues have restricted the high-quality development of statistical survey work. Combining work practice, this paper explores the methods and paths to strengthen the standardized management of grassroots auxiliary investigators from the perspectives of optimizing the recruitment mechanism, innovating training methods, improving incentive mechanisms and enhancing legal awareness. The purpose is to improve the overall quality of the team, consolidate the grassroots foundation, and provide strong support for high-quality statistical data.

Keywords： grassroots auxiliary investigators; management optimization; team construction; work efficiency; system improvement

引言

辅助调查员在统计调查对象与基层统计机构中间发挥着桥梁和纽带作用，是源头数据采集和基层数据质量控制的第一线力量^[1]，其业务能力的高低，直接影响到统计数据的真实性、准确性、完整性和及时性。随着统计管理体制改革的不断深入和新时代下统计工作面临的诸多问题和挑战，统计调查工作对辅助调查员的调查能力和统计素养提出更高要求。辅助调查员职能作用发挥与否，将直接影响到统计源头数据质量，统计数据的公信力也与其密不可分。本文从实际出发，结合多年基层统计工作经验，分析当前辅助调查员管理中存在的问题和面临的挑战，提出进一步完善和优化辅调员的思路和建议，助推其职能作用的充分发挥。

一、当前基层辅助调查员管理难点

随着调查任务的不断增加和统计调查手段的更新，基层统计力量与统计调查需求之间的矛盾日益显现。辅助调查员业务能力与职业素养制约着基层统计质量和统计服务水平的提升和优化^[2]。

（一）基层辅助调查人员的稳定性不足

一是基层辅助调查员大部分都是兼职，工作时间受限^[3]。聘用的辅调员以村（社区）人员为主，需要兼顾自家农活、村（社区）日常事务、乡镇（街道办）工作任务等诸多方面，难以有更多精力和时间投入到统计调查工作中，达不到新时代背景下统计

对基层基础数据规范化和高质量的要求。二是村级辅助调查员流动性大，人员稳定性不足^[4]。调查补贴较低，是造成人员流失的主要原因。辅助调查员的调查补贴每月或者每季度按补助形式发放，发放标准每月100到200多元不等，相对较低调查补贴与统计繁杂的工作任务不相匹配，造成一部分辅调员缺少工作主动性和积极性，有其他较好的工作机会时，会首先放弃统计调查工作，给统计工作持续性开展调查增加了难度。样本轮换、村委换届也是人员不稳定的重要原因。一个业务熟练的辅助调查员需要不同专业人员进行多次业务培训和长时间的实际操作演练，但由于统计制度需要按周期对调查样本进行轮换，而成熟的辅调员不愿意到其它区域工作，给高质量开展统计调查工作带来影响。村委换届后，一些未当选人员也会出于面子等方面原因而不再继续开展统计调查工作。

（二）业务能力参差不齐，缺乏有效的监督手段^[5]

受工作任务多和调查补贴较低因素的影响，在选聘辅助调查员时对参加选聘人员的学历、年龄、工作经历要求的相对宽松，更多的是注重熟悉情况，基层经验多的村干部或者是街道（社区）人员担任。选聘出的辅助调查员文化程度普遍较低，年龄较大，电子信息技术水平不高，对新知识、新技术接受较慢。作为一名合格的辅助调查员，不仅要求熟悉调查制度、方案、流程和指标含义，熟练操作各类调查小程序，还需要具备一定的访户技巧和亲和力。受年龄、文化程度等因素影响，辅助调查员业务能力参差不齐，部分辅助调查员对开展统计调查的意义、目的理解不透彻，对业务知识掌握不全面、不准确，不能给予调查对象全面、正确的指导，导致基层调查数据失真。县区级统计机构对于辅助调查员缺乏科学、合理、操作性强的考核办法和有效的监督措施，辅助调查员的职责不能有效得到发挥，也是造成基层统计数据质量不高的重要原因之一。

（三）统计法制意识不强，依法统计观念需进一步加强

《统计法》《统计法实施条例》等法律法规明确规定了统计人员和调查对象的权利和义务，规定了不按规定报送统计报表、编造虚假统计数据的法律后果。在实际统计调查工作中，一是辅助调查员对相关的法律条款了解不透彻，不能利用法律武器有效推动各项统计调查工作的开展；二是基层统计执法力量薄弱，不具备独立开展统计执法检查的条件，主动进行执法检查的次数较少，对统计违法现象达不到有效的震慑作用；三是碍于情面不愿意“上纲上线”，解决问题侧重于靠人情、靠协调，统计法律法规为数据质量保驾护航的作用没有得到有效发挥。

（四）激励机制不完善，工作积极性难以充分调动^[6]

一是在日常对辅助调查员的管理中，主要侧重于业务知识培训和实操技能训练，对个人业务水平高低、工作质量好坏没有一个系统的、可操作性强的考核方案。对工作认真负责的人员没有有效的激励措施，对工作敷衍了事的人员也没有制定相应的惩罚措施，调查补贴都是一个标准，个人价值难以在统计工作中充分体现，造成辅助调查员在工作中缺乏主动性、积极性。二是先进模范示范带动作用没有得到有效发挥。精神激励匮乏，缺乏对工作成果的认可与表彰，长期下来会使辅调员产生厌战心理，工作

懈怠，工作积极性得不到正向强化。

二、做好辅助调查员管理的思考和建议

（一）优化人员选聘机制，夯实队伍基础^[7]

一是进一步明确选聘标准，制定科学合理的人员选聘标准，对责任心强、文化水平高、沟通能力强、熟悉当地情况的优先选择。例如，在农村可以选择返乡创业人员、年轻村干部、待业大学生等，他们对当地的住户情况、农业种植、养殖户较为熟悉，便于开展农村住户调查、劳动力、农产量等统计调查工作。在城镇地区，可吸纳有一定工作经验、时间相对充裕的社区工作者，确保能高质量完成城镇住户调查、人口抽样调查、工业品价格调查等任务。同时，建立严格的选聘流程，联合人社部门通过发布选聘公告，扩大知晓范围，吸引更多优秀的人员参加选聘，采取资格审查、面试考核、电子化小程序实操技能考试等环节，确保选聘人员能够胜任统计调查工作。

（二）创新培训方式，提升业务能力

一是开展分层分类培训。根据辅助调查员的岗位需求和业务水平，分层分类开展“接地气”的业务培训^[8]。对于新入职人员，开展统计法律法规、诚信统计职业道德培训、统计基础业务知识和调查技能的系统性培训，包括调查方案解读、数据采集上报流程讲解、问卷填写规范、APP操作流程、指标解释、统计原则等内容，帮助其快速掌握工作要领。对于经验丰富的老调查员，定期开展统计工作经验交流研讨会，分享先进工作经验和好的工作方法，充分发挥先进工作者的模范带头作用；探讨解决调查工作中遇到的复杂问题，不断提升业务能力和统计素养。二是探索创新培训方式。改变传统集中培训模式，充分利用网络新媒体传播优势，采用线上线下相结合、理论实践相结合的多元化培训方式，通过网络培训、小视频教学、案例分析、模拟调查等形式，增加培训的实用性和操作性。

（三）完善考核激励机制，激发工作动力

一是建立完善科学考核制度。结合工作实际，制定明确、可量化的考核指标体系，如调查数据质量、访户次数、差错率、上报及时性、基础工作规范性等方面对辅助调查员进行全面考核。形成有效的激励约束机制，充分调动辅助调查员工作积极性。二是优化激励措施，健全奖励机制。在确保辅助调查员补贴及时足额发入的前提下，建立与工作强度、工作质量匹配的薪酬增长机制。设立绩效资金，对表现优秀、工作成绩突出的辅助调查员给予奖励。在精神激励方面，定期开展考核评估，如在每个年度结束后开展优秀调查员评选，对表现优秀人员发文表彰、发放荣誉证书或者实物奖励，增加其职业荣誉感和成就感。同时，加强对调查员的人文关怀，关心他们的工作和生活，帮助解决实际困难，营造良好的工作氛围。

（四）强化法律意识，提升依法统计观念

一是充分利用统计开放日、12.4国家宪法日、12.8《统计法》宣传日活动契机，通过制作普法展板、海报，发放宣传手册，制作普法小视频等，开展多种多样的普法宣传活动，营造浓厚的统

计法治氛围。二是结合统计年报会、业务培训会，认真解读《统计法》《统计法实施条例》，明确辅助调查员的责任和义务，使依法统计调查观念入脑入心。通过播放统计违法案例警示教育片，深入浅出的对统计违法现象进行剖析，结合工作日常实际，对统计违法现象、弄虚作假行为开展学习和举例，增加对统计工作的敬畏之心和依法统计意识，确保统计调查源头数据真实、可靠。

参考文献

[1] 加强住户调查辅助调查员管理的思考 - 统计信息 [J]. 南平市人民政府 ,2018-09-19.

[2] 浅议加强基层辅助调查员管理的几点思考 - 统计制度方法研究 [J]. 国家统计局辽宁调查总队 ,2020年第9期 .

[3] 关于住户调查辅助调查员管理的几点思考 - 调查研究 [J]. 国家统计局洛阳调查队 2023-10-31.

[4] 构建“三项机制”加强住户调查辅助调查员队伍建设的实践与思考 - 制度方法研讨 [J], 国家统计局湖州调查队 ,2024-09-03.

[5] 新形势下关于加强基层辅助调查员队伍建设的几点思考 - 调查分析 [J]. 国家统计局赣州调查队 ,2021-09-17.

[6] 加强住户调查辅助调查员管理的思考 - 统计信息 [J]. 南平市人民政府 ,2018-09-19.

[7] 区统计局：倾力打造高质量住户调查辅助调查员队伍 [J]. 广元市朝天区人民政府 ,2025-04-14

[8] 贵定县统计局关于如何选配、管理、培训、使用好基层辅助调查员 [J]. 贵定县人民政府 ,2023-11-20.

政府投资项目跟踪审计与投资效益提升的协同机制探索

简结倩

佛山市顺德区审计发展中心，广东 佛山 528300

DOI:10.61369/ASDS.2025070011

摘 要： 随着政府投资项目规模的持续扩大，提升资金使用效益与加强全过程监督成为财政管理的关键目标。跟踪审计作为动态监督的重要手段，在推动投资项目规范实施、提升投资效益方面发挥着日益突出的作用。本文系统分析了跟踪审计与投资效益提升之间的协同关系，梳理了当前区级政府投资项目审计工作中存在的资源配置、协同监管、绩效评估等方面问题，并提出了强化审计资源整合、健全跨部门协同机制与完善绩效审计体系的实践路径，通过典型案例分析进一步验证路径有效性，旨在为优化政府投资项目管理提供可行参考。

关 键 词： 政府投资项目；跟踪审计；投资效益；协同机制

Exploration of the Synergistic Mechanism between Government Project Tracking Audit and Investment Efficiency Enhancement

Jian Jieqian

Audit Development Center of Shunde District, Foshan, Guangdong 528300

Abstract： With the continuous expansion of the scale of government investment projects, enhancing the efficiency of fund utilization and strengthening the supervision throughout the entire process have become key goals in fiscal management. As an important means of dynamic supervision, tracking audit plays an increasingly prominent role in promoting the standardized implementation of investment projects and improving investment efficiency. This paper systematically analyzes the synergetic relationship between tracking audit and the improvement of investment efficiency, sorts out the problems existing in the current audit work of district-level government investment projects in terms of resource allocation, collaborative supervision, and performance evaluation, and proposes practical paths to strengthen the integration of audit resources, improve cross-departmental collaboration mechanisms, and perfect the performance audit system. Through the analysis of typical cases, the effectiveness of the paths is further verified, aiming to provide feasible references for optimizing the management of government investment projects.

Keywords： government investment projects; tracking audits; investment efficiency; collaborative mechanism

引言

政府投资项目是推动我国经济社会发展的重要手段，具有规模大、覆盖面广等特点。但在实际运行过程中，仍然存在着投资效益低下，资源浪费，决策失误等诸多问题，制约着项目的可持续发展。跟踪审计是一种重要的全过程监管方法，它能及时发现风险隐患，纠正偏差，优化管理，逐步成为提高投资绩效的重要环节。构建跟踪审计与投资效益提升的协同机制，实现审计监督与项目治理的有机结合，实现从“事后问责”到“过程控制”“价值提升”的转变，增强政府投资项目的透明度与问责制，提高公共资金的使用效率，为高质量发展提供可靠的保证。

一、跟踪审计与投资效益提升的协同关系分析

跟踪审计作为政府投资项目管理的重要监督工具，通过对项目执行全过程进行动态监测和评估，可以对资金使用、项目进

度、质量控制等方面出现的异常和风险进行及时发现和揭示，推动完善风险预防机制，优化管理过程，为提高投资效益提供强有力的保证^[1]。提高投资效益，既要靠科学合理的决策，又要有高效率的资源分配，更要靠全过程的规范化管理与风险管控。二者的

作者简介：简结倩（1987-），女，本科，中级会计师，研究方向：投资管理。

协同作用体现在审计过程中对投资执行情况进行实时反馈，帮助发现管理漏洞、效益瓶颈，促使有关部门及时采取纠正措施，保证项目目标达成，资源利用率最大化。同时，投资效益的提高反过来又能给审计工作提供明确的评价准则和指导，增强审计结果的针对性、实效性，推动审计功能由单纯的监督向价值创造的转变^[2]。引入信息技术，通过数据共享、智能分析等手段，将审计与工程管理深度结合，提升决策科学性与执行透明度，形成动态闭环协同机制，促进政府投资项目效益持续优化，提高公共资源使用效率。

二、政府投资项目跟踪审计的现状分析

当前，区级政府投资项目跟踪审计已呈现出多元化发展态势，同时也面临着一些挑战。在制度层面上，随着我国《审计法》和地方性法规的不断完善，跟踪审计的法定地位逐渐确立，部分区级审计机构已经尝试把审计节点嵌入到工程全生命周期中，涵盖项目立项、招投标、建设、竣工等各个环节，初步实现了由传统事后审计向事前、事中、事后全过程动态监管的转变^[3]。然而，受区域审计资源有限和项目复杂性的矛盾制约，审计深度和广度不平衡，特别是在民生工程和基础设施等重大工程中，由于缺乏专业人才，风险预警效果不理想。技术应用上，发达地区的区县审计机关逐步开展大数据分析模式的数字化审计工具，而大部分地区仍然依靠人工核查，数据孤岛问题制约着审计效率的提高。在协同机制方面，审计部门与发改、财政、住房等部门之间的联动还停留在个案层面，信息共享平台建设相对滞后，制约了监管合力的发挥^[4]。此外，部分地区还存在“重程序、轻绩效”的倾向，更多地关注资金使用效率和政策目标的实现。未来，需要通过优化资源配置、加强技术赋能、完善跨部门协作等途径，推动区内跟踪审计由“合规性”向“价值性”转变。

三、政府投资项目投资效益管理存在的问题

（一）审计资源配置不足

区级政府投资项目跟踪审计在资源分配上面临显著瓶颈，主要表现为审计力量薄弱、专业结构单一及技术手段滞后。在人力资源方面，区级审计机构普遍存在编制紧张、专业人才匮乏的问题，特别是缺乏具有工程造价、信息技术、绩效评价等复合型人才，难以满足日趋复杂的政府投资项目管理需求。在审计任务繁重的情况下，一些地区仍然依靠传统的抽样审计方式，很难实现对重大工程的全过程和精细化监管。技术层面上，虽然部分发达地区已善用数字化审计工具，但大部分地区仍以人工核查为主，缺少智能化审计体系的支持，导致审计效率低，难以发现隐蔽资金浪费、合同履约风险等问题^[5]。另外，由于审计经费有限，部分地区审计机构很难引入第三方专业机构进行辅助审计，这对审计的覆盖范围和准确性造成了更大的限制，导致一些民生工程和重大建设工程的关键环节没有得到有效的监管，影响到对投资效益的客观评估。

（二）跨部门协同机制不畅

区级政府投资项目管理涉及发改、财政、住建、审计等多个部门，而当前各部门之间的信息共享和协作机制还没有建立起来，这就使得跟踪审计很难起到系统的监督作用。一方面，目前审计机关和项目管理部门之间的数据交互仍然是被动的，缺少实时共享的平台，导致审计介入相对滞后，很难及时发现和纠正工程建设中存在的资金挪用、超概算等问题。另一方面，部分区县存在着“条块分割”的现象，审计建议整改工作依赖于被审计者的自觉性，缺乏有效的跨部门监督机制，使得审计结果无法充分发挥作用。比如，在招投标监管和合同履行等关键环节，审计机关和行业主管部门之间存在着监督标准不统一，甚至出现了监管盲区，导致一些项目在执行过程中出现了偏差，导致资金使用效率下降^[6]。另外，绩效评估数据分散于不同主体（如财政、项目等）手中，使得审计部门无法获得完整的信息，不利于对投资综合效益的准确判断。

（三）绩效审计体系不完善

当前，区级政府投资项目跟踪审计主要以过程合规性审核为主，缺乏对绩效指标（如经济性、效率性、效益性）的关注，很难全面反映项目的实际效益。在审计内容方面，部分区县侧重于招投标合规性、资金拨付规范性、工程质量优劣等，缺乏对长周期指标（如建后的社会经济效益、可持续发展等）的系统性评价。如一些民生项目竣工后，由于缺乏后续的维护管理而处于闲置状态，而审计环节又没有提前预警，造成了资金的浪费。在评价标准上，尚未建立起科学统一的绩效审计指标体系，特别是未将公益项目（如教育、医疗等）的社会效益定量化，导致审计结论偏重于工程合规性而忽略政策目标。另外，受数据获取能力的限制，部分区级审计机构没有将宏观经济政策和区域发展规划结合起来进行综合分析，导致审计建议缺乏前瞻性，很难实现政府投资项目整体效益的根本性提升。

四、构建政府投资项目跟踪审计与效益提升的实践路径

（一）强化审计资源整合，提升跟踪审计专业化水平

要提高跟踪审计的专业化水平，就必须从组织结构、人才储备等方面进行整合。在组织层次上，要建立科学、合理的审计资源统筹机制，打破条块分割、部门壁垒，使审计力量得到统一配置。可以在区级层面成立“重大项目审计协调中心”，对涉及不同行业、不同部门的重点项目进行统筹安排，集中优势资源，提升整体审计效率和深度。在人力资源配置方面，要加快建设专业化的复合型审计人才队伍，促进工程造价、工程管理、预算执行、数字信息和绩效评价等方面的专家共同参与，解决传统审计师专业覆盖不足的问题。通过引入社会审计力量，如依法依规委托具有资格的第三方审计机构，提高重大工程审计的专业性和广度。与此同时，审计的技术能力也要同步提升，要加快数字化转型的速度，将智能分析平台引入进来，运用区块链、AI技术构建审计信息系统，实现对合同履行、支付凭证、招投标过程等关键

节点的自动化抓取和实时监测，形成一个全流程、多节点、无死角的审计跟踪机制^[7]。结合业务系统的数据整合能力，建立动态风险预警模型，及时识别和干预可能出现的异常资金流和超预算风险。通过健全审计组织机构，优化资源配置，强化技术支持，促进区级审计机构由传统审计向全流程专业审计转变，为保障政府投资项目效益提供有力支持。

（二）健全跨部门协同监管机制，增强审计监督合力

由于政府投资项目资金来源多，实施周期长，参与主体多，单一的审计机构很难做到面面俱到。为此，必须建立完善的跨部门协同监管机制，强化审计监督的整体合力。在现有的电子政务平台基础上，区政府可以建立一个统一的政府投资项目监管信息系统，将项目从立项到审批，设计，施工，验收，运营等各个阶段的数据进行集中归集与动态更新。在信息交互机制方面，要将审计部门和发改委、财政局、住建局等重点监管机构的信息隔离开来，保证项目规划、资金拨付、施工进度、合同履行情况等关键节点信息的实时共享，提升审计部门对工程全寿命周期的感知能力。在具体的协同方法方面，要明确各部门之间的职责界限，形成闭环的协同工作流程，从发现的问题到整改的再到反馈的全过程。如对审计发现的资金违规支出，由财政部门牵头进行整改；对不规范的招投标管理，由建设部门或行业主管部门负责纠正；审计部全程跟踪整改效果，形成“发现—交办—落实—反馈—再监督”闭环机制。另外，在审计实施层次上，可推行“联合审计+联合检查”模式，针对重大工程项目（如工程变更、重大设备购置等）实施协同监管，由多个部门组成联合工作组，统一检查标准及处理规范，避免因职能重叠或空白而导致责任推诿和失效。

（三）完善绩效审计体系，推动投资效益全面评估

绩效审计不仅关注资金使用的合规合法问题，更聚焦于资金配置是否合理、项目目标是否实现、项目运营是否可持续等关键维度。在审计内容上，要把评价内容拓展到整个工程的全过程，包括施工前的可行性论证、施工过程中的预算控制和质量监督、竣工后的运营效益和社会效益。特别是在教育、医疗、生态环保等公益性、民生类项目中，要结合行业特点，建立“服务覆盖率”“群众满意度”“运营成本控制率”等可量化、可比较的绩效指标体系，从而对项目实施效果进行精确评价。在制定标准时，要充分吸收上级审计机关颁布的绩效审计准则和行业参照准则，根据当地经济发展水平、项目类型，实行分级管理，避免“一刀切”造成的评价失真^[8]。与此同时，在绩效审计过程中，还需要加强数据分析和系统支持能力的建设，将大数据挖掘和可视分析技术引入到项目的资金流、材料流和信息流中，对项目的资金流、材料流和信息流进行集成分析，发现投资回报率不高、资源配置过多的潜在问题，并提出切实可行的优化方案。通过将审计结果与政府决策机制紧密结合，把审计结果作为预算安排、项目审批及后续投入的重要依据，促进审计结果有效地转化为管理成果，使财政资金的投入产出最大化和社会公共价值不断提高。

五、典型案例分析

（一）案例概况

某区政府提出以水美城、以水兴城的水环境治理战略，以联围系统治理模式，有效解决联围内的水污染防治。2021—2024年率先启动第一批、第二批围水治水的重大工程项，包括水污染防治和水环境整治工程。项目建设总投资概算约187亿元，计划建设周期2021—2025年，涉及全区12个堤围，合计9个水体综合整治项目。该区审计局2023年至2024年对第一批、第二批工程项目进行跟踪审计，覆盖项目审批、招投标、建设、资金使用等各个环节，重点揭示项目统筹规划谋划，工程建设管理及治水成效中的问题，推动完善水体综合整治政策措施，提升治水实效。

审计人员践行投资审计“三个转变”，以“政策—项目—资金”为主线，运用“5W1H”等多角度分析，利用大数据分析方式方法，揭示该项目建设过程中问题的本质和症结。例如审计人员抽查部分勘察钻孔岩芯样，并对现场拍摄图片进行对碰，发现部分岩芯样本重复图片，存在弄虚作假的疑点。审计人员随即扩大核查范围，但面临初步勘察报告超1万张图片，数量巨大的问题，审计人员则利用python编程语言与GPS定位系统“双线嵌入”的手段，抓取定位信息，导入勘察平面设置CAD图对比，结合走访现场确定定位信息，成功筛选出1000多个钻孔出现偏差，揭示勘察报告的准确性存疑，影响后续设计的有效性的问题。此外，审计人员还发现初步设计支护方案设计优化不彻底，导致增大投资；施工图设计调整滞后，5.5亿元结存资金未投入使用；未按规定及时收回共管资金利息等问题。

审计人员提出审后整改处理意见并出具审计报告，要求被审计单位根据审计问题整改类型的时限开展整改工作。被审计单位对审计报告提出的问题分析，按照审计整改意见对审计发现的问题落实整改，并向审计人员提交整改报告。随后审计人员开展整改“回头看”，为避免被审计单位事前作出补救措施，审计人员通过突击检查方式，现场随机选取部分钻孔位置，核查是否有敷衍整改情况。核查发现，根据清单中钻孔的定位信息，审计人员现场却没找到部分钻孔，被审计单位提供的钻孔清单中部分钻孔存在虚增情况。通过质询，审计人员发现被审计单位以整改范围面广，人手不足为由，直接让第三方开展整改，期间未监督核查第三方工作开展情况，也未对第三方出具的报告进行核实，仅以第三方出具的整改情况书作为该问题的整改，从而揭示被审计单位未尽整改主体责任，整改监督机制失效，存在敷衍整改。审计人员通过发出督办函，纳入绩效考核等方式，压实被审计单位整改主体责任，督促被审计单位全面核查第三方整改情况，对确实存在虚增的及时核减相应的服务费，并对勘察单位作严肃处理，加强项目管控力度。

（二）审计介入与投资效益协同效果分析

审计介入政府投资项目建设过程，不仅有效揭示建设过程中的关键性问题，还通过发挥审计整改监督职能，督促被审计单位严格落实审计整改，将审计整改成果有效转化，运用到后续治水工程当中，最终达到提高整体项目投资效益的效果：一是增收节

支。对审计发现虚增投资成本、多计服务费等问题，被审计单位通过优化设计，重新核算，累计节约投资成本1.78亿元，核减多计服务费708万元；被审计单位收回审计提出应收未收共管账户利息，并进一步与施工单位签订补充协议，明确利息的归属，要求后续每季度按时上缴归属于被审计单位的利息，累计已收取利息1600万元。二是推动健全机制。被审计单位制定设计优化导则，为第二批工程项目设计方案编制提供有力支撑，提高项目投资效益；根据审计发现采购活动中出现的漏洞，完善内部采购活动制度，堵塞漏洞，有效控制日后采购活动的成本，减低项目投资成本。三是加强项目管控力度。借助审计整改之“手”，被审计单位严格落实项目管控主体责任，开展全面自查，严肃查处弄虚作假虚增费用的行为，切实维护政府投资利益。（见表1）

表1 该项目跟踪审计关键效益指标对比表（单位：亿元）

指标维度	审计前	审计整改后	变动率
节约投资成本	0	1.78	100%
资金使用效益	5.5	1.2	-78.18%
采购服务成本	1.2	1.13	-5.8%
共管资金利息收益	0	0.16	100%

（三）建议与启示

该案例验证跟踪审计在投资控制中的直接价值，更揭示出机制协同的深层需求。建议加强对政府投资项目审计管理，将“审

计手”从“事后抓”向“事前管”的方式转变，并善用大数据信息技术，以边建边审，审改结合方式实现项目全过程闭合循环管理。在组织层次上，推行“审计督导+专家组”模式，将复杂难解问题引入第三方复核机制，以消除监管盲区。案例启示：项目跟踪审计需要由单点纠偏向系统性治理转变，及时发现工程建设中的薄弱环节，压实建设方监督责任，实时整改，强化执行，建立评价体系，有效防范工程风险和规范资金使用，形成“发现问题—优化改进—效益提升”的良性循环。

六、结语

跟踪审计对提高政府投资项目管理的透明度，提高资金的使用效率有很大的促进作用。通过对审计资源进行优化分配，加强部门协同监督，健全绩效审计制度，可以有效地打通监督链，提高审计工作的针对性和时效性，从而推动投资效益的系统性提高。实践证明，科学地嵌入审计机制，可以有效地规避管理盲区，实现项目全过程的可控、可评估和优化。未来应加快推进智能审计技术的应用，建立动态风险预警机制，促进政府投资项目的事后补救向事前防范、过程控制转变，实现政府投资项目高质量发展的目标。

参考文献

[1] 陈金东. 法治思维视角下政府投资项目全过程跟踪审计绩效提升研究 [J]. 国际商务财会, 2024, (24): 52-55+65.
[2] 邓玉艳, 罗莎, 薛洁, 等. 全过程跟踪审计在政府投资 EPC 项目中的应用分析 [J]. 会计师, 2024, (21): 78-80.
[3] 刘爱利. 政府投资项目跟踪审计的优化探讨 [J]. 理财, 2024, (09): 24-25.
[4] 王猛. 基于协同理论的政府投资项目跟踪审计模式 [J]. 中国经贸导刊, 2024, (10): 85-87.
[5] 褚业娴. 经济法视角下政府投资项目全过程跟踪审计研究 [J]. 南昌师范学院学报, 2024, 45(01): 129-133.
[6] 周轶. 浅析政府投资项目全过程跟踪审计路径优化 [J]. 建设监理, 2023, (12): 59-61.
[7] 刘华生. 遂宁市船山区政府投资项目跟踪审计研究 [D]. 四川大学, 2023.
[8] 黄涛. 政府投资项目跟踪审计质量管理研究 [D]. 云南大学, 2023.

大湾区经管类概率统计课程的就业与升学导向教改研究

刘超, 彭舒心, 林诗敏, 张钰笛^{*}
深圳大学 经济学院, 广东 深圳 518000
DOI:10.61369/ASDS.2025070012

摘 要 : 粤港澳大湾区建设作为国家重大发展战略, 伴随其产业持续转型与升级, 对兼具理论知识与实践能力的复合型人才需求日益迫切。概率统计课程作为经管类专业的核心基础课程, 在培养学生数据分析能力方面发挥着重要作用, 然而高校课堂现行教学模式与大湾区当前发展需求存在显著错位。本文运用文献分析法与问卷调查法, 聚焦概率统计课程教学体系中存在的问题, 并提出课程改进建议。研究表明, 课程教学的主要问题体现在三个方面: 一是课程内容与学生就业及升学的双重需求脱节; 二是教学内容与大湾区特色经济场景的关联薄弱; 三是传统教学方法的学习成效欠佳。基于此, 本文构建了兼顾就业与升学双重需求、融合大湾区特色场景、关注教学方法与考核方式且更具过程关怀的课程改革框架。该框架旨在为大湾区经管类人才培养提供理论参照与实践路径, 进而更好地服务于大湾区建设发展战略。

关 键 词 : 粤港澳大湾区; 概率统计; 教学改革; 人才培养

Research on Employment and Further Study-Oriented Teaching Reform of Probability and Statistics Courses for Economics and Management Majors in the Greater Bay Area

Liu Chao, Peng Shuxin, Lin Shimin, Zhang Yudi^{*}
College of Economics, Shenzhen University, Shenzhen, Guangdong 518000

Abstract : As a major national development strategy, the construction of the Guangdong-Hong Kong-Macau Greater Bay Area, along with its continuous industrial transformation and upgrading, has an increasingly urgent demand for interdisciplinary talents who possess both theoretical knowledge and practical capabilities. Probability Theory and Statistics, as core foundational courses for economics and management majors, play a significant role in cultivating students' data analysis abilities. However, the current teaching mode in university is significantly misaligned with the current development needs of the Greater Bay Area. This paper adopts literature analysis and questionnaire survey methods to focus on the problems existing in the teaching system of probability and statistics courses and puts forward suggestions for curriculum improvement. The research shows that the main problems in curriculum teaching are reflected in three aspects: first, the curriculum content is disconnected from the dual needs of students' employment and further study; second, the teaching content has a weak connection with the characteristic economic scenarios of the Greater Bay Area; third, the learning effect of traditional teaching methods is not satisfactory. Based on this, this paper constructs a curriculum reform framework that takes into account the dual needs of employment and further study, integrates the characteristic scenarios of the Greater Bay Area, pays attention to teaching methods and assessment methods, and is more caring about the process. This framework aims to provide theoretical reference and practical paths for the cultivation of economics and management talents in the Greater Bay Area, thereby better serving the construction and development strategy of the Greater Bay Area.

Keywords : Guangdong-Hong Kong-Macau Greater Bay Area; probability and statistics; teaching reform; talent cultivation

基金项目: 深圳大学本科教学改革研究项目(一般项目): 经管类概率统计课程数字化教学方法创新策略与应用实践研究(编号: JG2025070), 主持人: 刘超。
作者简介:
刘超, 深圳大学经济学院, 助理教授, 硕士生导师, 研究方向: 函数型数据分析、异常点检测;
彭舒心, 深圳大学经济学院, 本科生, 研究方向: 会计学、时空数据分析;
林诗敏, 深圳大学经济学院, 硕士研究生, 研究方向: 函数型数据分析。
通讯作者: 张钰笛, 深圳大学经济学院, 硕士研究生, 研究方向: 统计建模, 时空数据分析; 邮箱: 2410125003@mails.szu.edu.cn。

引言

概率论与数理统计课程作为揭示随机现象统计规律的基础性学科，不仅是经管类专业知识体系的核心组成部分，更是培养学生数据思维、数据分析与决策能力的关键环节，为学生理解经济运行规律、解决实际商业问题提供了重要的理论与方法支撑^[1]。概率统计课程由概率论和数理统计两部分组成，主要包含随机变量及其分布、随机变量的数字特征、参数估计、假设检验、方差分析等内容。随着数据逐渐成为经济社会发展的核心生产要素，这门课程在经管类人才培养过程中的价值也愈发突出。

粤港澳大湾区位于中国南部，由深圳、广州、珠海等9座城市与香港、澳门两个特别行政区组成，是中国开放程度最高、经济活力最强的区域之一。粤港澳大湾区是国家重大发展战略的核心区域，集聚众多优秀企业，是高端人才实现价值的重要平台。高质量人才是社会进步的关键支撑，也是企业发展和生存所必须的核心生产资料。然而，当前概率统计人才培养与大湾区发展需求存在不相适应的问题。基于此，本文立足于粤港澳大湾区发展对复合型数据人才的迫切需求，探索经管类专业概率统计课程教学困境的改革路径。本文首先阐述研究的背景与意义，接着指出当前课程中存在的问题，并提出一套体系化的课程改进建议，最后对全文进行总结。

一、研究背景及研究意义

粤港澳大湾区（以下简称“大湾区”）的建设是新时代国家改革开放背景下的重大发展战略，该战略的目标是把大湾区打造成为一个充满活力的世界级城市群，并将其打造为具有全球影响力的国际科技创新中心。2019年中共中央、国务院印发《粤港澳大湾区发展规划纲要》（以下简称“《规划纲要》”），明确提出要“加快发展现代服务业，大力发展特色金融产业，加强金融科技载体建设”，推动形成“金融+科技+制造”的复合产业生态。

我们正处在一个由数据、算法和智能技术定义的全新时代。数据不再是冰冷的数字，而是关键生产要素。在此背景下，大湾区的经济形态正经历一场深刻的数字化与智能化变革。相关行业报告显示，人工智能与大数据分析被认为是未来3-5年内对大湾区产生最大影响的技术。相应地，劳动力市场对人才的需求也发生了根本性转变：市场不再满足于单一专业人才，而是需要能够熟练分析数据并能应用于大湾区特色场景的“复合型人才”。

对于经管类专业而言，数据分析能力早已不是一项“加分项”，而是安身立命的“基本功”。概率论与数理统计课程，作为现代数据科学的理论基石，是学生专业能力培养的必学课程^[1]，该课程对毕业生未来核心竞争力的塑造至关重要，也直接关系到他们能否在湾区未来的经济建设中占有一席之地。然而，初步调研与既有文献显示，当前高校经管类专业的概率统计课程教学普遍存在与时代需求脱节的问题：传统的教学模式往往侧重于抽象的理论推导与公式记忆，忽视了知识与实践的结合，导致学生学习兴趣不高、应用知识的能力不强^[2-4]。

尽管过去有不少学者呼吁过概率统计的教学改革，但前人的讨论缺少一个关键视角，即聚焦于大湾区这一国际科技创新中心的改革方向。为此，本文讨论经管类专业概率统计课程存在的问题，指出大湾区高校课程内容与学生就业和升学的双重需求相脱节的现状，探究课堂教学如何与大湾区特色经济场景相结合，以及如何构建一个既能夯实理论基础、强化应用技能，又能深度融入大湾区特色的教学新范式。

二、经管类统计课程存在的问题

概率统计课程教学在大湾区人才培养战略中具有重要地位，但知识的“学”与“用”之间仍存在明显的差距，我们必须正视当前教学实践与理想目标之间的鸿沟。通过对大湾区高校经管专业学生进行问卷调查，本文指出当前大湾区高校课程教学存在的三个主要问题，这些问题共同影响了人才培养的质量。

（一）课程内容与学生就业升学双重需求脱节

调查数据显示，经管类专业超50%的毕业生进入金融、科技等高附加值行业，但现有课程仍以理论讲授为核心，对财务数据分析、Python建模、数据可视化等实操技能训练不足。学生问卷反馈显示，37.5%的学生计划直接就业（其中84.21%期望进入金融/财会行业），32.14%计划升学（其中90%选择经管相关专业）。就业群体需具备企业级数据的分析能力以适应岗位要求，升学群体则需深化计量经济学等理论知识以应对考研或学术研究，而现有课程难以兼顾这两类需求，导致课堂教学内容与学生发展需求相脱节。

（二）教学内容与大湾区经济实践联系薄弱

调查问卷显示，39.29%的学生认为课程案例与大湾区实际问题结合不够紧密，缺乏区域针对性。当前教学多采用全国通用案例（如宏观GDP分析），对深港资金流动监测、珠三角制造业数字化转型等区域性特色场景缺乏深度挖掘。这种脱节导致学生难以将统计工具应用于本地产业实际问题，与《粤港澳大湾区发展规划纲要》提出的“产学研深度融合”要求存在偏离。

（三）统计课程的学习效果有待提升

一方面，概率分布、假设检验等重点内容的理解与掌握依赖高等数学知识（如积分、求导、分布函数），但学生因数学基础薄弱或知识遗忘而普遍难以衔接相关内容。问卷显示，55.36%的学生反映“概率统计”章节难度大，35.71%的学生对“回归分析”内容理解困难。另一方面，传统教学停留在公式推导，缺乏真实案例支撑，导致学生无法建立统计理论与实际问题之间的联系；同时，课程内容与学生的专业领域结合较少，使得学生难以认识到学习概率统计的重要性和应用价值^[5]。

上述问题给经管类学生的发展与大湾区建设发展均带来了负面影响。就学生群体而言，所学技能与企业需求脱节可能加剧就业压力；于区域发展而言，专业人才缺口正制约大湾区产业升级进程。因此，探索更有效的统计课程改进路径显得尤为关键——既要满足学生就业与升学的双重需求，又要立足于大湾区区域经济发展，切实增强教学实效性，这对促进学生职业发展、支撑湾区经济建设具有现实意义^[5]。

三、课程改进建议

为解决概率统计课程在教学中存在的问题，本文遵循“问题导向、需求驱动、区域嵌入”的核心原则，提出一套体系化的课程改进方案。该方案从课程框架、教学内容、教学方法、考核方式及过程关怀五个角度出发，以期切实突破当前教学瓶颈，实现教学质量的实质性提升，为大湾区培养出更多符合产业发展需求的复合型数据人才。

（一）建立兼顾就业与升学需求的课程框架

课程内容应紧密对接学生的未来职业规划发展。对于占比较高的就业意向学生（占比37.5%），特别是期望进入金融/财会行业（占就业意向群体的84.21%）的群体，亟需强化数据分析工具（如Excel、R、Python）的实操能力培养，将证券/基金从业资格、大数据财务分析师等证书考核要点（如假设检验用于风险评估）融入教学环节，为学生设置贴近岗位需求的实践作业（如财务数据分析、股价预测等）。可参考已有文献中关于“增加数据挖掘、数据仓库等处理海量数据的分析方法”的建议增设大数据分析模块，契合《新质生产力下社会对数智化财务人才的素质需求研究》提出的“技术+业务”复合能力要求^[3]。对于32.14%的升学群体（其中约90%选择经管相关专业深造），则需要深化回归分析、时间序列分析、抽样推断等理论知识的深度，并加强Stata/R等学术软件应用及实证研究能力训练。

（二）融合大湾区特色场景，构建产学研合作平台

为改变教学内容与大湾区实践脱节的现状，需将区域性特色场景融入课程教学过程中。具体包括开发一系列基于大湾区真实经济问题的教学案例：如利用深港资金流动数据分析外汇风险模型、基于珠三角制造业产能数据进行产能预测回归分析、应用湾区上市公司数据研究ESG评级与财务绩效的相关性等，以响应“高校课程需对接湾区跨境金融、智慧物流等新兴业态”的区域化教育主张^[4]。为落实《粤港澳大湾区发展规划纲要》中“建立以企业为主体、市场为导向、产学研深度融合的技术创新体系”的要求，可以引入腾讯、微众银行等湾区企业的真实数据作为学生实践作业的分析对象，既有助于学生理解抽象统计理论，又能加深其对大湾区企业发展的认知。同时，探索与港澳高校合作开展“粤港澳贸易流监测”等联合研究课题，并借鉴《湾区资本市场与财务治理》课程关于“校企共建实战平台”的经验，设立校内“数智业务实战中心”或校外实训基地^[6]。此外，通过举办聚焦大湾区热点问题的“数据挑战赛”（如澳门旅游消费预测、广深科创走廊专利分析），以赛促学，推动知识向实践能力转化。

（三）创新教学方法，破解学习难点

针对学生普遍反映学习困难的“概率统计”（55.36%认为难）和“相关与回归分析”（35.71%认为难）等章节，必须改变教学模式，摒弃单一的“讲授法”。正如现有文献所指出的：“教学内容与实际应用脱节会导致学生被动学习”^[2]。应采用差异化教学设计：对于“概率统计”中的抽象概念和依赖高等数学基础的部分（如积分、分布函数），可事先录制微课供学生课前复习巩固，课堂上则结合生活化案例降低理解门槛^[2]。对于“回归分析”内容，应设计贯穿“基础理论—应用实践—深化拓展”的教学链条。同时积极推广混合式教学模式，科学划分内容边界，精心设计“哪些内容适合面授、哪些适合在线学习、哪些适合翻转课堂”，有效解决课时不足和学生基础差异的问题^{[2][7]}。

（四）推动考核方式转型，强化实践能力评价

将案例分析、数据建模等实践任务纳入评价体系，减少机械性计算题占比。理论知识的学习效果最终需要通过实践来检验和提升，因此应增加课程中软件实操讲解和练习时长，特别是学生期望强化的数据清洗、可视化、建模编程等技能，减少机械枯燥计算类作业占比；探索过程性评价改革，提高案例分析报告、实操项目完成度等在总评成绩中的比重^[8]；改变“重理论轻实践”的现状，设计更多基于真实数据或大湾区场景的分析任务；鼓励学生运用Tableau等可视化工具呈现分析结果，如制作大湾区某产业分析报告、金融风险预警仪表盘等^[9]。这种“学中练、练中学、练后展”的模式契合“可视化成果驱动学习动机”策略，能显著提升学习成就感^[5]。

（五）关注学习过程与个体差异，激发内在学习动力

问卷结果显示，学习专注力、学习习惯、学习动机及数学基础对统计课程学习效果的影响较大（图1）。在教学中应将课程内容与学生未来的职业发展目标紧密联系，注重激发学生的内在学习动机^[9]。对于数学基础薄弱的学生，在涉及如概率分布等依托高等数学知识的关键章节，提供前置知识的补充学习资源（如微课、习题集）^[10、11]。同时，问卷结果显示，80.36%的学生认为实时数据互动（如股票市场模拟）可以提升课程趣味性，64.29%的学生认为可视化的成果展示可以提升课程趣味性。因此，课程教学可融入相关内容，以增强学生的学习兴趣并提升学习效果。

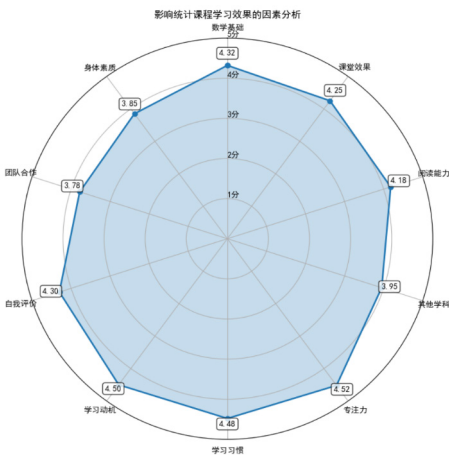


图1：影响统计课程学习效果的因素分析雷达图

四、结语

概率统计课程的教学改革从来不是孤立的课堂试验，而是回应区域发展需求的必然选择。本文立足于大湾区产业升级对复合型数据人才的迫切需求，研究当前课程教学中存在的结构性问题，并提出了针对性的改进方案。研究发现，课程内容与学生就业和升学的双需求相脱节、教学案例与大湾区经济实践分离、传统教学方法的学习效果不理想，是影响人才培养质量的核心问题。本文据此构建了“问题导向、需求驱动、区域嵌入”的改革体系为上述问题提供了解决方案，为大湾区高校经管类概率统计

教学改革提供理论补充和实践参考。本文所提的关于课程框架、教学方法及评价机制等具体建议，能够为大湾区高校经管类专业的教学改革提供有效参考，有助于培养符合大湾区区域发展需求的复合型人才，为大湾区的可持续发展提供动力。

课程改革是一个动态优化、持续完善的过程。本文提出的框架仍需在教学实践中进一步检验和完善。未来的研究可结合不同高校的办学特色及学生群体的异质性，针对区域产业需求探索个性化教学模式，并对改革成效进行长期追踪。通过这些实践探索，有望为大湾区培养更多兼具理论与实践能力的高素质经管人才，最终服务于区域经济社会高质量发展战略。

参考文献

[1] 孙静娟、杨光辉、杜婷编著. 统计学 (第四版) [M], 清华大学出版社, 2021.

[2] 刘光明. 对“概率论与数理统计”课程教学改革的几点思考 [J]. 教育教学论坛, 2025(26): 81-84.

[3] 杨盼盼. 大数据时代经管类专业经济统计学课程改革研究 [J]. 现代商贸工业, 2024(7): 197-198.

[4] 黄水香. 高等教育服务粤港澳大湾区发展需求分析 [J]. 肇庆学院学报, 2025(4): 113-118.

[5] 江淑芳. 数字经济背景下高校经管专业人才培养素养能力提升策略研究 [J]. 经济师, 2024(1): 187-188.

[6] 陈晓珊. 粤港澳大湾区智能财会融合型人才培养路径探索——以《湾区资本市场与财务治理》课程为例 [J]. 商业会计, 2025, (12): 138-141.

[7] 金春红, 隋振婷, 曹颖. 数学文化视角下经管类概率统计课程教学改革研究 [J]. 产业与科技论坛, 2024, 23 (21): 158-160.

[8] 杜健, 梁红梅, 胡红娟. 基于成果导向的“概率论与数理统计”教学改革研究 [J]. 高等教育研究学报, 2023, 46 (02): 110-116.

[9] 石凯, 刘洪江. 地方高等院校概率统计课程教学改革模式探析 [J]. 大学教育, 2022, (10): 127-130.

[10] 吴云顺, 彭艳芳. 师范专业认证背景下概率统计课程教学改革初探 [J]. 科教导刊, 2021, (23): 47-49+56.

[11] 徐丽, 张振华. 大数据时代经济统计专业课程思政融合教学评价研究 [J]. Applied Statistics and Data Science, 2025, 1(2): 96 - 102.

高中“双新”驱动的“五维联动”生涯教育实践

李春波, 陈璟

桂林市桂林中学, 广西 桂林 541000

DOI:10.61369/ASDS.20250700135

摘 要 : 在“双新”(新教材、新高考)改革背景下,高中教育需从“分数导向”转向“素养导向”,生涯规划教育成为培养学生个性化发展与终身发展能力的核心任务。桂林中学以数据技术为支撑,构建“课程体系、师资建设、家校协同、实践体验、技术支撑”的“五维联动”生涯教育实践模型,通过“四阶一体”课程体系、专业化师资队伍、深度家校社协同、多元实践体验及数据赋能技术,有效提升学生选科科学性与职业规划能力。实践表明,该模型使学生科学选科率达85%,90%学生明确职业倾向,形成可推广的高中生涯教育实践范式,为“双新”背景下基础教育人才培养提供“桂中方案”。

关 键 词 : “双新”驱动;生涯规划教育;五维联动;数据赋能;选科指导

The Practice of "Five-Dimensional Linkage" Career Education Driven by "Double New Initiatives" in Senior High Schools

Li Chunbo, Chen Jing

Guilin Middle School, Guilin, Guangxi 541000

Abstract : Against the background of the "Double New Initiatives" (new textbooks and new college entrance examination) reform, senior high school education needs to shift from a "score-oriented" approach to a "competency-oriented" one. Career planning education has become a core task for fostering students' personalized development and lifelong learning capabilities. Relying on data technology, Guilin Middle School has constructed a "Five-Dimensional Linkage" practical model for career education, which encompasses "curriculum system, teacher development, home-school collaboration, practical experience, and technological support". Through the "four-stage integrated" curriculum system, professional teacher teams, in-depth home-school-community collaboration, diverse practical experiences, and data-enabled technology, the school has effectively improved the scientific nature of students' subject selection and their career planning abilities. Practice shows that this model has enabled 85% of students to make scientific subject selections and 90% of students to clarify their career orientations. It has formed a replicable practical paradigm for senior high school career education and provided the "Guizhong Solution" for talent cultivation in basic education under the background of the "Double New Initiatives".

Keywords : "Double New Initiatives" driven; career planning education; five-dimensional linkage; data empowerment; subject selection guidance

引言

在“双新”(新教材、新高考)改革纵深推进与新高考“3+1+2”模式全面实施的背景下,高中教育正经历从“分数导向”到“素养导向”的深刻转型。《普通高中课程标准(2020年修订版)》明确将生涯规划教育列为培养学生适应社会、高等教育及职业发展能力的核心任务,要求学校突破传统育人模式,助力学生个性化素养发展,实现生涯育人新模式。

然而,新高考赋予的选科自主权也带来决策复杂化、学科与专业匹配度不足等挑战,加之学生自我认知模糊、职业意识淡薄,传统经验式指导难以满足精准化需求。桂林中学立足这一现实,以数据技术为抓手,构建“课程体系、师资建设、家校协同、实践体验、技术支撑”的“五维联动”生涯教育模型,探索出一条“数据赋能、全程渗透”的实践路径,为破解“双新”背景下高中生涯教育难题提供了可借鉴的范式。

基金项目:广西教育科学“十四五”规划2023年度广西教育信息化教学应用实践共同体专项课题《数据驱动下高中的精准教学决策、规划学生发展和助力教师成长的研究》(2023ZJY552)。

作者简介:

李春波,桂林市桂林中学,信息中心副主任,中小学高级职称,邮箱:64321930@qq.com;

陈璟,桂林市桂林中学,生涯发展中心副主任,中小学高级职称,邮箱:530255254@qq.com。

一、双新驱动素养导向的大数据时代高中学生生涯教育面临的问题

（一）政策与教育变革驱动

新高考“3+1+2”模式与“双新”政策的落地，标志着高中教育进入转型关键期。《普通高中课程标准（2020年修订版）》明确要求学校为学生适应社会生活、高等教育和职业发展奠定基础，生涯规划教育被纳入核心育人任务^[1]。新高考赋予学生选科自主权，但也带来选科决策难、学科与专业匹配度低等问题。2019年国务院办公厅《关于新时代推进普通高中育人方式改革的指导意见》提出“到2022年，选课走班教学管理机制基本完善”^[2]；2021年广西壮族自治区印发《广西深化普通高等学校考试招生制度改革实施方案》，明确自2021年入学的高一新生起实施新高考，要求学校建立科学的选课指导体系^[3]。政策倒逼高中教育突破传统“文理分科”局限，构建以学生发展为核心的生涯教育体系。

（二）学校现实需求

桂林中学作为广西首批示范性高中，在“双新”实施中面临三重挑战。学生层面，多数学生自我认知模糊，职业意识淡薄，选科存在“从众”“唯分数”倾向，如2021届高一学生中，65%选科依赖家长建议，仅30%能结合自身兴趣选择；学校层面，生涯课程碎片化，缺乏系统性设计，师资多为班主任兼职，专业化水平不足，学科教学与生涯教育融合度低；社会层面，家长过度关注学业成绩，忽视学生职业适配性，本地校企合作资源零散，学生职业体验机会匮乏。这些问题导致学生生涯规划能力薄弱，难以适应新高考与未来职业发展需求。

（三）数据赋能的迫切性

新高考背景下，选科赋分规律、高校专业选考要求、学生兴趣测评数据等信息呈爆发式增长，传统经验型指导已无法满足精准化需求。例如，广西“3+1+2”模式下有12种选科组合，不同组合对应数百个高校专业，仅靠教师人工梳理难以高效匹配学生需求。同时，学生学业成绩、社会实践、心理特质等数据分散，缺乏整合分析，导致生涯指导针对性不足。因此，引入大数据技术，构建“数据驱动”的生涯教育支持体系，成为破解上述难题的关键路径。

二、“五维联动”生涯教育实践模型构建

桂林中学以“双新”驱动为指引，以“数据赋能”为核心，整合“课程、师资、家校、实践、技术”五大维度，形成闭环式生涯教育实践模型，实现“知己—知彼—决策—行动”的全链条指导。

（一）构建“四阶一体”生涯课程体系

基于学生认知规律与生涯发展阶段，设计“基础认知—深度探索—决策实践—终身发展”的四阶课程，实现生涯教育与高中三年学业的深度融合。

1. 基础认知阶段（高一）：聚焦“自我认知与政策理解”，开设新高考政策解读课，解析“3+1+2”模式下选科规则与高校专

业要求；引入霍兰德职业兴趣测试、MBTI 性格测试，帮助学生初步定位职业倾向；结合学校近三年选科数据与广西高校录取数据，开展“智慧选科专题讲座”，例如通过分析“物理+化学+生物”组合对应96%的专业覆盖率、“历史+地理+政治”组合对应58%的专业覆盖率，引导学生理性选科。

2. 深度探索阶段（高二）：侧重“职业探索与能力匹配”，课程包含大学专业分类解析、选科策略指导（如SWOT分析法、决策平衡单的应用）；开展“职业万花筒”主题班会，邀请医生、律师、工程师等家长或校友分享职业经历；组织学生参与“模拟政协”社团活动，模拟职业场景，提升社会认知。此外，在物理课中结合“大国工匠”案例讲解工程类职业路径，在语文课中通过“我的职业理想”写作任务引导价值观反思，实现学科教学与生涯教育的融合。

3. 决策实践阶段（高三）：聚焦“升学规划与职业准备”，开设志愿填报策略课，教授“职业目标倒推法”，如根据目标职业确定大学专业，再反推选科要求；开展模拟面试、简历制作培训，提升学生求职竞争力；以班级为单位组织“家长—学生生涯对话会”，结合学生生涯画像与高校招生动态，共同制定升学计划。

4. 终身发展阶段（全程渗透）：开发《职业生涯规划初探析》《遇见未来的自己》等校本课程，将生涯教育融入日常教学；建立“生涯成长档案”，记录学生从高一到高三的测评结果、实践经历与目标调整，培养学生终身规划意识。

（二）打造专业化生涯师资队伍

通过“内培外引、角色赋能”，构建“专职+兼职+外聘”的三维师资队伍，提升生涯指导专业化水平。

1. 内培外引，夯实基础：定期邀请华南师范大学、广西师范大学的生涯规划专家，以及企业HR开展培训，内容涵盖生涯理论（如舒伯生涯发展理论）、测评工具使用、选科数据分析等；2021—2024年累计开展培训28场，覆盖教师373人次。成立“学生生涯中心”，联合心理教研组开发《高中生生涯导航手册》，梳理新高考政策、职业信息、高校专业要求等实用内容，为教师提供指导工具。

2. 角色赋能，明确职责：推行“心理教师生涯导师制”，要求5名专职心理教师每人对接3—4个班级，协助班主任每学期开展2次生涯主题班会，如“自我认知与职业匹配”“选科决策模拟”等；设计“班主任生涯彩虹图”工程，通过workshops帮助班主任掌握舒伯生涯发展理论，将生涯指导融入班级管理，例如在家长会中加入“生涯规划沟通技巧”培训，引导家长关注学生兴趣与职业适配性。

3. 学科融合，全员参与：鼓励学科教师在教学中渗透生涯元素，如数学教师结合统计学知识分析选科赋分概率，生物教师讲解生物医药类职业发展前景；建立“学科+生涯”教案库，收录优秀融合案例36个，推动生涯教育从“专人负责”向“全员参与”转变。

（三）深化家校社协同机制

构建“家长导师团+校企合作基地+高校联盟”的协同网络，

整合社会资源，拓展生涯教育场景。

1. 家长导师团，搭建家庭桥梁：招募律师、医生、工程师、高校教师等职业背景的家长82人，成立“家长导师团”，定期开设“职业微课堂”，如医生家长讲解医学专业与职业发展路径，律师家长分享法律行业现状；每年组织“亲子生涯规划日”，通过“职业角色扮演”“生涯决策游戏”等活动，增强亲子间的生涯沟通，2024年参与家庭达420户。

2. 校企合作，拓展实践场景：与桂林南药股份有限公司、桂林理工大学科技园等3家企业共建“生涯实践基地”，安排学生寒暑假参与岗位见习，如在药企观察药品研发流程，在科技园了解人工智能产业应用；2022-2024年累计组织500人次参与见习，学生撰写的《生物医药产业见习报告》《人工智能职业体验反思》等12篇社会实践报告获市级奖项。

3. 高校联盟，衔接升学路径：与桂林理工大学、广西师范大学、桂林电子科技大学等3所高校建立合作，开展“大学专业体验营”，组织学生走进实验室、旁听专业课，提前感知学术氛围与专业要求；邀请桂林中学毕业的优秀大学生返校宣讲，分享大学学习与职业规划经验，2024年累计开展宣讲16场，覆盖学生1800人次。

（四）开展多元化职业生涯实践体验

以“实践育人”为导向，设计“课程体验+社团活动+研学实践+高校交流”四位一体的实践体系，强化学生职业认知与动手能力。

1. 课程体验，深化理论认知：依托“四阶”课程，开展“选科模拟决策”“职业能力测评”等实践活动，如使用决策平衡单分析“物理+化学+地理”与“历史+政治+地理”两种组合的利弊，提升学生决策能力。

2. 社团活动，模拟职业场景：“模拟政协”社团围绕“青少年职业教育”“新高考选科政策”等议题开展调研，撰写模拟提案；“创客社团”与企业合作开展项目式学习，如设计校园智能储物系统，了解产品设计与工程职业；2024年“模拟政协”社团提案获桂林市青少年模拟政协提案大赛一等奖。

3. 研学实践，拓展社会视野：组织高一学生每年到桂林市综合实践基地开展三天研学，参与“职业体验工坊”（如茶艺师、电工、烘焙师）、“生涯规划工作坊”等活动；2023-2024年累计参与研学学生1200人次，92%学生反馈“通过实践明确了兴趣方向”。

4. 高校交流，衔接高等教育：联合桂林理工大学开展“大学科学营”，组织学生参观地质博物馆、参与化学实验；与广西师范大学合作开展“师范生—高中生结对”活动，让师范生分享大学专业学习经验，帮助高中生提前了解高校生活与专业要求。

（五）数据赋能的精准化技术支持

引入大数据与人工智能技术，构建“数据采集—分析—应用”的闭环，实现生涯指导精准化。

1. 建立学生生涯画像：整合学生学业成绩（如各学科排名、历次考试分数）、兴趣测评（霍兰德、MBTI结果）、社会实践（见习经历、社团活动）、职业倾向等数据，通过科大讯飞智慧校园系统生成“学生生涯画像”，直观呈现学生优势、短板与职业适

配方向。例如，学生张某的画像显示其美术特长突出、空间思维能力强，系统推荐“建筑设计”“视觉传达”等相关专业与“历史+地理+艺术”选科组合。

2. 智能选科决策支持：利用大数据分析广西近五年选科赋分规律，挖掘不同选科组合的竞争优势，如“物理+化学+生物”组合高分段竞争激烈，但专业覆盖率高；“物理+生物+地理”组合竞争较小，适合中等成绩学生。系统对接教育部“阳光高考”平台，实时更新全国高校专业选考要求，学生输入目标专业后，系统自动推荐匹配的选科组合，2024届高一学生使用该系统后，科学选科率从2021届的45%提升至85%。

3. 动态监测与调整：建立生涯教育数据监测平台，跟踪学生选科调整、职业目标变化、实践参与情况，定期生成《生涯发展报告》，为教师、家长提供干预建议。例如，若学生多次调整选科组合，系统提示教师与家长关注其决策困惑，通过一对一指导帮助明确方向。

三、高中“双新”驱动的“五维联动”生涯教育实践成效

（一）学生层面：生涯规划能力显著提升

1. 选科科学性大幅提高：2024届高一学生中，85%能结合兴趣、能力与高校要求选择科目，“从众选科”比例从2021届的55%下降至10%；典型案例如学生张某，初期受同学影响选择“物理+化学+生物”组合，通过兴趣测评与选科数据推荐，发现自身美术特长与“历史+地理+艺术”组合更匹配，调整后学习动力显著提升，高二美术作品获市级奖项。

2. 自我认知与职业意识增强：90%学生通过测评与课程明确自身优势与职业倾向，较2021届提升50个百分点；学生职业体验参与率从2021届的30%提升至85%，累计500人次参与岗位见习，12篇社会实践报告获市级奖项，职业认知从“模糊”转向“清晰”。

3. 升学与就业衔接更顺畅：2023届毕业生中，78%录取专业与高中选科、职业倾向一致，较2021届提升32个百分点；部分学生通过职业体验明确升学目标，如学生李某参与药企见习后，立志报考药学专业，最终被中国药科大学录取。

（二）教师层面：生涯指导能力与科研水平双提升

1. 专业化水平显著提高：373名教师均接受生涯规划培训，5名心理教师成为专职生涯导师，班主任掌握舒伯生涯发展理论^[4]与选科指导方法；家长满意度调查显示，92%家长认可教师的生涯指导服务，较2021届提升40个百分点。

2. 科研与课程开发成果丰硕：学生生涯中心与心理教研组联合开设校本课程3门，发表《高中生生涯规划中的德育教育融入方法探思》《数据赋能下高中选科指导的实践路径》等相关论文5篇；开发《高中生生涯导航手册》《“学科+生涯”融合教案集》等教学资源，被广西多所高中借鉴使用。

（三）学校与社会层面：形成生涯教育品牌效应

1. 学校影响力扩大：桂林中学成为广西教育基础信息化融合

创新实验校，获全国重点 AI 项目“农村地区教师教学能力智能评测与教学精准辅助技术”平台应用试点学校、北京师范大学“AI 大先生循证教研赋能联盟单位”；2024 年举办的生涯规划系列讲座被桂林市教育局网站专题报道，吸引广西 10 余所高中到校考察学习。

2. 家校社协同生态完善：“家长导师团”开展活动 10 场，覆盖学生 1200 人次；与 3 家企业、3 所高校建立长期合作，形成“职业体验—高校衔接—就业支持”链条，为学生生涯发展提供持续支持；2025 年参评桂林市教育局的“桂林教育数字化转型优秀案例”，多所全区兄弟学校到校交流推广。

四、高中“双新”驱动的“五维联动”生涯教育实践的经验启示与未来展望

（一）经验启示

1. 政策引领是前提：“双新”与新高考政策为生涯教育提供方向指引，学校需主动对接政策要求，将生涯教育纳入整体育人体系，避免“碎片化”“形式化”^[5]。

2. 数据赋能是关键：大数据技术能整合分散信息，实现选科指导、职业匹配精准化，破解传统经验式指导的局限，是“双新”背景下生涯教育提质增效的核心支撑。

3. 协同育人是保障：仅靠学校难以满足学生多样化生涯需求，需整合家长、企业、高校资源，构建“家校社”协同网络，拓展生涯教育场景与资源。

4. 学生主体是核心：生涯教育需尊重学生个性差异^[6]，通过“四阶”课程、多元实践，引导学生主动探索“我想做什么”“我

能做什么”“我该做什么”，培养终身规划能力。这就是班主任在管理实践中融入德育教育的价值所在。

（二）未来展望

1. 深化数据赋能，优化模型：引入 AI 技术升级“智能选科系统”，增加动态预测功能，如根据学生成绩变化预测选科竞争力；联盟广西 10 所兄弟学校构建“共同体学校高中生生涯”数据库，实现校际资源共享与数据互通。

2. 强化师资专业化，提升能力：选派 10 名教师参加国家级生涯规划师培训，培养骨干师资；扩大“学科+生涯”教案库规模，开发“生涯教育微课”，提升学科教师融合能力。

3. 完善评价体系，持续改进：设计“生涯素养评价量表”，从自我认知、职业探索、决策能力、实践能力四个维度评估学生生涯素养，纳入综合素质评价；建立家校双向反馈机制，每学期开展生涯教育满意度调查，根据反馈调整方案。

4. 拓展实践资源，助力终身发展：新增 5 家企业实践基地，覆盖人工智能、生物医药、文化创意等新兴领域；与区外高校（如华南理工大学、中山大学）建立合作，开展“跨省大学体验营”，帮助学生了解不同地区高校与职业环境，实现“从学业选择到终身发展”^[7]的跨越。

桂林中学以“双新”驱动为契机，通过课程重构、数据赋能、家校社协同，形成了科学化、系统化的高中“双新”驱动的“五维联动”生涯教育实践体系。未来将持续深化实践，融合“五育”，促进学生的全面发展，提升学生的综合能力素养，助力学生职业生涯实现“从学业选择到终身发展”的跨越，为数智化时代人才培养和广西基础教育信息化融合创新实验校建设提供“桂林中学方案”。

参考文献

- [1] 国务院办公厅. 关于新时代推进普通高中育人方式改革的指导意见 [Z]. 2019.
- [2] 中华人民共和国教育部. 普通高中课程标准（2020 年修订版）[S]. 北京：人民教育出版社，2020.
- [3] 广西壮族自治区人民政府. 广西深化普通高等学校考试招生制度改革实施方案 [Z]. 2021.
- [4] 舒伯. 职业生涯发展理论 [M]. 北京：中国轻工业出版社，2018.
- [5] 王树铨. 高中教师学科生涯辅导工作的现状与对策分析 [J]. 新课程，2022（20）：18-19.
- [6] 汤磊. 高中班主任在管理实践中融入德育教育的价值 [J]. 中学课程辅导，2022（13）：54-56.
- [7] 孙丹. 高中生生涯规划中的德育教育融入方法探思 [J]. 中小学教育，2024（05）：23-25.

融合 Shiny 与大语言模型的数理统计教学创新研究

赖廷煜¹, 胡志明^{2*}

1. 广西师范大学 数学与统计学院, 广西 桂林 541006

2. 上海财经大学浙江学院, 浙江 金华 321013

DOI:10.61369/ASDS.2025070014

摘 要 : 本文针对传统数理统计教学存在理论抽象、学生被动及与实践脱节等问题, 提出一种融合编程技术 (R 语言与 Shiny 库) 与大语言模型的创新教学框架。该框架通过交互式可视化工具将抽象理论转化为动态应用, 支持学生以“观察 – 猜想 – 验证”路径主动构建统计思维; 同时采用项目驱动 (PBL) 教学法, 依托大语言模型提供个性化项目设计及全过程辅助支持, 降低学习难度, 促进理论与实践深度融合, 旨在提升教学效果, 培养符合时代需求的数据分析人才。

关 键 词 : 数理统计; 教学改革; R 语言; Shiny; 大模型; 项目驱动学习

Innovative Teaching of Mathematical Statistics Integrating Shiny and Large Language Models

Lai Tingyu¹, Hu Zhiming^{2*}

1.School of Mathematics and Statistics, Guangxi Normal University, Guilin, Guangxi 541006

2.Zhejiang College, Shanghai University of Finance and Economics, Jinhua, Zhejiang 321013

Abstract : This paper addresses issues in traditional mathematical statistics education—such as abstract theory, passive learning, and disconnection from practice—by proposing an innovative teaching framework that integrates programming technology (R language and Shiny library) with large language models. The framework transforms abstract statistical theories into interactive and dynamic visual applications, enabling students to actively develop statistical thinking through an "observe–hypothesize–verify" inquiry–based learning process. Furthermore, a project–based learning (PBL) approach is adopted, supported by large language models to assist instructors in designing personalized multi–level projects and providing students with scaffolding support—including code assistance, idea generation, and problem debugging—throughout project implementation. This reduces the difficulty of project–based learning and fosters deep integration of theory and practice. The study aims to enhance teaching effectiveness and cultivate versatile data analysis talents suited to contemporary demands.

Keywords : mathematical statistics; teaching reform; R language; Shiny; large language model; project–based learning

引言

随着大数据和人工智能的快速发展, 统计学已逐渐成为推动科学研究和社会应用的重要方法论工具。作为统计学中的核心课程, 数理统计的教学目标不仅在于传授公式定理, 更在于帮助学生理解其中的逻辑关系, 并培养其分析与解决实际问题的能力。然而, 目前国内多数高校仍沿用传统的“讲授–练习”教学模式, 过度强调推导与记忆。这种方式使学生难以真正理解抽象概念, 也削弱了学习主动性。同时, 教材与考试内容多依赖简化和理想化的数据, 与现实中嘈杂、不完整且结构复杂的数据差距显著, 导致学生在面对实际问题时常常无所适从。正因如此, 推动数理统计教学改革已成为学界持续关注的重要议题。

近年来, 开源统计软件的广泛应用为教学改革提供了新的技术契机。其中, R 语言因其在统计建模与可视化方面的强大功能, 逐渐被引入课堂教学。郭林祥等^[1]提出在实验教学中使用 R 语言, 以帮助学生更直观地理解核心定理, 并同时提升编程能力。宋述芳等^[2]指

基金项目: 广西自然科学基金项目 (2025GXNSFAA069336); 广西自然科学基金项目 (2025GXNSFAA069453); 2024 年度人工智能与先进制造中外人文交流研究院项目 (CCIPERGZN202439); 2023 年度上海财经大学浙江学院发展基金 (23YJC630054)。

作者简介: 赖廷煜, 广西师范大学数学与统计学院, 博士, 研究方向为函数型数据分析, 非参数检验。

通讯作者: 胡志明, 上海财经大学浙江学院, 副教授, 研究方向为数理统计及应用。

出，R 语言不仅能够改善学习效果，还能为跨学科研究提供支持。刘钰等^[3]则展示了 R 在空间统计课程中的可视化优势，使学生能够更直观地把握复杂的空间数据关系。与此同时，汪浩等^[4]在大数据与人工智能背景下建议将 R 贯穿于课程全过程，以提升 IT 专业学生的实践能力和综合素养。这些研究结果共同表明，将 R 引入教学不仅能弥补传统教学的不足，还能在理论与实践之间建立有效的连接。

作为 R 语言的重要扩展工具，Shiny 能够将统计理论转化为交互式、动态化的演示。相关研究已证明其在医学和公共卫生等领域的成功应用。王思娜等^[5]基于 R-Shiny 开发了临床试验分配误差系统，使统计推断更具可操作性。崔怡丹等^[6]利用 Shiny 构建了临床数据分析平台，展示了交互式系统在实际数据处理中的应用潜力。胡晓雯等^[7]则在传染病数据可视化研究中验证了 Shiny 的扩展性和适用性。总体而言，Shiny 在缓解数理统计中抽象性强、直观性不足等问题方面效果显著，同时能够提升学生的学习兴趣与课堂参与度。

除了编程与可视化工具外，教学方法的创新同样至关重要。李倩^[8]指出，项目驱动学习（Project-Based Learning, PBL）能够显著提升应用型课程的实践效果，并增强学生的自主性。金鑫等^[9]基于 BOPPPS 模型提出了一种新的教学设计，强调课堂结构化与学生主动参与的重要性。宋丽娟等^[10]在贝叶斯公式的教学中引入研究型方法，验证了探究式学习在促进概念理解方面的有效性。王泽龙等^[11]则主张通过个性化教学模式培养学生的创新思维。张慧星等^[12]结合新工科背景推动数理统计的实践改革，突出了理论知识与工程应用的结合价值。同时，杨晓丽^[13]与闫兰香^[14]从国家战略和“新文科”建设的角度出发，强调统计学课程改革应服务于社会发展和时代需求。

近年，人工智能在数学与统计教育中的应用受到关注。García-López 等在芬兰的实证研究显示^[15]，GenAI 的“有用性”对学生使用意向影响最大，“愉悦感”能增强对“有用性”和“易用性”的感知，而加入“兼容性”后，模型解释力显著提升，突出了教育环境差异的重要性。Setälä 等^[16]基于技术接受模型（TAM）研究发现，学生对生成式 AI 工具的接受度主要取决于其“有用性”、“易用性”和“兼容性”，若工具能降低使用门槛并嵌入课程，学习积极性与参与度将显著提升。这些最新研究表明，人工智能，尤其是大语言模型，正在成为推动高等教育教学模式转型的重要力量。

综观现有研究，数理统计教学改革已在 R 语言的引入、Shiny 的应用和 PBL 方法等方面取得了明显进展，但仍存在一些不足：第一，多数研究侧重单一工具或方法，缺乏系统整合；第二，教学成效评估多以短期实验为主，缺乏长期与量化分析；第三，虽然人工智能工具已逐渐进入教育领域，但如何与统计课程有机结合尚缺乏深入研究。本文利用 Shiny 和大语言模型（LLM）创造性地解决上述问题。具体来说，利用 Shiny 提升学生对抽象理论的理解，同时可以提高编程实践能力；利用大语言模型（LLM）的自然语言处理与代码生成能力，推动个性化与项目化学习的实施；进一步，建立科学的评估体系，从多维度检验教学改革的效果与可持续性。本文所提出的 Shiny 与 LLM 融合创新框架，旨在为培养符合时代需求的复合型数据分析人才提供新的思路。

一、传统数理统计教学方法的弊端分析

传统教学模式在长期的实践中形成了一套成熟的知识传授体系，但其固有的弊端在当前强调能力和创新的教育背景下日益凸显。

（一）理论的抽象性与学习的被动性

数理统计的理论体系建立在严格的数学基础之上，其核心概念，如概率分布、抽样分布、统计量等，都具有高度的抽象性。例如，中心极限定理是数理统计的灵魂，它阐述了“大量随机变量之和的分布近似于正态分布”这一深刻规律。在传统课堂上，教师通常通过黑板板书，一步步进行严格的数学证明。学生在课堂上忙于抄录笔记，课后则需要花费大量时间去理解和消化这些复杂的推导过程。由于缺乏直观的体验，学生很难理解该定理的适用条件、收敛速度及其在实践中的巨大威力。学习变成了单向的知识灌输，学生始终处于被动接收的地位，缺乏主动探究的欲望和机会。当他们面对“为何样本均值的抽样分布是正态的”这类根本性问题时，往往只能复述定理的结论，却无法形成深刻的、基于直觉的理解。

（二）知识与实践的脱节

传统教学的另一个显著弊端是理论与实践的割裂。教学内容严格遵循教材的章节体系，从概率论基础到参数估计，再到假设检验，线性的知识结构固然清晰，但也导致了知识的碎片化。学生学习了大量理论，却不知道这些理论的应用场景是什么。课后的习题大多是为了巩固课堂所学的公式计算，例如给定数据求期望、方差，或进行指定的 T 检验、F 检验等。这些习题的设计目标是训练计算能力，而非解决问题的能力。学生很少有机会接触到数据分析的全流程，包括如何提出问题、收集数据、清洗数据、探索性分析、模型选择、结果解释和报告撰写。这种“纸上谈兵”式的学习，使其在未来的工作和科研中难以适应真实、复杂的数据分析任务。

（三）评估方式的单一性

与教学模式相对应，传统的评估方式也极为单一，通常以期中、期末的闭卷考试为主。考试内容侧重于对概念的记忆和公式的熟练应用，能够有效考察学生对知识点的掌握程度，但却无法衡量其统计思维、编程能力和解决复杂问题的综合能力。这种评估方式向学生传递了一种导向：学习的目标就是为了通过考试。

因此，学生可能会采取应试驱动的策略，如考前突击背诵、刷题等，而忽略了对知识内涵的深入理解和能力的长期培养。

二、基于 R Shiny 的数理统计可视化教学设计

为了克服理论的抽象性，我们引入了 R 语言及其 Shiny 框架，旨在将静态的、抽象的统计理论转化为动态的、可交互的可视化应用，构建一种“观察－猜想－验证”的探究式学习模式。

（一）为何选择 R 和 Shiny

R 语言是专为统计计算和图形显示而设计的编程语言，被誉为“统计学家的通用语”。它拥有极其丰富的统计分析函数库和世界领先的可视化工具（如 ggplot2），是数理统计教学与实践的理想平台。Shiny 是 R 语言的一个扩展包，它允许开发者仅用 R 代码就能快速构建功能强大的交互式 Web 应用，而无需掌握 HTML、CSS 和 JavaScript 等前端技术。这极大地降低了教师开发教学工具的技术门槛。教师可以轻松地将一个统计模拟过程封装成一个带有滑动条、下拉菜单、复选框等控件的网页应用，学生通过操作这些控件，可以实时地观察参数变化对统计结果的影响。

（二）“观察－猜想－验证”的探究式学习模式

我们设计的教学模式颠覆了传统的“理论先行”路径，倡导一种由具体到抽象的认知过程：

1. 观察（Observe）：学生首先接触的不是冰冷的公式，而是一个精心设计的 Shiny 应用。他们可以像玩游戏一样，自由拖动滑块，改变参数，观察图形的实时变化。这个过程充满了趣味性和探索性，能够瞬间抓住学生的注意力。

2. 猜想（Conjecture）：在反复的互动操作中，学生会逐渐发现一些稳定的模式和规律。例如，在中心极限定理的演示中，他们会发现，无论原始总体是什么分布（均匀分布、指数分布甚至是偏态分布），只要样本量 n 足够大，样本均值的分布形态总是趋向于对称的钟形。基于这些观察，学生会自然地提出自己的猜想或假设。

3. 验证（Verify）：带着观察得出的猜想和疑问，学生被激发了强烈的求知欲，主动回到课本中去寻求理论解释。此时，学习数学证明不再是被动的任务，而是为了验证自己发现的内在驱动。教师的角色也从知识的灌输者转变为学生探究过程的引导者和启发者。

（三）教学案例

1. 案例一：中心极限定理（Central Limit Theorem）

我们设计一个 Shiny 应用，界面包含以下控件：

总体分布选择：一个下拉菜单，允许学生选择不同的总体分布，如正态分布、均匀分布、指数分布、泊松分布等。

样本容量（ n ）：一个滑动条，允许从 2 到 500 之间动态调整每次抽取的样本大小。

模拟次数：一个输入框，设定重复抽样的次数（如

10000 次）。

应用的核心展示区会实时绘制出样本均值的抽样分布直方图，并叠加一条理论上的正态分布密度曲线，见图 1，左边对应样本容量 5，右边样本容量 363。

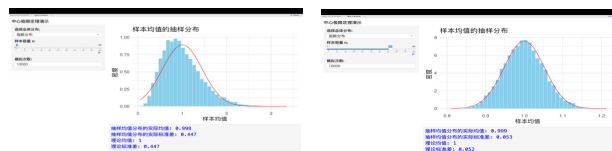


图 1. 中心极限定理的 shiny 实现

学生通过操作可以发现：当 n 很小时，抽样分布的形状与总体分布相似。

随着 n 的增大，无论总体分布是什么，抽样分布都迅速地 toward 正态分布逼近。抽样分布的中心始终稳定在总体均值附近，且其离散程度（方差）随着 n 的增大而减小。通过这个应用，中心极限定理的深刻内涵——“均值的正态化”和“方差的缩减”——变得一目了然。

2. 案例二：置信区间的理解（Confidence Intervals）

置信区间是数理统计中最容易被误解的概念之一。学生常常错误地认为“95% 置信区间意味着总体参数有 95% 的概率落在这个区间内”。为了澄清这个概念，我们设计了一个模拟应用：

设定一个已知的总体均值 μ （例如 $\mu=10$ ）。可以设置置信水平（如 90%, 95%, 99%）和样本量 n 。点击“抽样”按钮，应用会从总体中随机抽取 100 个样本，并为每个样本计算出一个置信区间，然后将这 100 个置信区间可视化地绘制出来。

那些成功“捕获”了真实总体均值 μ 的区间用绿色表示，未捕获的用红色表示。



图 2. 置信区间的 shiny 实现

学生会清晰地看到，当置信水平为 95% 时，100 个区间中大约有 95 个是绿色的，5 个是红色的。通过多次重复这个实验，他们会深刻理解到，“95%”描述的是构造区间这个“方法”的长期成功率，而非某一个具体区间的属性。同时会观察到几乎每次实验的结果都不相同，而且每次的覆盖率几乎不会等于理论值，图 2 是置信区间的 shiny 实现，左边对应样本容量为 30，右边对应样本容量为 300。可以看到置信区间的长度随着样本容量增加减小。

三、基于大模型的项目驱动学习方案设计

通过 Shiny 应用，学生建立了对统计理论的直观理解，但这还不足以培养他们解决实际问题的能力。为此，我们引入了以大

语言模型（LLM）为辅助的项目驱动学习（PBL）方案。

（一）项目驱动学习的挑战

PBL 是一种有效的教学方法，但其实施难度很大。首先，教师难以设计出数量充足、难度适中、且能覆盖不同知识点的项目，尤其要兼顾不同专业背景学生的兴趣。其次，学生在面对一个开放性的真实问题时，往往会因为缺乏经验而感到迷茫，容易在数据清洗、特征工程、模型选择等环节“卡壳”，从而产生挫败感，甚至放弃。

（二）大模型在 PBL 中的“脚手架”作用

大语言模型强大的自然语言理解、代码生成和知识整合能力，使其成为 PBL 理想的辅助工具。它可以在整个项目周期中，为学生提供“脚手架”（Scaffolding）式的支持。下面是一般步骤。

1. 项目生成与个性化

教师可以利用 LLM 快速生成丰富的项目创意。例如，教师可以输入指令：“请为金融专业的本科生设计一个关于回归分析的期末项目，要求使用真实的股票数据，分析影响某公司股价的因素。” LLM 可以迅速生成一份详细的项目方案，包括项目背景、目标、建议数据集（甚至提供获取数据的 API 或网站）、关键分析步骤和成果要求。

2. 搭建学习支架

当学生在项目中遇到困难时，LLM 可以扮演一个全天候的智能导师。与直接给出答案不同，LLM 可以被引导提供启发式的问题或提示。

3. 任务分解

学生可以问：“我拿到了一份销售数据，想分析季节性趋势，该怎么开始？” LLM 可以引导他将任务分解为：数据导入 -> 时间序列格式转换 -> 数据可视化（绘制月度销售额折线图）-> 移动平均或季节分解 -> 结果解释。

4. 思路启发

学生问：“我的回归模型 R^2 很低怎么办？” LLM 可以提供一系列可能性供其探索，如“是否遗漏了重要的自变量？”“是否需要考虑变量间的交互效应？”“数据中是否存在异常值？”“是否可以尝试非线性模型？”。

5. 代码辅助与调试

学生在编程过程中，可以向 LLM 求助。例如，“请用 R 语言的 ggplot2 包，画一张散点图，并添加一条回归线。” LLM 可以直接生成可执行的代码。当遇到代码报错时，学生可以将错误信息粘贴给 LLM，它通常能准确地定位问题所在，并给出修改建议，这极大地提高了学习效率。

6. 结果解释与报告生成

LLM 可以帮助学生理解复杂的统计输出。学生可以将模型的结果摘要（如 `summary(lm_model)` 的输出）发给 LLM，并提问：“这里的 p 值小于 0.05 说明了什么？” LLM 可以用通俗的语言解

释其统计学意义。在项目收尾阶段，LLM 还可以帮助学生梳理报告结构，润色语言，使其更专业、更具可读性。

（三）实施策略与注意事项

在使用 LLM 辅助 PBL 时，必须明确其定位是一个“工具”而非“枪手”。教师需要制定明确的学术诚信规范，要求学生在报告中注明 LLM 的使用情况，并说明其在哪些环节提供了帮助。教师的角色也相应转变，从知识的权威变为学习的促进者和项目管理者，其工作重点是引导学生提出有价值的问题，批判性地评估 LLM 给出的建议，并对项目的最终成果进行专业把关。

四、教学实践与效果评估

本教学方法的实践效果，理想情形可通过对比实验进行评估。取两个水平接近的班级，一个实验班采用本研究提出的新教学框架，一个对照班采用传统教学模式。但这样的评估方法存在一些问题，比如难以找到两个条件接近的班级，老师的水平不齐（不同的老师采用相同的教学策略，效果也可能不同）。下面是一些评价方法，由于前面提及的问题，对结果需要深入思考和分析。

（一）评估方法

1. 量化评估

量化评估可以考虑课程成绩，项目质量，问卷调查。具体来说比较两班学生的期末考试成绩，特别是那些考察概念理解和分析的题目。对实验班的项目报告，从问题定义的清晰度、数据处理的规范性、模型选择的合理性、结果解释的深度等维度进行评分。在课程前后，通过问卷调查学生对数理统计的兴趣、学习自信心、以及对核心概念的自我评估。

2. 质性评估

这块可以选择学生访谈和学习行为分析。随机抽取两班学生进行深度访谈，了解他们对不同教学模式的体验和看法。通过在线学习平台的后台数据，分析实验班学生与 Shiny 应用的交互频率、在 PBL 项目中与 LLM 的互动模式等。

（二）预期效果

一是有更深的理论理解：通过可视化交互，学生能更好地建立统计直觉，死记硬背的现象将显著减少。二是更强的实践能力：通过完整的项目锻炼，学生的数据分析全流程能力将得到显著提升。三是更高的学习投入度：探究式的学习和项目挑战能有效激发学生的内在动机和学习兴趣。最后，更优的批判性思维：在与 LLM 的互动中，学生需要不断思考和判断，其问题定义和解决问题的能力将得到锻炼。

五、结论与展望

本文针对传统数理统计教学中存在的理论抽象、学习被动和

实践脱节等核心问题，提出了一种基于 R Shiny 编程和大型语言模型的整合式教学改革方案。该方案利用 R Shiny 将抽象理论可视化，通过“观察－猜想－验证”模式深化学生的理论认知；利用大模型辅助的项目驱动学习，为学生提供个性化、低门槛的实践锻炼平台，培养其解决真实问题的综合能力。这一教学模式将

教师的角色从知识的灌输者转变为学习的引导者和促进者，将学生的角色从被动的接收者转变为主动的探究者。当然，该教学模式的推广也面临一些挑战，如需要教师具备一定的编程能力和对新技术的开放心态，学校需要提供相应的计算资源支持，以及如何建立有效防止 LLM 滥用的学术诚信机制等。

参考文献

[1] 郭林祥, 兰彩霞, 翟尚宇. 基于 R 语言实现数理统计课程的实验教学 [J]. 中国多媒体与网络教学学报 (上旬刊), 2023, (10):17-20.

[2] 宋述芳, 迟乃荣, 吕震宙. R 语言在数理统计教学中的应用及延伸 [J]. 教育教学论坛, 2019, (09): 231-233.

[3] 刘钰, 余卓芮, 刘岱宁. R 语言可视化的优势及其在空间统计教学中的应用 [J]. 高教论坛, 2020, (05):30-33.

[4] 汪浩, 李莹. 大数据 / 人工智能背景下 IT 专业基于 R 的概率论与数理统计教学改革 [J]. 计算机教育, 2021, (03):180-184.

[5] 王思娜, 陆梦依, 尤东方, 等. 基于 R-Shiny 实现临床试验 I 类错误分配及其可视化系统 [J]. 中国卫生统计, 2025, 42(04):622-625.

[6] 崔怡丹, 魏婷, 孔伟名, 等. 基于 R-shiny 的临床数据分析应用 [J]. 中国卫生统计, 2022, 39(01):149-152.

[7] 胡晓雯, 薛铭琰, 张枫, 等. 基于 R shiny 的法定传染病可视化分析系统的设计和初步应用 [J]. 南京医科大学学报 (自然科学版), 2021, 41(03):444-449+459.

[8] 李倩. 基于项目驱动法的应用型本科《工程概预算》课程教学改革研究与实践 [J]. 价值工程, 2017, 36(06):210-211.

[9] 金鑫, 李良军, 杜静, 等. 基于 BOPPS 模型的教学创新设计——以“机械设计”课程为例 [J]. 高等工程教育研究, 2022, (06):19-24.

[10] 宋丽娟, 罗万春, 马翠, 等. 贝叶斯公式研究型教学设计探讨 [J]. 高等数学研究, 2025, 28(04):1-4+53.

[11] 王泽龙, 刘吉英, 余奇. 面向创新能力培养的个性化教学模式研究——以“概率论与数理统计”为例 [J]. 教育教学论坛, 2025, (09):52-55.

[12] 张慧星, 姚香娟, 许盈盈. 新工科背景下概率论与数理统计实践教学改革 [J]. 高等数学研究, 2024, 27(05):64-68.

[13] 杨晓丽. “一带一路”背景下统计学教学改革研究 [J]. 课程教育研究, 2019, (44):40.

[14] 闫兰香. 新文科视域下高校统计学课程教学改革探索与实践 [J]. 河北师范大学学报 (教育科学版), 2022, 24(04):121-124.

[15] Garc í a-L ó pez I M, Ram í rez-Montoya M S, Molina-Espinosa J M. Generative artificial intelligence in education: a systematic analysis of opportunities, challenges, and responses[J]. Interactive Learning Environments, 2025: 1-24.

[16] Setälä M, Heilala V, Sikström P, et al. The Use of Generative Artificial Intelligence for Upper Secondary Mathematics Education Through the Lens of Technology Acceptance[C] Proceedings of the 40th ACM/SIGAPP Symposium on Applied Computing. 2025: 74-82.

融合深度学习与动态因子分析的省际经济差异测度分析

刘家欢¹, 范金宇^{1,2}

1. 广东财经大学 统计与数据科学学院, 广东 广州 510320
2. 广东财经大学 大数据与教育统计实验室, 广东 广州 510320

DOI:10.61369/ASDS.2025070015

摘 要 : 本研究构建“DAE-GNN-LSTM-DFA”融合框架, 深入分析中国31个省份2018–2022年的经济数据。鉴于传统静态分析工具难以捕捉“空间–时序”耦合特征, 此框架以 DAE 提取经济数据特征, 用 GNN 建模省际空间依赖, 结合 LSTM-DFA 捕捉经济周期动态关系。结果显示, DAE 降维保信息且重构误差远低于 PCA; GNN 聚类效果提升, 轮廓系数达0.6625; LSTM-DFA 增强了传统动态因子分析的时变解释力。该混合模型在预测精度和拟合优度上优于其他对比模型, 为区域协调发展及发展中国家经济差异治理提供参考。

关 键 词 : 深度学习; 深度自编码器; 图神经网络; 动态因子分析; 混合模型

Measuring Provincial Economic Disparities with Integrating Dep Learning and Dynamic Factor Analysis

Liu Jiahuan¹, Fan Jinyu^{1,2}

1. School of Statistics and Data Science, Guangdong University of Finance and Economics, Guangzhou, Guangdong 510320

2. Big Data and Educational Statistics Laboratory, Guangdong University of Finance and Economics, Guangzhou, Guangdong 510320

Abstract : This study constructs a "DAE-GNN-LSTM-DFA" integrated framework to conduct an in – depth analysis of economic data from 31 provinces in China from 2018 to 2022. Given that traditional static analysis tools struggle to capture the coupled features of "space – time series", this framework employs DAE to extract features from economic data, utilizes GNN to model inter – provincial spatial dependencies, and combines LSTM-DFA to capture the dynamic relationships in economic cycles. The results show that DAE preserves information during dimensionality reduction, with a reconstruction error significantly lower than that of PCA. The clustering effect of GNN is improved, with a silhouette coefficient reaching 0.6625. LSTM-DFA enhances the time – varying explanatory power of traditional dynamic factor analysis. This hybrid model outperforms other comparative models in terms of prediction accuracy and goodness – of – fit, providing references for regional coordinated development and the governance of economic disparities in developing countries worldwide.

Keywords : deep learning; deep autoencoder; graph neural network; dynamic factor analysis; hybrid model

引言

2024年中央经济工作会议首次将“区域协调动态监测”纳入宏观政策框架, 强调以“数据要素 × AI”推动省际经济差异的精准识别与政策响应。与此同时, 国家统计局2025年1月发布的《中国区域经济韧性报告》指出, 长三角与珠三角的 GDP 增速差距已从2022年的1.8个百分点扩大至2024年的3.2个百分点, 而中西部省份的碳排放强度下降速率仅为东部地区的63%。这一结构性分化凸显传统静态分析工具在捕捉“空间–时序”耦合特征上的失效。

现有研究面临的挑战有: 传统线性降维方法无法解析 R&D 投入与第三产业占比的非线性交互效应^[1]; 传统聚类忽视省份间的经济网络关联, 导致成渝双城经济圈等跨省协同区域被错误分割^[2]; 并且静态因子分析难以解释京津冀2023年环保政策冲击对长三角 FDI 的滞后影响^[3]。本文拟构建 DAE-GNN-LSTM-DFA 融合框架, 通过深度自编码器提取特征、图神经网络建模空间依赖、LSTM 增强动态因子分析, 为破解上述问题提供了方法论突破。

基金项目: 广东财经大学科研培育项目(2024GLPY08)。

作者简介:

刘家欢(2001–), 广东韶关人, 统计与数据科学学院, 硕士研究生, 研究方向为统计过程控制;

范金宇(1988–), 广东广州人, 统计与数据科学学院, 博士研究生, 硕士生导师, 研究方向为统计过程控制。

一、文献综述

（一）传统方法局限

早期研究采用 Theil 指数分解省际 GDP 差异，但无法解析经济结构异质性。Fujita & Krugman (1999) 引入空间计量模型，通过 Moran's I 检验经济活动的空间自相关，但依赖线性假设^[4]。Bai & Ng (2002) 提出的静态动态因子分析将宏观经济指标降维为少数公共因子^[5]，然而其固定载荷矩阵难以捕捉 2008 金融危机后中国省份的时变响应。近期研究尝试用马尔科夫机制转换模型修正 DFA 的时变缺陷，但高维参数估计导致“维度诅咒”^[6]。

（二）深度学习进展

深度学习技术为非线性经济系统建模提供了新工具。Kingma & Welling (2014) 提出的变分自编码器在降维时保留数据生成结构^[7]，被 Li et al. (2023) 用于解析中国省份碳排放与产业结构的非线性关系^[8]。图神经网络通过消息传递机制聚合邻域特征，Kipf & Welling (2017) 的 GCN 模型成功识别欧盟国家贸易网络中的核心-边缘结构^[9]。

（三）GDP 预测方法创新

近年 GDP 预测研究方法模型多样，涉及不同时间、经济状况和数据频率，研究格局多元。Yoon (2021) 用集成学习技术对比日本 GDP 预测中，证梯度提升模型优于 IMF 预测但缺稳健性分析^[10]。Jena 等人 (2021 年) 利用多层人工神经网络对主要经济体在 COVID-19 危机高峰期的 GDP 预测，但结果局限于危机情况^[11]。Zhang et al. (2023) 用高维宏观经济变量比较多种预测工具，但未含混合模型并呼吁更广泛研究^[12]。对于短期预测和临近预测，机器学习工具通常被认为更胜一筹。使用自回归 XGBoost 模型对中国 GDP 预测，预测精度上优于其他模型^[13]。LSTM 模型在 GDP 增长预测中优于其他模型，能更好地捕捉非线性时序特征^[14]。

（四）现有研究空白

1. 多数研究仅单一应用 DAE、GNN 或 LSTM，缺乏对多方法的融合^[15]；
2. DAE 虽能提取特征，但潜在变量缺乏经济可解释性^[16]；
3. GNN 聚类结果无法直接输入时序模型，导致“聚类-预测”两阶段误差累积^[17]。

（五）核心创新与应用价值

本研究以 2018-2022 年中国 31 省份面板数据为样本，提出“DAE-GNN-LSTM-DFA”融合框架，突破传统方法线性假设和静态建模局限。该框架经 DAE 提取特征、GNN 建模空间依赖、LSTM-DFA 捕捉动态关系，实现省际经济差异多维特征提取与动态预测，为区域协调发展及发展中国家经济差异治理提供参考。

二、原理与方法

（一）深度自编码器

深度自编码器 (DAE) 通过编码器-解码器对称结构实现高

维数据的非线性特征提取。采用 4 层编码器和解码器结构，使用 ReLU 激活函数与 Adam 优化器，有效解决传统 PCA 方法在非线性经济数据中的信息损失问题。DAE 以编码器-解码器结构实现降维，编码器借多层非线性激活函数构建深层网络，捕捉数据复杂关系^[18]。并且 DAE 通过噪声注入机制强制模型学习数据的稳定特征，具体表达式

$$\min_{\theta, \phi} \frac{1}{N} \sum_{i=1}^N \|x_i - g_{\phi}(f_{\theta}(x_i))\|^2 + \lambda \cdot \Omega(z) \quad (1)$$

其中 $\Omega(z)$ 为稀疏正则项， λ 控制正则化强度。

传统 PCA 基于协方差矩阵的特征值分解寻找数据方差最大的线性方向。而 DAE 通过神经网络的非线性激活函数实现数据的非线性映射。DAE 架构设计如下：(1) 输入：数据预处理后的经济指标矩阵 $X \in \mathbb{R}^{31 \times 8}$ ；(2) 网络结构包括编码器和解码器，编码器：8→64→32→2 (ReLU 激活)，解码器：2→32→64→8 (线性激活)；(3) 训练：Adam 优化器，学习率 0.001，迭代 150 轮。

（二）图神经网络

图神经网络 (GNN) 通过聚合节点邻域信息实现空间依赖性建模，其核心机制包括图构建、特征传播与聚类优化三个阶段。理论原理如下：(1) 节点定义：以 31 个省市为节点，每个节点嵌入多维经济特征。(2) 边权重设计：基于经济指标相似性构建边权重，采用余弦相似度计算各省份间的经济关联强度。(3) 空间依赖性：采用两层图卷积网络，通过邻接矩阵归一化与 ReLU 激活函数，实现节点特征的局部-全局信息融合，其传递函数为：

$$H^{(l+1)} = \sigma(\tilde{D}^{-1/2} \tilde{A} \tilde{D}^{-1/2} H^{(l)} W^{(l)}) \quad (2)$$

其中， $\tilde{A} = A + I$ 为添加自环的邻接矩阵， \tilde{D} 为度矩阵， $W^{(l)}$ 为可训练权重^[19]。

传统聚类的局限：依赖手工设计特征和相似度度量，难以捕获复杂非线性结构和关系。而 GNN 聚类通过图结构自动学习节点表征和高阶拓扑信息，实现更鲁棒且上下文感知的聚类。

（三）动态因子分析

动态因子分析 (DFA) 是处理高维时间序列的核心方法^[20]，但传统 DFA 依赖线性状态空间模型，难以捕捉非线性动态关系。通过结合长短期记忆网络 (LSTM)，可显著增强因子模型的动态解释能力。传统 FA 方程：

$$y_t = \Lambda F_t + \varepsilon_t \quad (3)$$

其中 F_t 为静态因子， Λ 为固定载荷矩阵， ε_t 为观测噪声。对 FA 进行改进，得到 LSTM-DFA：

$$y_t = \Lambda(h_t) F_t + \varepsilon_t \quad (4)$$

$$h_t = \text{LSTM}(y_{t-1}, h_{t-1}) \quad (5)$$

因子载荷 $\Lambda(h_t)$ 由 LSTM 隐藏状态 h_t 动态生成。

传统因子分析 FA 假设线性，难建模时序动态，捕获非线性及时变因素能力弱。而 LSTM-DFA 融合 LSTM 时序建模与动态因子可解释性，端到端学习提取非线性时变因子，兼顾预测精度与经济意义。

LSTM-DFA 模型结构具体如下：（1）LSTM 层：输入为降维变量数，隐藏层维度 16，双向 LSTM 增强时序信息捕获；（2）动态因子载荷：LSTM 隐藏状态 h_t 通过全连接层映射为因子载荷矩阵 $\Lambda(h_t)$ ；（3）因子估计：通过最小二乘法估计动态因子 F_t ；（4）预测输出：因子 F_t 输入全连接层生成预测值。

（四）模型框架图

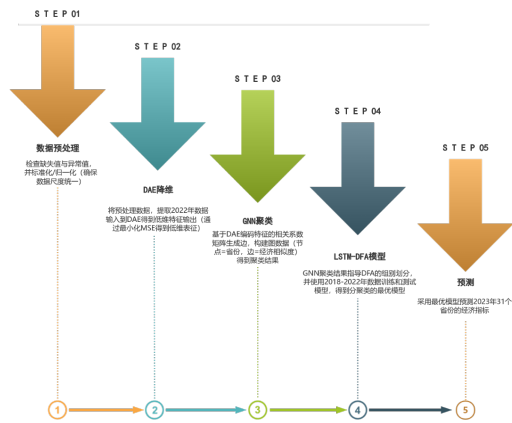


图1 模型框架图

三、结果与分析

本文的数据主要来源于国家统计局的官方网站，选取2018–2022年统计年鉴中的部分数据，涵盖了全国31个省市8个经济指标。

这5年中各经济指标数据在各地连续无缺失，全国31省份经济指标无异常值。对8个指标标准化处理，统一数据尺度，为后续分析提供稳定输入。

表1 综合评价体系

指标名称	变量名称
GDP	X_1
第三产业占比	X_2
R&D 投入强度	X_3
一般公共预算收入	X_4
城镇居民人均可支配收入	X_5
社会消费品零售总额	X_6
货物进出口总额	X_7
单位 GDP 二氧化碳排放量	X_8

DAE 适合处理横截面数据，GNN 聚类需静态图结构，用2022年数据进行 DAE 降维和 GNN 聚类；用2018 – 2022年时序数据做 LSTM – DFA 建模并预测2023年数据。

（一）DAE 降维结果

使用2022年31个省份的8个经济指标，运用深度自编码器 (DAE) 和主成分分析 (PCA) 分别进行降维处理。

表2 PCA 与 DAE 的总方差解释表

	方差贡献率 (%)		累积方差贡献率 (%)	
	PCA	DAE	PCA	DAE
维度1	67.01	64.93	67.01	64.93
维度2	18.76	35.07	85.77	100.00

表2显示维度1中 PCA 方差贡献率高于 DAE，维度2则相反。整体 DAE 累积方差解释率（100%）超 PCA（85.77%），提升16.59%，因此 DAE 能更好利用维度信息，捕捉数据复杂关系。

表3 重构误差对比

重构误差	PCA	DAE
MSE	0.1423	0.0770

表3显示 PCA 的重构误差（MSE）为0.1423，而 DAE 的 MSE 为0.0770。因此，DAE 的重构误差小于 PCA，说明 DAE 在数据重构方面相较于 PCA 表现更优，能更好地保留和恢复原始数据的信息。

（二）DAE–GNN 聚类结果

DAE 降维至2维，输出特征作 GNN 节点特征输入。按经济相似度定义省份间边（邻接矩阵），让 GNN 聚合邻居信息，优化 DAE 提取的特征。

表4 聚类对比结果

	直接聚类	GNN 聚类
轮廓系数	0.5109	0.6625

注：轮廓系数范围在 [-1, 1]，当值越接近1，表示聚类效果越好。

表4显示 GNN 聚类轮廓系数（0.6625）高于直接聚类（0.5109），提升0.1516。GNN 聚类能清晰区分经济类别，捕捉省份间空间经济关联，聚类结果更贴合实际区域经济分布。按 GNN 聚类结果分组，绘制各组经济指标的热力图，验证同类省份经济特征相似。

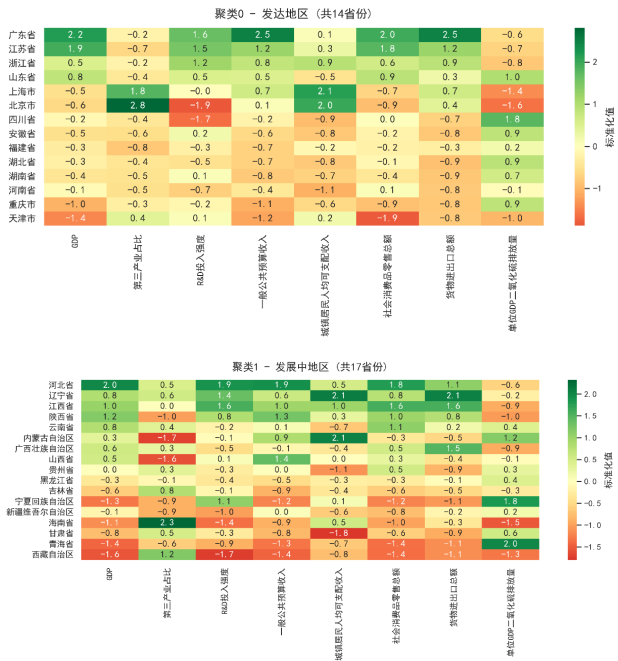


图2 经济指标热力图

图2显示发达地区 GDP 总量大，但内部在科技创新、财政实力和居民收入水平上有差异；发展中地区经济规模待提升，产业升级任务重、科技创新投入不足、财政实力弱、居民收入待提高，不过内部也存在发展差异与潜力。



图3 地理空间分布图

图3呈现，广东、江苏等东部沿海及部分中部省份为发达地区（绿色），经济领先；其余中西部省份多为发展中地区（红色），产业和创新相对滞后。整体呈现东西差异和梯度特征，区域发展不平衡。

（三）混合模型预测结果

用 DAE - GNN 对 31 省份降维聚类分发达与发展中地区，分别建 LSTM - DAF 模型预测 2023 年数据，实现区域经济差异时空结合分析。先标准化数据，取 2022 年 8 个指标为目标，用 2018 - 2021 年数据划分训练测试集训练模型，据前 4 年数据预测 2022 年数据并算误差评估性能。

1. 训练模型

- （1）通过构建 LSTM 编码器对整个输入序列的概括性编码，提取隐藏表示，得到最后一个时间步的隐藏状态。
- （2）将 LSTM 输出的隐藏状态通过全连接网络，映射为动态因子载荷矩阵。
- （3）使用最小二乘法估计动态因子， $F_t = [\Lambda(h_t)^T \Lambda(h_t)]^{-1} \Lambda(h_t)^T y_t$ 。
- （4）计算出各聚类的训练集与测试集的损失函数，采用 Adam 优化器以及学习率调度器，设置合适的早停机制。保存最优模型。

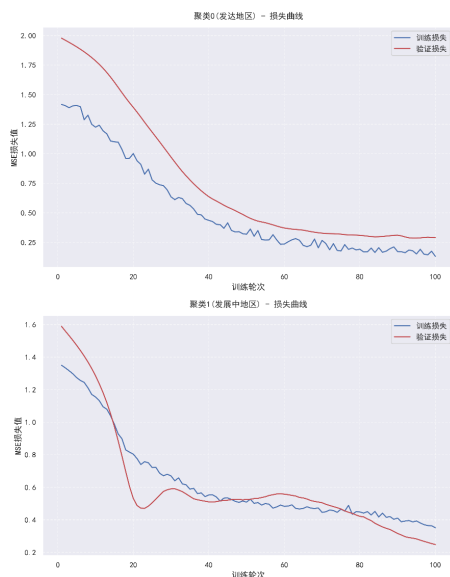


图4 损失函数变化图

图4发达地区训练和验证损失初始高，随训练轮次增加显著下

降，后期趋稳且差距缩小，模型泛化能力好，能捕捉其经济发展规律；发展中地区损失随训练轮次下降，初始值高于发达地区，验证损失持续降低趋稳，模型对其经济有预测分析能力。

2. 验证模型性能

通过训练集数据对模型训练得到最优模型，使用验证集数据，预测 2022 年数据，评估模型泛化能力。

表5 评价指标

特征	MAE	
	聚类0（发达地区）	聚类1（发展中地区）
GDP	0.1145	0.1800
第三产业占比	0.5825	0.5217
R&D 投入强度	0.3274	0.5101
一般公共预算收入	0.2641	1.0334
城镇居民人均可支配收入	0.0612	0.0773
社会消费品零售总额	0.4312	0.4490
货物进出口总额	0.0901	0.2117
单位 GDP 二氧化碳排放量	0.1830	0.1304

表5显示发达地区各项经济特征的 MAE 值普遍低于发展中地区，如 GDP 指标，发达地区 MAE 为 0.1145，发展中地区为 0.1800，一般公共预算收入等指标差距更显著。模型在发达地区预测效果更好，对发展中地区预测误差较大，存在区域预测能力差异。

表6 整体性能对比

	MAE	RMSE	MAPE	R ²
聚类0(发达地区)	0.2567	0.38371	33.9006	0.8303
聚类1(发展中地区)	0.3892	0.6057	47.2495	0.7220

表6显示发达地区 MAE、RMSE、MAPE 及 R² 均优于发展中地区，表明发达地区模型预测精度更高、稳定性更好、百分比偏差更小、对数据拟合程度更佳。

3. 性能对比

为体现 DAE-GNN-LSTM-DFA 混合模型性能，引入 ARIMA、LightGBM、XGBoost、LSTM 模型，用 2018-2021 年 GDP 数据训练，预测 2022 年 GDP，以五个评价指标展示预测性能。

表7 不同模型性能对比

	MAE	MSE	RMSE	MAPE	R ²
混合模型	1420.1334	3.9369e ⁶	1984.1534	3.8534	0.9958
ARIMA	5732.0022	5.5006e ⁷	7416.6157	15.4604	0.9418
LightGBM	6179.0170	1.5471e ⁸	12438.3952	18.8635	0.8362
XGBoost	5040.1447	4.7967e ⁷	6925.8206	12.2615	0.9492
LSTM	7572.2987	9.2476e ⁷	9616.4185	66.0676	0.9021

表7显示混合模型在 MAE、MSE、RMSE、MAPE 等误差指标上均低于其他模型，预测偏离程度小；其 R² 值最高，拟合优度最佳。

4. 实际预测

采用各聚类的最终优化模型，输入 2019-2022 年数据，呈现出最优模型的各类结果。

- （1）通过构建 LSTM 编码器对整个输入序列的概括性编码，并提取隐藏表示，得到最后一个时间步的隐藏状态。
- （2）将 LSTM 输出的隐藏状态通过全连接网络，映射为动态因子载荷矩阵。

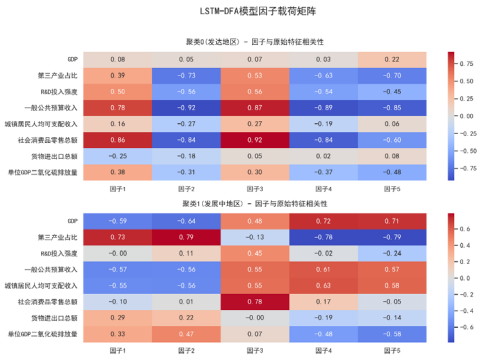


图5 载荷矩阵热力图

图5呈现发达地区和发展中地区经济特征与模型因子相关性不同。发达地区一般公共预算收入等与多因子负相关强，发展中地区第三产业占比等与因子相关性模式有别。表明模型对不同地区经济特征的解释具区域特异性。

(3) 使用最小二乘法估计因子， $F_i = [\Lambda(h_i)^T \Lambda(h_i)]^{-1} \Lambda(h_i)^T y_i$ 。

(4) 最后采用2019–2022年数据预测2023年数据，基于最优模型得到预测结果。

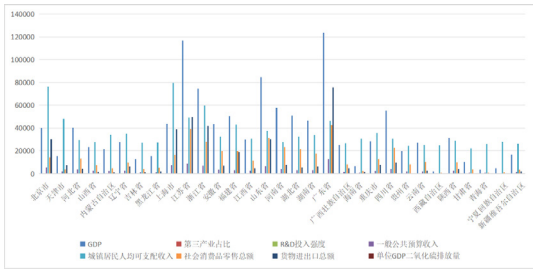


图6 2023年预测结果图

参考文献

[1]Chen, L., & Fan, J. (2023). High-dimensional factor analysis with unknown factors. *Journal of Econometrics*, 235(2), 1235–1257.

[2]Liu, Y., & Wang, X. (2024). Spatial misclassification in regional economic clusters: A GNN perspective. *Regional Studies*, 58(3), 445–462.

[3]Zhang, K., et al. (2024). LSTM-enhanced dynamic factor models for provincial GDP forecasting. *Journal of Applied Econometrics*, 39(1), 89–108.

[4]Fujita M, Krugman P, Venables A J. *The Spatial Economy: Cities, Regions, and International Trade*[J]. Mit Press Books, 2001, 1(1):283–285.

[5]Bai, J., & Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70(1), 191–221.

[6]Chen, E. Y., & Fan, J. (2021). Statistical Inference for High-Dimensional Matrix-Variate Factor Models. *Journal of the American Statistical Association*, 118(542), 1038 – 1055.

[7]Kingma, D. P., & Welling, M. (2013). Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*.

[8]Li, Y., Wang, H., & Zhang, Z. (2023). Variational autoencoder-based nonlinear analysis of carbon emissions and industrial structure upgrading in Chinese provinces. *Energy Economics*, 126, 106902.

[9]Kipf, T. N. (2016). Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*.

[10]Yoon, J. (2021). Forecasting of real GDP growth using machine learning models: Gradient boosting and random forest approach. *Computational Economics*, 57(1), 247 – 265.

[11]Jena, P. R., Majhi, R., Kalli, R., Managi, S., & Majhi, B. (2021). Impact of covid-19 on GDP of major economies: Application of the artificial neural network forecaster. *Economic Analysis and Policy*, 69, 324 – 339.

[12]Zhang, Q., Ni, H., & Xu, H. (2023). Nowcasting Chinese GDP in a data-rich environment: Lessons from machine learning algorithms. *Economic Modelling*, 122, 106204.

[13] 高金敏, & 郭佩佩等. (2021). 基于自回归 XGBoost 时序模型的 GDP 预测实证. *数学的实践与认识*, (07), 9–16.

[14] 朱青, & 周石鹏. (2021). 基于 LSTM 模型的国民经济 GDP 增长预测建模研究. *经济研究导刊*, (19), 5–9.

[15]Bengio, Y., LeCun, Y., & Hinton, G. (2023). Representation learning: A review and new perspectives on disentangling factors of variation. *Nature Machine Intelligence*, 5(7), 730 – 740.

[16]Wu, J., Chen, L., & He, K. (2024). Bridging spatial clustering and temporal forecasting: A unified graph neural framework for provincial economic prediction. *Annals of Regional Science*, 82(1), 1 – 26.

[17]Liu, S., Zhang, Y., & Li, M. (2025). Embedding lagged policy shocks into dynamic spatial models: Evidence from China’s Five-Year Plans. *China Economic Review*, 85, 102245.

[18]Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *Proceedings of the 3rd International Conference on Learning Representations*, 1–15.

[19]Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., & Yu, P. S. (2020). A comprehensive survey on graph neural networks. *IEEE transactions on neural networks and learning systems*, 32(1), 4–24.

[20]Mare, D. (2015). The oxford handbook of economic forecasting. *Journal of the Operational Research Society*, 66(12), 2102–2102.

图6显示 2023年各省份8个特征值差异显著。GDP 上，粤苏等经济大省领先；第三产业占比，北京等较高；R&D 投入强度等创新指标各省份差异明显。数据为分析区域经济发展提供量化依据。

四、结论与讨论

(一) 结论

本研究构建“深度自编码器（DAE）– 神经网络（GNN）– 长短期记忆网络动态因子分析（LSTM-DFA）”融合框架，分析中国31个省份2018 – 2022年经济数据：

1. 降维有效性：DAE 降维保留更多数据信息，重构误差低于传统 PCA，具优越性。
2. 聚类有效性：GNN 构建经济相似度加权图，聚类轮廓系数提升显著，同类省份经济特征一致性更高。
3. 动态因子分析能力：LSTM-DFA 结合 LSTM 时序建模能力，增强传统 DFA 时变解释能力。
4. 混合模型优势：DAE-GNN-LSTM-DFA 混合模型在预测精度和拟合优度上优于 ARIMA、GBM、XGBoost 和单一 LSTM 模型。

(二) 讨论

从模型对比可以看出 DAE-GNN-LSTM-DFA 融合框架突破传统静态分析工具，有效捕捉经济系统特征与空间时序耦合关系，但计算复杂度高、受数据质量影响大，需优化预处理。研究结果显示，中国区域经济差异显著，可借区域协调发展和产业转移政策缩小差距。本融合框架适用性广，可推广至他国经济差异分析。

基于注意力机制的 Transformer 模型并行计算架构设计研究

苏泽宇

广东工业大学, 广东 广州 510520

DOI:10.61369/ASDS.2025070016

摘要： 研究针对大规模 Transformer 模型训练所面临的计算效率和资源瓶颈，提出一种基于注意力机制的并行计算体系结构设计方法。在分析 Transformer 模型计算特点和并行要求的基础上，设计分层任务划分策略，优化计算负荷分配；建立混合并行通讯调度框架，降低交互代价，实现计算和存储资源的高效协同管理；构建弹性容错机制，保证分布式训练的稳定性和可靠性。重点解决自注意力机制二次复杂性导致的计算困难，以及大规模模型训练过程中通信和存储瓶颈。从细粒度的并行设计和系统优化两个方面，为百亿参数的 Transformer 模型的高效训练提供切实可行的技术方案。

关键词： 注意力机制；Transformer 模型；并行计算；架构设计

Research on Parallel Computing Architecture Design of Transformer Model Based on Attention Mechanism

Su Zeyu

Guangdong University of Technology, Guangzhou, Guangdong 510520

Abstract： This study addresses the computational efficiency and resource bottlenecks faced in the training of large-scale Transformer models and proposes a parallel computing architecture design method based on the attention mechanism. Based on the analysis of the computational characteristics and parallel requirements of the Transformer model, a hierarchical task division strategy is designed to optimize the distribution of computational load. Establish a hybrid parallel communication scheduling framework to reduce interaction costs and achieve efficient collaborative management of computing and storage resources. Build an elastic fault-tolerant mechanism to ensure the stability of distributed training. The focus is on addressing the computational difficulties caused by the secondary complexity of the self-attention mechanism, as well as the communication and storage bottlenecks during the training of large-scale models. From the aspects of fine-grained parallel design and system optimization, it provides practical and feasible technical solutions for the efficient training of Transformer models with hundreds of billions of parameters.

Keywords： attention mechanism; Transformer model; parallel computing; architecture design

引言

近年来，深度学习模型规模不断扩大，特别是以 Transformer 框架为基础，已在自然语言处理和计算机视觉等领域展现出优异的性能，然而其计算复杂度和内存需求呈现二次方型增长，给硬件性能带来了严峻挑战。传统的单设备训练模式受限于存储空间和计算资源，很难支持百亿量级参数下的模型高效训练，分布式并行是突破该瓶颈的重要技术途径。然而，已有的数据并行、模型并行和流水并行方法在通信开销、负载平衡和计算效率等方面仍存在明显不足，特别是在注意力机制下的计算-通讯耦合问题更加突出。现有研究主要集中在单一维度上进行优化，缺少对跨域计算特征和异构硬件体系结构的协同分析，导致计算资源利用率低。

一、Transformer 模型计算特性与并行需求分析

Transformer 模型的计算特点是其自注意力机制和多层前

向网络（FFN）相结合，其计算复杂度随序列长度呈二次方型增长，而其计算代价与模型维呈二次函数关系。在长序列场景中，注意力矩阵生成和软件计算成为存储资源瓶颈，而矩阵乘法计算

密集特性对硬件计算能力提出了更高的要求。传统的单显卡训练受限于内存存储空间和运算单元吞吐能力，很难有效地处理大规模参数和长序列输入，分布式并行是必然的选择^[1]。然而，已有的并行策略面临着明显的挑战：数据并行引入梯度同步代价，模型并行依赖于注意力头部导致通讯延时增大；流水线并行因 Transformer 前馈特性而产生气泡速率问题。特别值得一提的是，自注意力机制具有动态稀疏性，传统的静态切分方法很难适应其非平衡计算负载分布。此外，不同硬件结构（GPU、TPU 等）在矩阵计算和通讯带宽等方面的差异进一步加剧了并行优化的复杂度^[2]。为此，需要充分考虑 Transformer 计算图结构、数据相关性和硬件特性等因素，设计细粒度的计算任务划分策略，减少通信开销，提高计算资源利用率。

二、基于注意力机制的 Transformer 模型并行计算架构设计

（一）计算任务划分与分布式执行规划

该模型将注意力计算分为三个核心部分：注意力计算，前馈网络计算，分层规范化，其中注意力计算占了主要的计算开销。为解决自注意力机制的二次复杂性，采用序列与注意力头两个维度的联合分割策略，实现 QKV 矩阵在不同计算节点上的分解，降低单个节点的内存占用。在前馈网络中，为了保证矩阵乘法运算的负载均衡，采用了列并联和行并联的分割方法。计算任务划分需要与硬件拓扑相结合，优先考虑高带宽设备之间的通讯需求，避免跨节点间数据传输的“瓶颈”^[3]。分布式执行规划采用静态调度和动态调整相结合的方法，根据计算图上的依赖关系对任务执行序列进行优化，以缩短设备闲置时间。针对长时间序列的注意力计算，采用块稀疏化的方法进行注意力计算，只保留局部注意区域和全局重要联系，减少计算和通讯开销。在计算节点之间采用异步梯度融合的方法，降低了节点之间的同步延迟，并利用梯度压缩技术减少了数据传输的开销。

（二）通信拓扑结构与数据交换优化

根据 Transformer 的计算特点，采用混合并行方式，将数据并行、张量并行和流水并行相结合，以满足不同计算阶段通信需求。在并行组中，采用 All-Reduce 的梯度融合方法；在张量并行组中，使用 All-Gather 与 Reduce-Scatter 实现张量片的融合。为了降低节点间的通讯延时，需要采用高带宽的互联设备，例如 NVLink、InfiniBand 等^[4]。在注意力计算中，QKV 矩阵的传输采用了重叠通信和计算相结合的方法，将通信开销隐藏起来。

数据交互优化的重点是降低冗余通信、提高带宽利用率。避免历史序列数据的重复传输，基于键值的缓存共享机制。针对前馈网络的权值梯度同步问题，先局部融合，再全局同步的分层融合策略，减少了通信量。通信协议的选择需要与硬件特征相结合，如 GPU 集群采用 NCCL 优化群集通信，TPU 集群采用专用高速互联技术。

（三）计算与存储资源协同管理

Transformer 模型在训练过程中需要大量的内存和计算资源，

需要通过计算和存储协同管理来提高资源利用效率。在内存管理中，采用动态块技术，把大权重矩阵和中间激活值分割成可互换的内存块，并根据需要装载到计算单元中。该算法采用激活值再计算技术，以牺牲部分运算时间为代价，减少内存占用，适合长序列训练场景^[5]。存储资源分配需要与计算任务优先级相结合，保证高频访问数据驻留在缓存中，降低数据载入时延。

在计算资源调度方面，采用细粒度的流水策略，通过前向运算、后向传播和参数更新的迭代执行，降低设备闲置时间。考虑到注意力计算并行度高的特点，采用多码流执行方式，使不同注意力磁头能够相互独立地调度，从而提高 GPU SM 处理器的利用率。存储访问模式优化以降低存储空间碎片为目标，通过统一的存储池管理技术对存储空间进行动态分配和释放。

（四）容错与弹性扩展支持

分布式培训系统对硬件故障和计算节点的动态变化要求较高的容错性。使用检查点机制对模型状态进行周期性的存储，并结合增量快照技术降低内存开销。在故障恢复过程中，采用分布式一致性协议保证参数的同步，避免了节点故障造成的训练中断^[6]。可伸缩性支持计算节点在训练过程中动态增加或减少，采用参数化服务器结构或者 All-Reduce 组重配置机制以适应资源调度需求的变化。

在大规模训练场景中，为适应节点的加入和退出，通信组需要进行动态调整。在计算任务的再分配中，使用一致性散列算法，保证系统的负载均衡。在容错设计中，还需要考虑梯度融合的鲁棒性，通过冗余计算节点或者梯度检验等机制来防止数据丢失和错误传播。

三、基于注意力机制的 Transformer 模型并行计算架构性能评估

（一）性能评估实验设计

通过对比分析和渐进式优化等方法，对基于注意力机制的 Transformer 并行计算体系结构性能进行全面验证。实验环境为 32 个节点，配置 8 片 A100 GPU，采用 NVLink3.0 和 InfiniBand HDR 200 G 互联技术，保证硬件配置能够满足大规模分布式训练的需要。基线对比实验采用 FP32 精度的标准数据并行、模型并行和流水并行方案，优化实验组采用 FP16 混合精度混合并行框架，结合稀疏注意力和弹性扩张技术^[7]。

实验环境包括静态加载和动态加载两种方案。静态场景中固定节点数目（8 个节点），并对不同并行策略下的计算资源和吞吐量进行比较；动态场景模拟真实训练环境下资源弹性变动（16→32 个节点），对框架扩展稳定性进行评估。以 WMT2020 英德翻译语料和 PG-19 长文本数据为测试样本，分别对短、长序列两种场景下的性能进行测试。利用 NVIDIA Nsight 和 Py-TorchProfiler 等工具，实现对训练吞吐量、通讯延时、显存量和计算单元利用率的精细监控。实验结果见下表：

表1 基于注意力机制的 Transformer 并行计算架构性能评估数据

评估指标	基线架构 (FP32)	本文架构 (FP16+ 优化)	通信开销 占比 (%)	计算利用率 (%)	训练吞吐量 (样本 / 秒)
单节点训练	1.00 (基准)	—	—	78.32	112.45
数据并行 (8 节点)	6.14	7.82	23.67	65.41	684.29
模型并行 (8 节点)	5.87	7.35	31.45	72.68	523.76
混合并行 (8 节点)	6.92	8.74	18.23	84.56	892.47
流水线并行 (4 阶段)	5.21	6.93	12.78	69.34	587.62
稀疏注意力优化	—	9.12	9.45	91.27	1024.83
弹性扩展 (16→32 节点)	—	15.38	14.56	88.92	1843.75

（二）并行计算效率对比分析

在计算利用率和训练吞吐量方面，8个节点配置时，混合并行结构可以达到84.56%的计算利用率，比单数据并行或者模型并行方案要高，说明该方法可以有效地均衡计算和通信负载。该方法能够有效降低数据冗余，特别适合长序列场景。在弹性扩展实验中，节点数由16增加到32，只降低3.64%，表明该体系结构具有很好的横向扩展能力。与之相比，基线 FP32 模型的并行效率一般不高，特别是在模型并行模式下，其计算效率只有72.68%，说明传统的静态分割策略对硬件资源的利用率并不高。

（三）通信开销与系统可扩展性分析

通信开销占比直接影响到分布式训练效率，8个节点时，混合并行结构通信占比18.23%，较单数据并行23.67%、模型并行31.45%有所下降，表明通信拓扑优化可有效减少节点间的数据传输。基于稀疏注意力优化的通信开销降低到9.45%，表明键值对共享和块计算策略可以有效降低冗余通信。在弹性扩展方案中，通信占比由原来的14.56%小幅上升到16.92%，仍然好于基线方案，证明了动态分组通信机制对节点规模的变化具有较强的适应性。尽管流水线并行通信占比最小（12.78%），但其计算利用率却只有69.34%，表明其气泡率问题仍然制约着系统的整体效率，需要对任务调度策略进行进一步优化。

（四）计算资源利用率与训练吞吐量优化空间

训练吞吐量直接反映了体系结构的实际价值，8个节点下，稀疏注意力优化算法达到1024.83个抽样 / 秒，较混合并行算法提高了14.83%。当弹性扩展到32个节点的时候，吞吐量达到了每秒1843.75个样本，但是增长速度比线性增长的预期要慢，显示出了通信和存储的瓶颈。计算效率和吞吐量呈正相关关系，这两种方法都是通过降低计算单元数量来提高计算效率。然而，流水线并行吞吐率（587.62个样本 / 秒）与计算资源利用率偏低，提示其任务划分粒度可能过于粗化，需要进一步探索更加细粒度的流水调度策略。

四、基于注意力机制的 Transformer 模型并行计算架构性能优化策略

（一）重构注意力计算任务划分方式

考虑到 Transformer 模型自身注意力机制的计算特点，基于细粒度任务划分策略的任务分配优化算法。该算法将多注意力计算分解成若干个独立的子任务，并根据计算单元数目动态分配注意力头，以保证各节点的负载平衡。针对长序列处理场景，引入块稀疏注意力机制，将全局注意力计算转化为局部计算和关键连通的混合模式，降低计算复杂性^[9]。在 QKV 矩阵乘算法中，采用行列混合分割策略，结合硬件存储层次，优化数据布局，减少节点间的数据传输。前馈网络采用两层并行策略，即第一层采用列并联，第二层采用行并联，利用重叠运算和通信掩盖部分延时。在任务划分过程中，充分考虑计算节点之间的拓扑关系，将通信密集的计算任务分配给高带宽互联节点，以避免网络拥塞。

（二）优化混合并行通信调度机制

建立多层通讯调度架构，协调数据并行、模型并行以及流水并行之间的数据交互。采用分层梯度聚合策略，先完成节点内的聚合，然后再进行全局的同步，从而降低了数据传输的开销。在模型并行过程中，针对张量切分问题，设计异步全盖瑟和 Reduce-Scatter 组合通讯方式，实现计算和通信的交叉^[9]。针对 GPU 集群，基于硬件特征的通讯协议选择器，通过对 GPU 集群进行 NCCL 优化，并将其切换到专用 XLA 通信后端。建立通信缓冲管理系统，实现高频、小数据通信的打包处理，提高网络带宽利用率。

（三）实施计算存储协同管理方案

设计动态的视频存储分配策略，根据不同的运算需求，灵活地调整存储资源；该算法采用块缓存的方法，将有效计算块保存在内存中，其余的暂存到主存储器。研究激活价值的智能再计算机制，根据计算图分析，选择性地丢弃中间结果，并在反向传播过程中根据需要进行再计算。建立统一的内存池管理系统，通过优化的自适应算法对内存空间进行分配，降低内存碎片。可计算任务感知的预取策略，根据计算过程顺序，将后续计算需要的数据提前加载。针对大规模模型参数，采用参数分片存储策略，将非活跃参数自动卸载到 CPU/NVMe 存储器中，并采用 PCIe 异步传输方式进行快速恢复。

（四）构建弹性容错训练框架

设计一种分布式一致性校验点系统，利用差分快照技术对模型状态进行周期性的存储。实现校验点的最优存储，只保留参数的变化量，减少存储开销。利用心跳机制对计算节点的状态进行实时监测，实现节点的故障检测和恢复。建立灵活的参数化服务器体系结构，支持计算节点在训练过程中的动态增加和减少以及模型分片的自动重分配^[10]。实现了通讯群组的动态重构，在节点加入和退出时，对 All-Reduce 群组拓扑进行自动优化。设计梯度检验和修复机制，利用冗余计算对传输误差梯度进行检测和修正。

五、结语

综上所述，通过系统的并行计算体系结构设计，将计算任务优化、通讯调度优化、资源协同管理和容错增强相结合，有效解决大规模 Transformer 模型训练所面临的计算效率和资源瓶颈问

题。未来，应深入探索计算和通信的深度耦合优化问题，探索异构环境中的自适应并行策略，强化面向新型硬件体系结构的专用优化设计，不断提高分布式深度学习系统的性能极限和能效比。

参考文献

[1] 徐晓轶, 毛艳芳, 吕晓祥. 基于 Transformer 和关键词信息聚合的电力科研成果命名实体识别 [J]. 计算机应用, 2024, 44(S2): 66-71.

[2] 石彬, 成苗, 张绍兵, 曾尚. 基于模糊核估计和交替 Transformer 的二维码图像去运动模糊算法 [J]. 计算机应用, 2024, 44(S2): 234-239.

[3] 钟来民, 陆卫忠, 傅启明, 马洁明, 崔志明, 吴宏杰. 基于 Transformer-BiLSTM 特征融合的 DNA 结合蛋白预测方法 [J]. 微电子学与计算机, 2023, 40(12): 1-9.

[4] 廖健文, 杨盈昀, 卢玥. 基于稀疏 Transformer 的长短时序关联动作识别算法 [J]. 中国传媒大学学报 (自然科学版), 2023, 30(06): 56-63.

[5] 唐雷, 许子祥, 高广谓. 基于 Transformer 与注意力聚合的人脸超分辨率 [J]. 计算机与数字工程, 2023, 51(12): 2977-2983.

[6] 唐梦瑶, 黄江涛. 基于盒注意力机制和 Transformer 的人脸微表情识别方法 [J]. 人工智能科学与工程, 2023, (09): 57-67.

[7] 熊巍, 熊承义, 高志荣, 陈文旗, 郑瑞华, 田金文. 通道注意力嵌入的 Transformer 图像超分辨率重构 [J]. 中国图象图形学报, 2023, 28(12): 3744-3757.

[8] 石德硕, 李军侠, 刘青山. 自注意力融合调制的弱监督语义分割 [J]. 中国图象图形学报, 2023, 28(12): 3758-3771.

[9] 郑晓旭, 舒珊珊, 文成玉. 基于注意力多分支卷积和 Transformer 的手写文本识别 [J]. 成都信息工程大学学报, 2023, 38(06): 649-655.

[10] 刘华咏, 黄聪, 金汉均. 注意力增强的视觉 Transformer 图像检索算法 [J]. 电子测量技术, 2023, 46(23): 50-55.

基于日内 GARCH 模型 LAD 估计的混成检验

朱华锋¹, 陈燕珊², 张兴发²

1. 韶关学院 数学与统计学院, 广东 韶关 512005

2. 广州大学 岭南统计科学研究院, 广东 广州 510006

DOI:10.61369/ASDS.2025070017

摘要 : 基于 GARCH 模型的 LAD 估计和混成检验理论, 本文将 LAD 估计引入日内高频数据 GARCH 模型, 得到了对应的混成检验统计量以及估计的渐近正态性。数值模拟和实证研究的结果显示, 相比低频数据模型, 基于高频数据模型的估计效果更好, 对应的波动率代表模型能够更好的捕捉波动率信息, 参数估计的精确度也更高。

关键词 : 最小绝对值偏差估计 (LADE); GARCH 模型; 混成检验

Portmanteau Test Based on LAD Estimation of Intraday GARCH Model

Zhu Huafeng¹, Chen Yanshan², Zhang Xingfa²

1.School of Mathematics and Statistics, Shaoguan University, Shaoguan, Guangdong 512005

2.Lingnan Research Institute of Statistical Science, Guangzhou University, Guangzhou, Guangdong 510006

Abstract : Based on the LAD estimation of the GARCH model and the theory of portmanteau test, this paper introduces the LAD estimation into the GARCH model of intraday high-frequency data, obtains the corresponding portmanteau test statistics and the asymptotic normality of the estimation. Compared with the low-frequency data model, the results of numerical simulation and empirical research show that the estimation effect based on the high-frequency data model is better, the corresponding volatility representative model can better capture fluctuation information, and the accuracy of parameter estimation is also higher.

Keywords : Least Absolute Deviations Estimation(LAD); GARCH model; portmanteau test

引言

波动率是衡量金融市场风险的重要工具之一。关于金融市场波动率的模型, 自从1982年 Engle^[1] 提出 ARCH 模型, Bollerslev^[2]1986 年将其推广为 GARCH 模型后, 研究者们在此基础上提出了一系列统称为 GARCH 族模型的条件异方差模型。迄今为止, GARCH 族模型的理论及其应用仍然是金融计量学和时间序列分析领域的研究热点。

方便起见, 我们将 GARCH (1,1) 模型表示如下:

$$y_t = \sigma_t \varepsilon_t, \quad (1)$$

$$\sigma_t^2 = \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2, \quad (2)$$

其中, $\omega > 0, \alpha \geq 0, \beta \geq 0$; $\varepsilon_t \sim i.i.d.(0, \sigma^2)$; 且 ε_t 与 $y_s (s \leq t)$ 独立; y_t 与 σ_t^2 分别表示某资产第 t 天的对数收益率与波动率。

随着计算机科学技术的发展, 人们可以容易地获取金融资产的日内交易数据并对其进行统计与分析, 学者们也开始基于高频数据对 GARCH 族模型进行研究。Visser^[3] 在 GARCH (1,1) 模型基础上, 融入日内高频数据信息, 提出如下模型:

$$Y_t(u) = \sigma_t Z_t(u), \quad (3)$$

作者简介:

朱华锋, 韶关学院数学与统计学院, 博士研究生, 研究方向为时间序列、金融统计;

陈燕珊, 广州大学岭南统计科学研究院, 研究方向为时间序列;

张兴发, 副教授, 博士研究生, 广州大学岭南统计科学研究院, 研究方向为时间序列、金融统计、环境统计、机器学习。

$$\sigma_t^2 = \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2, \quad (4)$$

其中, $Y_t(u)$ 与 σ_t^2 为某资产第 t 天的日内对数收益率与波动率; $Z_t(\cdot)$ 为相互独立且分布相同的随机过程。文章还进一步提出了如下高频数据 GARCH 波动率代表模型:

$$H_t \triangleq H(Y_t(u)) = \sigma_t H(Z_t(u)), \quad (5)$$

$$\sigma_t^2 = \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2. \quad (6)$$

GARCH 族模型的估计方法主要有拟极大似然估计 (QMLE)、拟极大指数似然估计 (QMELE) 和最小绝对值偏差估计 (LAD)。关于 QMLE 及其渐近性质, Lee 和 Hansen^[4] 对其进行了比较充分的研究。但是, Hall 和 Yao^[5] 发现 QMLE 并不总是满足渐近正态性, 且其收敛速度有时候会小于 \sqrt{n} 。为解决以上问题, Peng 和 Yao^[6] 提出了三种 LAD 估计, 文章中还未加证明的介绍了 QMELE。随后, Li 和 Li^[7] 对 QMELE 的渐近性质进行了证明。此后, 研究者们通常采用 QMLE、QMELE 和 LAD 三种方法对 GARCH 族模型的参数进行估计。

随着日内高频交易数据越来越容易获得, 学者们开始对基于高频数据的波动率模型理论及其应用进行研究。李莉丽等^[8] 和李莉丽等^[9] 分别使用 QMLE 和 QMELE 对高频数据 GARCH (1,1) 模型进行了估计, 研究显示引入高频数据能够提升估计的精度。陈燕珊等^[10] 基于高频数据 GARCH 模型 QMELE 的性质, 提出了一种 Portmanteau Q 检验统计量。

关于混成检验的研究, Box 和 Pierce^[11] 率先提出了混成检验 (又称 Q 检验) 方法。之后, Ljung 和 Box^[12] 提出了更接近卡方分布的 LB 检验统计量。McLeod 和 Li^[13] 将 LB 统计量用于时间序列模型的检验, 对 ARMA 模型的拟合优度进行了检验。之后, Li 和 Mak^[14] 对 ARCH 模型和 GARCH 模型进行了混成检验。Chen 和 Zhu^[15] 针对 GARCH 族模型构建了一种基于秩的 S(M) 混成检验统计量。迄今为止, 混成检验仍然是 GARCH 族模型检验的常用方法之一。

本文的内容安排如下: 第一小节简要介绍高频 GARCH 模型的 LAD 估计及其渐近正态性; 第二小节分别基于低频和高频数据对 GARCH 模型的 LAD 估计进行混成检验; 第三小节和第四小节为数值模拟和实证分析; 最后对文章进行简要总结。

一、基于高频数据 GARCH 模型的 LAD 估计

本文将使用如下高频波动率代表 GARCH (p, q) 模型:

$$H_t = \sigma_t z_{H,t}, \quad (7)$$

$$\sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i y_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2, \quad (8)$$

其中 $z_{H,t} = H(Z_t(u))$ 。

首先对模型做一些变形, 假设存在一个常数 λ_H 使得 $H(Z_t(u))/\lambda_H$ 的平方满足中位数为 1。记 $\varepsilon_t^* = H(Z_t(u))/\lambda_H$, 易得 $\lambda_H = \sqrt{\text{median}(H(Z_t(u)^2))}$ 。

为方便, 令 $H_t \triangleq H(Y_t(u))$ 、 $\sigma_t^* = \sigma_t \lambda_H$, 则模型 (5) — (6) 可化为:

$$H_t = \sigma_t^* \varepsilon_t^*, \quad (9)$$

$$\sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i y_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2. \quad (10)$$

记 $\theta = (\omega, \alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q)$ 和 $\theta^* = (\omega^*, \alpha_1^*, \dots, \alpha_p^*, \beta_1^*, \dots, \beta_q^*)$, 其中 $\omega^* = \omega \nu_H^2$, $\alpha_i^* = \alpha_i \nu_H^2$, $\beta_j^* = \beta_j$, $i = 1, 2, \dots, p$, $j = 1, 2, \dots, q$ 。

由模型 (7) — (8) 可得, LAD 估计如下所示:

$$\begin{aligned} \hat{\theta}^* &= \arg \min_{\theta^*} \sum_{t=1}^n |\log(H_t^2) - \log(\sigma_t^{*2})| \\ &= \arg \min_{\theta^*} \sum_{t=1}^n l_t^* = \arg \min_{\theta^*} L_n^*. \end{aligned} \quad (11)$$

参考 Peng 和 Yao^[6] 的研究, 为得到参数 θ^* 的渐近性质, 需给出如下假设条件:

假设 1 记 θ^* 的真值为 $\theta_0^* = (\omega_0^*, \alpha_{10}^*, \dots, \alpha_{p0}^*, \beta_{10}^*, \dots, \beta_{q0}^*)$, 且满足:

$$E \left[\log \left(\sum_{i=1}^p \alpha_{i0}^* \varepsilon_t^{*2} + \sum_{j=1}^q \beta_{j0}^* \right) \right] < 0;$$

假设 2 记 Θ 为包含 θ_0^* 的紧集, 且对任意 $\theta^* \in \Theta$ 和某些 $\varpi > 0$, $\sigma_t^{*2}(\theta^*) \geq \varpi, a.s.$;

假设 3 对所有 $\pi \in \mathbb{R}^m$, 如果 $\pi'(\partial \sigma_t^{*2} / \partial \theta^*)_{i=1,2,\dots,m} = 0, a.s.$, 那么 $\pi = 0$;

假设 4 存在 $t_0 \in (0, 1)$, 使得 $E|y_{t_0}|^{2t_0} < \infty$, $E(H_{t_0}^{2t_0}) < \infty$, 且

$$E \left[\sup_{\theta^* \in \Theta} |\log \sigma_t^{*2}(\theta^*)| \right] < \infty, \quad E \left[\sup_{\theta^* \in \Theta} \left\| \frac{1}{\sigma_t^{*2}(\theta^*)} \frac{\partial \sigma_t^{*2}(\theta^*)}{\partial \theta^*} \right\| \right] < \infty,$$

$$E \left[\sup_{\theta^* \in \Theta} \left\| \frac{1}{\sigma_t^{*2}(\theta^*)} \frac{\partial^2 \sigma_t^{*2}(\theta^*)}{\partial \theta^* \partial \theta^*} \right\| \right] < \infty;$$

假设 5 对所有 $\theta^* \in \Theta$ 和某些 $\kappa > 0$, $\sigma_t^{*2}(\theta^*) \geq \kappa, a.s.$ 及某些常数 $\gamma \in (0, 1)$, 有

$$\begin{aligned} \sup_{\theta^* \in \Theta} \left| \widehat{\sigma_t^{*2}}(\theta) - \sigma_t^{*2}(\theta) \right| &\leq O(\gamma^l) R_l, \\ \sup_{\theta^* \in \Theta} \left\| \frac{1}{\widehat{\sigma_t^{*2}}(\theta)} \frac{\partial \widehat{\sigma_t^{*2}}(\theta)}{\partial \theta^*} - \frac{1}{\sigma_t^{*2}(\theta)} \frac{\partial \sigma_t^{*2}(\theta)}{\partial \theta^*} \right\| &\leq O(\gamma^l) R_l, \end{aligned}$$

$$\sup_{\theta \in \Theta} \left\| \frac{1}{\hat{\sigma}_t^2(\theta^*)} \frac{\partial^2 \hat{\sigma}_t^2(\theta)}{\partial \theta^* \partial \theta^*} - \frac{1}{\sigma_t^2(\theta^*)} \frac{\partial^2 \sigma_t^2(\theta)}{\partial \theta^* \partial \theta^*} \right\| \leq O(\gamma') R_t,$$

其中 R_t 是大于0的随机变量, 且对某些 $t \in (0,1)$, 有 $ER_t^4 = O(t)$ 。

假设6 中位数 $\text{median}(\log \varepsilon_t^2) = 0$, $\log \varepsilon_t^2$ 的密度函数 $f(x)$ 在 $x=0$ 处连续, 且满足: $f(x) > 0$, $\sup_{x \in R} f(x) < \infty$ 。

在以上假设条件下, 我们得到如下定理1。

定理1 在假设1—假设6条件下, $\hat{\theta}^*$ 的 LAD 估计具有渐近正态性, 即:

$$\sqrt{n}(\hat{\theta}^* - \theta_0^*) \xrightarrow{d} N(0, [2f^{*-2}(0)]^{-2} G^{*-1}), n \rightarrow \infty, \quad (12)$$

其中

$$G^* = E\left(\frac{1}{\sigma_t^{*4}} \frac{\partial \sigma_t^{*2}}{\partial \theta^*} \frac{\partial \sigma_t^{*2}}{\partial \theta^*}\right). \quad (13)$$

且

$$\sqrt{n}(\hat{\theta}^* - \theta_0^*) = \frac{G^{*-1}}{2f^*(0)\sqrt{n}} \sum_{t=1}^n \frac{\xi_t^*}{\partial \sigma_t^{*2}} \frac{\partial \sigma_t^{*2}}{\partial \theta^*} + o_p(1) \quad (14)$$

其中 $\xi_t^* = \text{sgn}(\varepsilon_t^{*2} - 1)$ 。

为了得到参数估计 $\hat{\theta}$, 先要得到 λ_H^2 的估计, 如下式所示:

$$\hat{\lambda}_H^2 = \frac{1}{n} \sum_{t=1}^n \frac{\hat{\sigma}_t^{*2}(\hat{\theta}^*)}{\hat{\sigma}_t^2(\hat{\theta})}, \quad (15)$$

其中, $\hat{\sigma}_t^{*2}$ 和 $\hat{\sigma}_t^{*2}$ 是分别基于模型 (1) — (2) 和模型 (7) — (8) 使用 LAD 方法得到的波动率估计。

于是, 由 (9) 式和 (13) 式, 可得参数估计 $\hat{\theta}$ 为:

$$\hat{\omega} = \frac{\hat{\omega}^*}{\hat{\lambda}_H^2}, \hat{\alpha}_i = \frac{\hat{\alpha}_i^*}{\hat{\lambda}_H^2}, \hat{\beta}_j = \hat{\beta}_j^*. \quad (16)$$

在定理1的基础上, 我们还可以得到服从 χ^2 分布的检验统计量, 如引理1所示。

引理1 当定理1的假设条件成立时, 可得如下服从卡方分布的检验统计量:

$$S(M) = n\hat{\rho}' \hat{\Lambda}^{-1} \rho \xrightarrow{d} \chi^2(m), n \rightarrow \infty,$$

其中 $\Lambda = I_m + \hat{p}_j \hat{H} \hat{G}^{-1} T' H' - 2\hat{p}_j \hat{\rho}_0' \hat{H} \hat{G}^{-1} H^{*'}'$, $\hat{p}_j = \hat{f}(\hat{0}) / \hat{f}'(\hat{0})$, $\hat{\rho}_0^* = \frac{1}{n} \sum_{t=1}^n \hat{\xi}_t \hat{\xi}_t'$,

$$\hat{H}^* = (\hat{H}_1^*, \hat{H}_2^*, \dots, \hat{H}_m^*), \quad \hat{H}_k = \frac{1}{n} \sum_{t=k+1}^n \frac{\hat{\xi}_{t-k}}{\sigma_t^{*2}} \frac{\partial \sigma_t^{*2}}{\partial \theta^*}, k=1, 2, \dots, m.$$

二、混成检验

对于 GARCH 模型, 由 (1) 式可得, 其误差为 $\varepsilon_t = y_t / \sigma_t$, 记基于秩的滞后 k 阶自相关函数为:

$$\hat{\rho}_k^* = \frac{\sum_{t=k+1}^n (\xi_t - E\xi_t)(\xi_{t-k} - E\xi_t)}{\sum_{t=k+1}^n (\xi_t - E\xi_t)^2}, \quad k=1, 2, \dots, m,$$

其中: $\xi_t = \text{sgn}(\varepsilon_t^2 - 1)$, 符号函数 $\text{sgn}(x) = I(x > 0) - I(x < 0)$ 。

为了方便以下的混成检验研究, 设定原假设与备择假设如下:

$$H_0: \rho_1^* = \rho_2^* = \dots = \rho_m^* = 0; \quad H_1: \text{存在某个 } \rho_i^* \neq 0, \quad i \in \{1, 2, \dots, m\}.$$

下面分别基于低频数据和高频数据对 GARCH 模型的 LAD 估计进行混成检验。

(一) 基于低频数据的混成检验

对于低频模型 (1) — (2), $\text{median}(\varepsilon_t^2) = 1$ 当时, $\{\xi_t\} \sim i.i.d.(0,1)$ 。

参考 Chen 和 Zhu^[15] 的研究, 记随机变量 $\{\xi_t\}$ 的滞后 k 阶样本残差自相关函数为:

$$\hat{\rho}_k^* = \frac{\sum_{t=k+1}^n (\hat{\xi}_t - \bar{\xi}_n)(\hat{\xi}_{t-k} - \bar{\xi}_n)}{\sum_{t=k+1}^n (\hat{\xi}_t - \bar{\xi}_n)^2}, \quad k=1, 2, \dots, m. \quad (17)$$

其中 $\hat{\xi}_t = \text{sgn}(y_t^2 / \hat{\sigma}_t^2(\hat{\theta}) - 1)$, $\bar{\xi}_n = n^{-1} \sum_{t=1}^n \hat{\xi}_t$ 。

由于 $\bar{\xi}_n = E(\xi_t) + o_p(1)$ 且 $\frac{1}{n} \sum_{t=1}^n (\hat{\xi}_t - \bar{\xi}_n)^2 = \text{var}(\hat{\xi}_t) + o_p(1) = 1 + o_p(1)$, 所以只需要考虑 $\hat{\rho}_k^* = \sum_{t=k+1}^n \hat{\xi}_t \hat{\xi}_{t-k}$ 即可。记 $\hat{\rho} = (\hat{\rho}_1, \hat{\rho}_2, \dots, \hat{\rho}_m)'$, 可得如下定理2。

定理2 在一定的正则条件下, $\sqrt{n}\hat{\rho}$ 满足渐近正态性, 即:

$$\sqrt{n}\hat{\rho} \xrightarrow{d} N(0, I_m - H\Sigma^{-1}H'), n \rightarrow \infty,$$

其中 $H = (H_1, H_2, \dots, H_m)'$, 且 $H_k = E\left(\frac{\xi_{t-1}}{\sigma_t^2} \frac{\partial \sigma_t^2}{\partial \theta}\right)$, $k=1, 2, \dots, m$,

$$\Sigma = E\left(\frac{1}{\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \theta} \frac{\partial \sigma_t^2}{\partial \theta'}\right).$$

证明: 见 Chen 和 Zhu^[15] 中定理2.1 的证明。

(二) 基于高频数据的混成检验

类似低频数据情形, 对于高频数据, 根据 Chen 和 Zhu^[15] 的研究, 可得如下样本残差的自相关函数:

$$\hat{\rho}_l^* = \frac{\sum_{t=l+1}^n (\hat{\xi}_t - \bar{\xi}_n)(\hat{\xi}_{t-l} - \bar{\xi}_n)}{\sum_{t=l+1}^n (\hat{\xi}_t - \bar{\xi}_n)^2},$$

其中 $\hat{\xi}_t = \text{sgn}(y_t^2 / \hat{\sigma}_t^2(\hat{\theta}) - 1)$, $\bar{\xi}_n = n^{-1} \sum_{t=1}^n \hat{\xi}_t$ 。

同理, 因为 $\bar{\xi}_n = E(\xi_t) + o_p(1)$, $n^{-1} \sum_{t=1}^n (\hat{\xi}_t - \bar{\xi}_n)^2 = \text{var}(\hat{\xi}_t) + o_p(1) = 1 + o_p(1)$, 所以只需要考虑 $\hat{\rho}_l^* = \sum_{t=l+1}^n \hat{\xi}_t \hat{\xi}_{t-l}$ 即可。记 $\hat{\rho} = (\rho_1, \rho_2, \dots, \rho_m)'$, 增加如下假设条件, 可得定理3。

假设7 对某些 $t \in (0,1)$, 有 $E|\varepsilon_t|^{2t} < \infty$ 。

定理3 在假设1—假设7条件下, 如果估计 $\hat{\rho}$ 的渐近方差是正定的, 那么:

$$\sqrt{n}\hat{\rho} \xrightarrow{d} N(0, \Lambda),$$

其中 $\Lambda = I_m + p_j \hat{H} \hat{G}^{-1} T' H' - 2p_j \hat{\rho}_0^* \hat{H} \hat{G}^{-1} H^{*'}'$,

$$p_j = \frac{f(0)}{f^*(0)}, \quad \rho_0^* = E(\xi_t \xi_t'),$$

$$H^* = (H_1^*, H_2^*, \dots, H_m^*), \quad H_k^* = E\left[\frac{\xi_{t-k}}{\sigma_t^{*2}} \frac{\partial \sigma_t^{*2}}{\partial \theta^*}\right], k=1, 2, \dots, m.$$

证明: 由于篇幅限制, 本文不给出具体的证明过程, 感兴趣

的读者可联系作者。

三、数值模拟

本小节通过数值模拟来验证本文所提理论的正确性。根据模型设定，需要对参数 θ 进行赋值，以生成 $Y_t(u)$ ，但由（3）—（4）式可知，生成 $Y_t(u)$ 之前，需先生成 $Z_t(u)$ 。参考 Visser^[3] 的方法，可生成符合条件的随机过程 $Z_t(u)$ ，具体公式如下：

$$d\Gamma_t(u) = -\delta(\Gamma_t(u) - \mu_\Gamma)du + \sigma_\Gamma dB_t^{(2)}(u),$$

$$d\Psi_t(u) = \exp(\Gamma_t(u))dB_t^{(1)}(u), \quad u \in [0,1],$$

$$Z_t(u) = \frac{\Psi_t(u)}{\sqrt{\text{median}[\Psi_t(1)]^2}}, \quad (18)$$

其中 $B_t^{(1)}(u)$ 和 $B_t^{(2)}(u)$ 由两个不相关的布朗运动生成。令 $Z_t(0) = 0$ ， $\Gamma_t(0)$ 由正态分布 $N(\mu_\Gamma, \sigma_\Gamma^2)$ 产生，设

$$\delta = \frac{1}{2}, \mu_\Gamma = \frac{1}{4}, \sigma_\Gamma = -\frac{1}{16}.$$

对于参数 u ，根据我国股票市场每天交易4个小时共240分钟的实况，只要令 $du = 1/240$ ，每分钟的交易情况即可模拟出来，进而可生成随机序列 $Z_t(u)$ 。此外，对于 GARCH (1,1) 模型，这里分别假设参数取值为如下两种情况来生成序列 $Y_t(u)$ ：
 $\theta_0 = (0.03, 0.06, 0.8)'$ 和 $\theta_0 = (0.02, 0.07, 0.9)'$ 。

进一步，要使用高频模型（7）—（8），先要通过选择适当的波动率代表对序列 $Y_t(u)$ 进行降维。参考陈燕珊^[10]，本文已取得波动率（简记为 RV）为波动率代表进行降维，采集样本频率选择每5分钟、每15分钟和每30分钟三种情况，记为 $H_t = RV5$ 、 $H_t = RV15$ 和 $H_t = RV30$ 。某资产每天开盘1分钟—240分钟的时间点记为 u_1, u_2, \dots, u_{240} ，并给定初值为 $Y_t(u_0) = Y_t(0) = 0$ 。

另外，为与低频数据模型的结果进行比较，低频情形下取 $H_t = |y_t|$ 。假设显著性水平为0.05，取自相关函数的滞后阶数为 $m = 6$ ，则对应 χ^2 检验的自由度也为6。通过模拟可得参数估计的结果，并计算对应的均根方误差（RMSE）。模拟具体结果见表1和表2^[16]。

表1 基于各波动率代表的参数估计 RMSE， $\theta_0 = (0.03, 0.06, 0.8)'$

		$ y_t $	RV30	RV15	RV5
n=300	ω	0.0776	0.0142	0.0086	0.0075
	α	0.0440	0.0152	0.0130	0.0108
	β	0.2472	0.0574	0.0400	0.0332
n=400	ω	0.0627	0.0118	0.0075	0.0065
	α	0.0375	0.0131	0.0113	0.0093
	β	0.2034	0.0488	0.0345	0.0293
n=500	ω	0.0564	0.0086	0.0068	0.0057
	α	0.0351	0.0117	0.0097	0.0080
	β	0.1873	0.0372	0.0302	0.0245

表2 基于各波动率代表的参数估计 RMSE， $\theta_0 = (0.02, 0.07, 0.9)'$

		$ y_t $	RV30	RV15	RV5
n=300	ω	0.0286	0.0075	0.0059	0.0051
	α	0.0349	0.0129	0.0108	0.0096
	β	0.0619	0.0227	0.0187	0.0162
n=400	ω	0.0268	0.0074	0.0057	0.0049
	α	0.0292	0.0108	0.0091	0.0080
	β	0.0519	0.0188	0.0157	0.0137
n=500	ω	0.0256	0.0077	0.0059	0.0048
	α	0.0241	0.0095	0.0081	0.0068
	β	0.0429	0.0166	0.0139	0.0115

表1和表2的结果显示，与基于低频数据的模型估计相比，高频数据模型的估计效果是更好的。根据参数估计的结果，进一步可求出对应检验统计量的值。参考陈燕珊^[10]，本文将重复1000次进行模拟，计算检验统计量的值大于 H_0 下 χ^2 检验统计量95%分位数的百分比，即为对应检验的经验水平。模拟所得结果如表3所示^[16]。

表3 各波动率代表模型基于 LAD 估计检验的经验水平

		n=300	n=400	n=500
$\theta_0 = (0.03, 0.06, 0.8)'$	$ y_t $	0.0420	0.0430	0.0450
	RV30	0.0410	0.0490	0.0470
	RV15	0.0480	0.0510	0.0520
	RV5	0.0430	0.0450	0.0500
$\theta_0 = (0.02, 0.07, 0.9)'$	$ y_t $	0.0360	0.0410	0.0450
	RV30	0.0390	0.0460	0.0510
	RV15	0.0470	0.0570	0.0520
	RV5	0.0430	0.0500	0.0500

根据统计建模理论，如果模型的拟合是充分的，那么残差序列应该为白噪声，统计量也就服从对应的 χ^2 分布；反之，如果残差是自相关的，统计量也就不服从对应的 χ^2 分布。基于此，我们设定模拟产生的随机误差项是自相关的，然后看检验统计量能否识别出拟合不充分的模型。参考陈燕珊^[10]，通过模拟产生自相关的误差序列，重复试验1000次，求出对应检验统计量的值大于 χ^2 分布95%分位数的百分比，即为对应的检验功效，模拟结果见表4^[16]。

表4 各波动率代表模型基于自相关的检验功效

	$\theta_0 = (0.03, 0.06, 0.8, \beta_2)'$				$\theta_0 = (0.02, 0.07, 0.9, \beta_2)'$		
	β_2	0.1	0.3	0.5	0.1	0.3	0.5
n=300	$ y_t $	0.036	0.099	0.517	0.041	0.072	0.201
	RV30	0.037	0.091	0.469	0.033	0.062	0.342
	RV15	0.040	0.062	0.301	0.044	0.092	0.491
	RV5	0.044	0.042	0.094	0.041	0.105	0.520
n=400	$ y_t $	0.037	0.043	0.124	0.039	0.089	0.292
	RV30	0.036	0.074	0.462	0.030	0.080	0.509
	RV15	0.043	0.112	0.639	0.034	0.117	0.669
	RV5	0.047	0.124	0.675	0.042	0.127	0.716
n=500	$ y_t $	0.056	0.067	0.168	0.053	0.101	0.378
	RV30	0.049	0.110	0.579	0.066	0.108	0.634
	RV15	0.061	0.148	0.744	0.063	0.151	0.797
	RV5	0.053	0.159	0.798	0.058	0.167	0.836

表4结果显示,相比低频情形,高频数据模的统计量的检验功效是逐渐变大的,并越来越接近于1。虽然自相关程度较低时也存在一定程度的误判情况,但总体来说高频数据模型对残差自相关的识别能力更强一些,发散效果也相对更好。

四、实证分析

本小节进行实证分析。参考陈燕珊^[16],选取如下两个数据集每天1分钟间隔的价格数据:2014年9月2日—2015年12月31日沪深300(CSI 300)共325个交易日和2013年9月18日—2015年10月13日上证50(SSE 50)共485个交易日。由于我国股票市场每天交易4个小时,如果记第t天每1分钟的价格为 $P_t(u)$,那么每天就有240个观测值。令 $P_t(l)$ 表示对应指数第t天最后1分钟的价格(即收盘价),对CSI 300和SSE 50的每日收盘价画时序图,如图1所示^[16]。

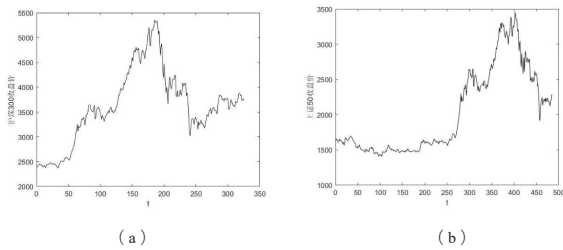


图1 CSI 300和SSE 50日收盘价时序图

图1(a)和(b)显示,CSI 300和SSE 50的日收盘价序列都是不平稳的。由时间序列平稳性相关理论可知,可以通过对价格数据取对数可达到序列平稳化的效果。

对于CSI 300和SSE 50的日内高频价格数据序列,也通过取对数得到对应的日内高频对数收益率。本文选取两大指数每日收盘价 $P_t(l)$,并对数据取对数得到CSI 300和SSE 50的每日收盘价对数收益率,然后画出对应时序图,如图2所示^[16]。

图2(a)和(b)显示,相比图1(a)和(b),虽然CSI 300和SSE 50两大指数的每日收盘价对数收益率仍有异方差性和波动率集聚等特性,但是波动性较原始数据大大降低,序列呈现出平稳的特征,因此可以考虑使用GARCH族模型对数据进行拟合。下面我们选择GARCH(1,1)模型对CSI 300和SSE 50两大指数的数据集进行建模与分析,并对低频和高频两种情形下的模型结果进行比较。

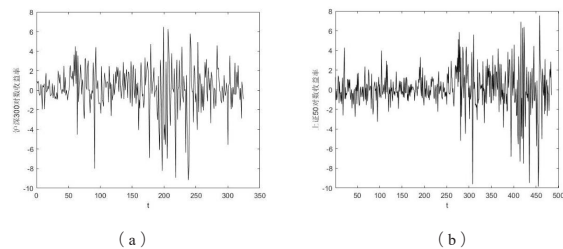


图2 CSI 300和SSE 50日间对数收益率时序图

参考陈燕珊^[16],实证研究同样选择已实现波动率(简记为RV)为波动率代表对数据进行降维,采集样本频率选择每5分

钟、每15分钟和每30分钟三种情况,记为 $H_t = RV5$ 、 $H_t = RV15$ 和 $H_t = RV30$ 。对于低频模型,令 $H_t = |y_t|$,以便与高频模型进行比较。计算可得参数的LAD估计结果如表5所示^[16]。

表5 CSI 300和SSE 50各波动率代表模型LAD估计结果

		$\hat{\omega}$	$\hat{\alpha}$	$\hat{\beta}$
CSI 300	$ y_t $	0.0408	0.0421	0.8476
	RV30	0.0472	0.0791	0.7493
	RV15	0.0654	0.0842	0.7193
	RV5	0.0494	0.0766	0.7536
SSE 50	$ y_t $	0.0274	0.0368	0.8569
	RV30	0.0121	0.0757	0.7799
	RV15	0.0264	0.0764	0.7567
	RV5	0.0109	0.0704	0.7951

下面以CSI 300数据集为例,分别在低频和高频两种情形下进行统计建模与分析。

首先,对于低频模型,取 $H_t = |y_t|$,可得如下GARCH(1,1)模型:

$$y_t = \sigma_t \varepsilon_t, \sigma_t^2 = 0.0494 + 0.0766 y_{t-1}^2 + 0.7536 \sigma_{t-1}^2.$$

其次,对于高频情形,以RV5为例,即取 $H_t = RV5$,可得如下GARCH(1,1)模型:

$$y_t = \sigma_t \varepsilon_t, \sigma_t^2 = 0.0408 + 0.0421 y_{t-1}^2 + 0.8476 \sigma_{t-1}^2.$$

同理可拟合出SSE 50指数在低频和高频两种情况下的GARCH(1,1)模型。利用所求模型即可求得CSI 300和SSE 50两大指数波动率的估计值,进而可画出对应波动率估计值的图形。为了便于与低频模型结果比较,日内高频情形下选择波动率代表为RV5的估计结果进行可视化操作,具体如图3所示^[16]。

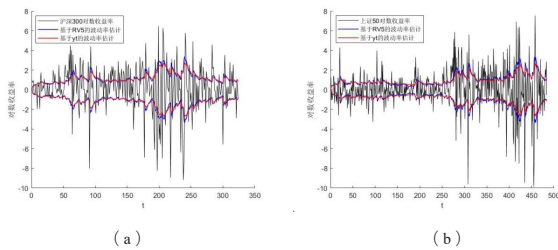


图3 CSI 300和SSE 50日内和日间波动率估计图

由图3(a)和(b)可以看出,相比低频数据情形,以波动率代表为RV5的日内高频GARCH模型能够更加充分的提取指数的波动率信息,所以对应的估计结果的精确性也越高。

跟模拟时一样,选取残差自相关函数的最大滞后阶数为 $m=6$,将所求数据带入检验统计量公式,可求得CSI 300和SSE 50对应统计量的值,具体数值如表6所示^[16]。

由表6结果可知,低频和高频数据两种情形下检验统计量的值都小于 χ^2 分布的临界值 $\chi_{0.95}^2(6)=12.5916$,故可认为两种模型对数据的拟合都是充分的。另外,由图3还能够发现,当原假设 H_0 成立时,高频数据GARCH模型的估计效果比低频数据模型更好一些。

表6 CSI 300和SSE 50各波动率代表的检验统计量结果

	$\tilde{Q}_e(y_t)$	$\tilde{Q}_e(RV30)$	$\tilde{Q}_e(RV15)$	$\tilde{Q}_e(RV5)$
CSI 300	7.8585	5.8365	6.7697	5.8390
SSE 50	2.4374	4.1757	4.7806	3.4218

进一步, 根据所得的日间低频模型和日内高频波动率代表模型, 我们还可以得到 CSI 300 和 SSE 50 两大指数的残差自相关图, 如图 4 所示^[16]。

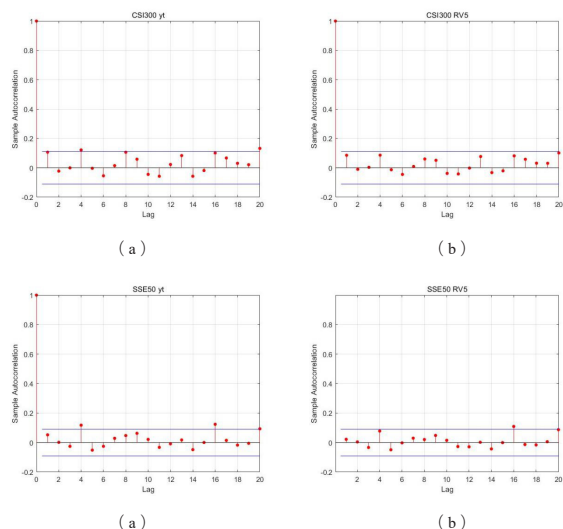


图4 CSI 300 和 SSE 50 的残差自相关图

图 4 结果显示, 与低频模型的残差自相关系数相比, 由以波动率代表为 RV5 的日内高频 GARCH 模型所得的残差自相关系数更小一些, 表明高频模型对波动率信息的提取更加充分, 这也与之前检验估计的效果图所得的结果是一致的。需要注意的是, 高频信息通常会包含一定程度的噪声, 因此选择信息的频率也并不是越高越好。通过对 CSI 300 和 SSE 50 两大指数数据的计算, 我们发现以 RV30 为代表的模型的检验统计量的值是小于以 RV15 为代表的模型。因为高频波动率代表模型的残差绝对值的估计可以通过 $H_t/\hat{\sigma}_t^*$ 求得, 所以对于不同的波动率代表模型, 想要看模型对高

频信息的提取效果, 也要对相应的残差自相关图进行对比。因为 $H_t \geq 0$, 所以在计算残差自相关系数时, 只要计算 $\text{sgn}(y_t)H_t/\hat{\sigma}_t^*$ 即可。下面利用 CSI 300 指数的数据, 分别在 RV5 和 RV15 两种高频波动率代表情形下, 计算 $\text{sgn}(y_t)H_t/\hat{\sigma}_t^*$, 可得对应的高频残差自相关图, 如图 5 所示^[16]。

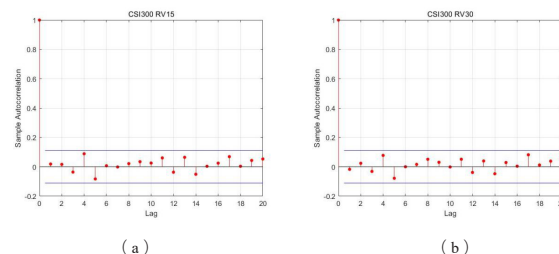


图5 CSI 300 指数的高频残差自相关图

图 5 结果显示, 与以 RV15 为代表的模型相比, 以 RV30 为代表的高频模型的自相关系数相对更小一些, 这同样与之前检验估计的效果图所得的结果是一致的。

五、结束语

本文基于 LAD 估计与混成检验的相关理论与研究, 将 LAD 估计引入日内高频数据的 GARCH 模型, 在一定的假设条件下得到了对应的混成检验统计量以及估计的渐近正态性质。数值模拟和实证研究的结果均显示, 与低频数据模型相比, 基于日内高频数据模型的估计效果相对更好一些, 拟合模型所得的残差自相关系数也更小一些, 说明高频波动率代表 GARCH 模型能够更加充分的提取波动率信息, 因此参数估计的精确性也更高一些。

参考文献

- [1] Engle R F. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*[J], 1982, 50(4): 987 - 1007.
- [2] Bollerslev T. Generalized autoregressive conditional heteroskedasticity[J]. *Journal of Econometrics*, 1986, 31(3): 307-327.
- [3] Visser M P. GARCH parameter estimation using high-frequency data[J]. *Journal of Financial Econometrics*, 2011, 9(1): 162-197.
- [4] Lee S, Hansen B. Asymptotic Theory for the GARCH (1,1) Quasi-Maximum Likelihood Estimator[J]. *Econometric Theory*, 1994, 10(1): 29-52.
- [5] Hall P, Yao Q. Inference in ARCH and GARCH models with heavy-tailed errors[J]. *Econometrica*, 2003, 71(1): 285-317.
- [6] Peng L, Yao Q. Least absolute deviations estimation for ARCH and GARCH models[J]. *Biometrika*, 2003, 90(4): 967-975.
- [7] Li G, Li W K. Least absolute deviation estimation for fractionally integrated autoregressive moving average time series models with conditional heteroscedasticity[K]. *Biometrika*, 2008, 95(2): 399-414.
- [8] 李莉丽, 张兴发, 李元, 等. 基于高频数据的日频 GARCH 模型估计 [J]. *广西师范大学学报 (自然科学版)*, 2021, 39(4): 68-78.
- [9] 李莉丽, 张兴发, 邓春亮, 等. 基于高频数据的 GARCH 模型拟极大指数似然估计 [J]. *应用数学学报*, 2022, 45(5): 652-664.
- [10] 陈燕珊, 张兴发, 田玥, 等. 基于高频数据的 GARCH 模型拟极大指数似然估计的一种 portmanteau Q 检验 [J]. *广州大学学报 (自然科学版)*, 2024, 23(5): 54-68.
- [11] Box G E P, Pierce D A. Distribution of Residual Autocorrelations in Autoregressive Integrated Moving Average Time Series Models[J]. *Journal of the American Statistical Association*, 65, 1509-1526.
- [12] Ljung G M, Box G E P. On a Measure of Lack of Fit in Time Series Models[J]. *Biometrika*, 1978, 65(332): 297-303.
- [13] McLeod A I, Li W K. Diagnostic Checking ARMA Time Series Models Using Squared-Residual Autocorrelations[J]. *Journal of Time Series Analysis*, 1983, 4(4): 269-273.
- [14] Li W K, Mak T K. On the squared residual autocorrelations in non-linear time series with conditional heteroskedasticity[J]. *Journal of Time series Analysis*, 1994, 15(6): 627-636.
- [15] Min C, Ke Z. Sign-based portmanteau test for ARCH-type models with heavy-tailed innovations, *Journal of Econometrics*, 2015, 189(2): 313-320.
- [16] 陈燕珊. 基于高频数据的 GARCH 模型的混成检验研究 [D]. 广州大学, 2024.

广东省放心消费“双承诺”活动满意度测评报告

——以江门市为例

王润良

五邑大学 经济管理学院, 广东 江门 529020

DOI:10.61369/ASDS.2025070018

摘 要 : 为检验广东省放心消费“双承诺”活动满意度,将美国顾客满意度指数模型与史密斯政策执行过程模型相结合,构建广东省放心消费“双承诺”活动满意度指数模型。以江门市为例,采集经营主体满意度样本353个、消费者满意度样本400个。运用样本数据对经营主体满意度指数模型和消费者满意度指数模型进行验证,获得正式测评模型及相关参数。运用问卷数据和模型参数进行满意度指数计算,得到江门全市总体满意度指数值为78.91,消费者满意度指数值为73.85。对满意度指数和问卷数据进行分析,发现放心消费“双承诺”活动存在以下问题:放心消费“双承诺”活动申请程序便利性不高;商家对“双承诺”活动的主动告知做得不到位;不同企业称号的经营主体满意度存在差异;不同身份消费者满意度存在差异;不同性别消费者满意度存在差异。针对这些问题,提出相应的对策建议。

关 键 词 : 放心消费“双承诺”;满意度指数;ACSI;史密斯政策执行过程模型;结构方程模型

Guangdong Province Assured Consumption "Double Commitment" Activity Satisfaction Evaluation Report — Take Jiangmen City as an Example

Wang Runliang

School of Economics and Management, Wuyi University, Jiangmen, Guangdong 529020

Abstract : To test the satisfaction level of the "Double Commitment" campaign for safe consumption in Guangdong Province, the American Customer Satisfaction Index model and the Smith policy implementation process model were combined to construct a satisfaction index model for the "Double Commitment" campaign for safe consumption in Guangdong Province. Taking Jiangmen City as an example, 353 samples of business entity satisfaction and 400 samples of consumer satisfaction were collected. The sample data were used to verify the business entity satisfaction index model and the consumer satisfaction index model, obtaining the formal evaluation model and related parameters. The satisfaction index was calculated using the questionnaire data and model parameters, and the overall satisfaction index value for the entire Jiangmen City was 78.91, with the consumer satisfaction index value being 73.85. By analyzing the satisfaction index and questionnaire data, it was found that the "Double Commitment" campaign for safe consumption has the following problems: the application process for the "Double Commitment" campaign is not very convenient; merchants do not actively inform about the "Double Commitment" campaign; there are differences in satisfaction among business entities with different titles; there are differences in satisfaction among consumers of different identities; and there are differences in satisfaction among consumers of different genders. In response to these problems, corresponding countermeasures and suggestions are proposed.

Keywords : assured consumption "double commitment"; satisfaction index; ACSI; Smith policy implementation process model; structural equation model

引言

2020年,广东省消委会在广东省市场监管局的指导支持下,在全省开展“放心消费承诺”“线下无理由退货承诺”(以下称“双承诺”)活动,4年多来,活动取得良好效果。根据2024年广东省民生实事和开展城乡居民放心消费行动的统一部署,省市场监管局指导省消委会在全省范围内继续开展放心消费“双承诺”活动。上半年,全省新增“放心消费承诺单位”24085家,新增“线下无理由退货

承诺店”3424家,目前,全省累计创建放心消费承诺单位9万家,线下无理由退货承诺店1.9万家,放心消费承诺单位共自行处理消费纠纷91.4万件,线下无理由退货承诺店共为消费者无理由退货商品83.3万件,退货金额达7533万元^[1]。根据广东省相关部署,江门市全域自2020年起开展放心消费“双承诺”活动。为检验活动效果,作者受江门市消费者权益保护委员会委托,开展了放心消费“双承诺”活动经营主体满意度和消费者满意度调查。

一、模型与方法

(一) 测评模型构建

美国顾客满意度指数模型(American Customer Satisfaction Index, ACSI)是顾客满意度测评的经典模型,它首先被用于私营部门顾客满意度测评,此后被借鉴用于公共部门公众满意度测评^{[2][3]}。ACSI由顾客期望、感知质量、感知价值、顾客满意度、顾客抱怨和顾客忠诚等6个潜变量以及9条因果路径构成^[4-6]。美国学者托马斯·史密斯在《政策执行过程》一书中提出了政策执行过程模型,认为政策执行过程受最优化政策、执行机构、目标群体和环境因素影响,同时四个影响素之间存在复杂的影响关系^[7-10]。本文认为,广东省放心消费“双承诺”活动的开展是典型的政策执行过程,因此将ACSI和史密斯政策执行过程模型相结合建立了广东省放心消费“双承诺”活动满意度测评模型(如图1)。其中,最优化政策指清晰合理、切实可行、符合事物发展客观规律的政策,如活动方案中有关参与活动经营主体经营活动和消费者消费行为的促进等内容;目标群体反映政策实施目标群体对政策的认可度与参与度,“双承诺”活动的目标群体指参与活动的经营主体和消费者两类;执行机构指政策执行机构的服务,本模型中经营主体满意度的执行机构是市场监管部门,消费者满意度的执行机构是参与“双承诺”活动的经营主体;环境因素指政策实施的物质、技术、人文等环境,如当地经济发展水平、消费者的行为习惯等。

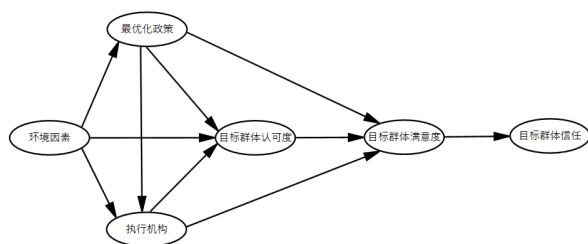


图1 广东省放心消费“双承诺”活动满意度测评模型

(二) 满意度指数计算方法

本文借鉴ACSI满意度指数计算方法进行满意度指数计算,各潜变量满意度指数公式如下^[11]:

$$A = \frac{\sum_{i=1}^k P_i \bar{X}_i - \sum_{i=1}^k P_i}{9 \sum_{i=1}^k P_i} \times 100 \quad (1)$$

其中, A 为各潜变量满意度指数, P_i 为观察变量的标准化路

径系数, \bar{X}_i 为第 i 个观察变量的平均值, k 为属于潜变量的观察变量个数, $A \in (0,100)$ 。

满意度指数 SI 计算公式如下:

$$SI = \frac{\sum_{j=1}^l A_j W_j}{\sum_{j=1}^l W_j} \quad (2)$$

其中, A_j 为第 j 个潜变量满意度得分, W_j 为第 j 个潜变量的总效应, l 为目标群体满意度前因潜变量个数,本文中 $l=4$ 。

二、数据

(一) 数据收集

根据测评模型设计《江门市“双承诺”活动经营主体满意度调查问卷》(以下简称“经营主体问卷”)和《江门市“双承诺”活动消费者满意度调查问卷》(以下简称“消费者问卷”),并在江门市下辖四市三区参与“双承诺”活动的经营主体经营场所开展问卷调查。其中,“经营主体问卷”共采集样本数据为353份,“消费者问卷”共采集样本数据400份。

(二) 数据预分析

1. 样本特征分析

针对经营主体样本数据进行频数统计,100%的调查对象参与了“双承诺”活动;调查对象在各辖区的分布比例如下:蓬江区占18.7%,江海区占10.2%,新会区占14.45%,台山市占17%,开平市占17%,鹤山市占11.33%,恩平市占11.33%,与各市(区)参与“双承诺”活动的经营主体数量分布相当;调查对象中,企业占比42.21%,个体工商户占比57.79%,个体工商户所占比略高于企业;调查对象经营主体经营类别主要分为商品类和服务类两大类,其中商品类经营主体占比为83%,服务类经营主体占比为17%;商品类经营主体以医疗及医药用品类(23.89%)、肉菜生鲜类(22.18%)和食品类(非餐饮服务,18.09%)为主,服务类以生活、社会服务类(50%)为主;调查对象中高新技术企业占比11.61%,专精特新中小企业占比20.11%。

针对消费者问卷数据进行频数统计,参与调查的400人中,江门市居民占80.5%,外地游客占19.5%;调查样本在蓬江区、江海区、新会区、台山市、开平市、鹤山市和恩平市的分布比例分别是20%、10%、20%、15%、15%、10%和10%,与各市(区)人口分布比例大致相当;98.5%的样本知晓江门市正在开展“双承诺”活动,100%的样本表示接受过参与“双承诺”活动商家的服务;样本中男性占44.5%,女性占55.5%,女性略多于男性;

调查样本以青年群体为主，其中18岁至25岁的年龄段占比高达50.5%，26至35岁年龄段占比13%；样本学历以大学本科为主，占比为49.5%；样本职业以学生群体和企业员工为主，分别占比46.25%和17%（注：调查期间正值暑假，因此学生群体较多）。

2. 信度与效度分析

经对经营主体问卷数据和消费者问卷数据进行信度分析，两个问卷数据中所有潜变量及模型整体的克隆巴赫 Alpha 系数均大于0.8，说明问卷数据可靠性较高。

经对经营主体问卷数据和消费者问卷数据进行效度分析，两个问卷数据 $KMO \geq 0.6$ ，Bartlett 球形度检验显著性概率均 < 0.001 。进一步进行因子分析，发现两个问卷数据的所有成分和模型整体的累积方差解释率均大于50%，说明问卷数据效度良好。

可见，经营主体问卷和消费者问卷数据信度、效度良好，可用于进一步分析。

（三）模型验证与参数计算

1. 经营主体模型验证与参数计算

根据图1所示测评模型建立经营主体结构方程测量模型，运用AMOS 软件运行该模型，进行模型优化，获得经营主体满意度测量模型输出结果和参数分别如图2、表1所示。其中，表1中的标准化路径系数是经营主体满意度指数计算的依据。

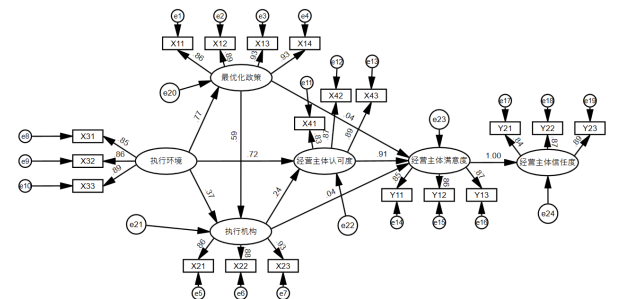


图2 经营主体满意度结构方程模型验证结果

表1 经营主体满意度结构方程模型标准化路径系数

路径			标准化路径系数
最优化政策	<---	执行环境	0.773
执行机构	<---	执行环境	0.37
执行机构	<---	最优化政策	0.592
经营主体认可度	<---	执行环境	0.718
经营主体认可度	<---	执行机构	0.243
经营主体满意度	<---	经营主体认可度	0.908
经营主体满意度	<---	最优化政策	0.042
经营主体满意度	<---	执行机构	0.041
经营主体信任度	<---	经营主体满意度	0.999
X11	<---	最优化政策	0.856
X12	<---	最优化政策	0.893
X13	<---	最优化政策	0.925
X14	<---	最优化政策	0.926
X22	<---	执行机构	0.879
X23	<---	执行机构	0.931
X41	<---	经营主体认可度	0.829
X31	<---	执行环境	0.852

路径			标准化路径系数
X32	<---	执行环境	0.855
X33	<---	执行环境	0.894
X21	<---	执行机构	0.863
X42	<---	经营主体认可度	0.869
X43	<---	经营主体认可度	0.889
Y11	<---	经营主体满意度	0.845
Y12	<---	经营主体满意度	0.863
Y13	<---	经营主体满意度	0.867
Y21	<---	经营主体信任度	0.844
Y22	<---	经营主体信任度	0.87
Y23	<---	经营主体信任度	0.889

根据表1计算各前因变量对经营主体满意度的影响效应如表2所示。该表参数亦为经营主体满意度指数计算的依据。

表2 经营主体满意度结构方程模型各潜变量对经营主体满意度的影响效应

潜变量	潜变量→经营主体满意度		
	直接效应	间接效应	总效应
最优化政策	0.04	0.15	0.19
执行机构	0.04	0.22	0.26
执行环境	—	0.91	0.91
经营主体认可度	0.91	—	0.91

2. 消费者模型验证与参数计算

同理，可建立消费者满意度结构方程测量模型，并获得消费者满意度指数模型输出结果和标准化路径系数，以及各潜变量对消费者满意度的影响效应（模型图和数据表略）。

三、结果

（一）经营主体满意度指数计算

运用 SPSS 软件计算经营主体模型观察变量平均值，然后将其和表1、表2数据代入公式（1）、（2）计算，可得经营主体满意度指数值如表3所示，可见江门市经营主体满意度指数为78.91分。

表3 江门市经营主体满意度指数计算表

潜变量	观察变量	编号	观察变量平均值	观察变量路径系数 P	潜变量满意度得分 A	潜变量总效应 W	满意度指数 SI
最优化政策（X1）	“双承诺”活动总体上积极有效	X11	4.31	0.856	78.69	0.19	78.91
	“双承诺”活动促进了营业额增长率的提高	X12	4.07	0.893			
	“双承诺”活动提高了顾客回头率	X13	4.1	0.925			
	“双承诺”活动促进了品牌推广	X14	4.12	0.926			

执行机构 (X2)	“双承诺”活动宣传及时到位	X21	4.17	0.863	78.33	0.26	
	“双承诺”活动申请程序便利	X22	4.09	0.879			
	“双承诺”活动推进手段有效	X23	4.14	0.931			
执行环境 (X3)	本地区产品和服务质量对“放心消费”的支撑度	X31	4.26	0.852	78.81	0.91	
	本地区居民消费心理对“放心消费”的支撑度	X32	4.09	0.855			
	本地区“放心消费”环境对外地游客的吸引力	X33	4.11	0.894			
经营主体认可度 (X4)	活动知晓度	X41	4.22	0.829	79.23	0.91	
	活动参与意愿	X42	4.13	0.869			
	活动认可度	X43	4.16	0.889			
经营主体满意度 (Y1)	总体满意度	Y11	4.29	0.845	80.74	—	—
	与预期相比满意度	Y12	4.18	0.863			
	与活动开展前相比满意度	Y13	4.22	0.867			
经营主体信任度 (Y2)	总体信任度	Y21	4.27	0.844	79.98	—	—
	继续参与意愿	Y22	4.16	0.87			
	向同行推荐度	Y23	4.17	0.889			

（二）消费者满意度指数计算

同理，计算江门市消费者满意度指数为73.85分（数据表略）。

（三）经营主体满意度指数分析

1. 经营主体满意度指数分析

江门市经营主体满意度指数78.91分，属良好水平，说明经营主体对于江门市推出放心消费“双承诺”活动持较高水平的欢迎态度。

就影响经营主体满意度指数的四个前因变量得分而言，最优化政策、执行机构、执行环境和经营主体认可度满意度指数分别为78.69、78.33、78.81和79.23分，得分最高的为经营主体认可度，得分最低的为执行机构，说明经营主体对“双承诺”活动持

较高的认可态度，而活动执行机构的满意度相对较低（图3）。

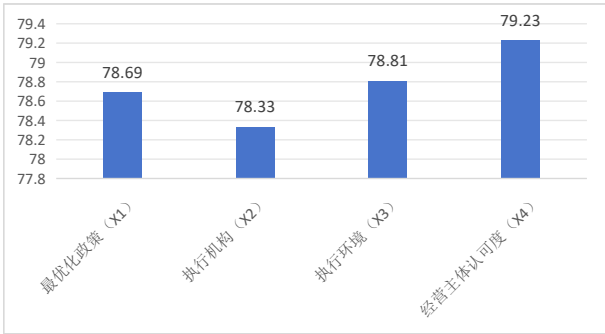


图3 江门市经营主体潜变量满意度指数比较分析

2. 不同类型经营主体满意度分析

经营主体类型与经营主体满意度的卡方检验结果如表4所示，可见四类经营主体中只有企业称号与经营主体满意度三个观察变量的渐进显著性（双侧）全部小于0.05，即不同企业称号的经营主体满意度存在差异；而其他三种经营主体类型的类别内不同细类经营主体满意度不存在显著差异，因此不做进一步分析。不同企业称号的经营主体满意度得分如

表5所示，可见三种企业称号类别中，“其他”类的经营主体满意度三个观察变量平均值均高于高新技术企业和专精特新中小企业，这是因为后二者实力较强、经营情况相对较好，“双承诺”活动给其带来的收益并不明显；而“其他”类多为无企业称号的经营主体，其实力较弱，平时的经营情况一般，放心消费“双承诺”活动给其带来的收益较为明显。

表4 不同经营主体类型与经营主体满意度卡方检验

经营主体类型	经营主体满意度	值	自由度	渐进显著性（双侧）
承诺类型	Y11	9.47	8	0.304
	Y12	6.626	8	0.577
	Y13	7.309	8	0.504
经营主体大类	Y11	4.844	4	0.304
	Y12	9.019	4	0.061
	Y13	9.749	4	0.045
经营主体经营类别	Y11	1.986	4	0.738
	Y12	0.613	4	0.962
	Y13	1.145	4	0.887
企业称号	Y11	23.867	8	0.002
	Y12	28.841	8	0.000
	Y13	20.045	8	0.010

表5 不同企业称号经营主体满意度比较分析

	V11满意度平均值	V12满意度平均值	V13满意度平均值
高新技术企业	4.07	3.95	4.02
专精特新中小企业	4.01	4.06	4.14
其他	4.41	4.25	4.27

（四）消费者满意度指数分析

1. 消费者满意度指数分析

江门市消费者满意度指数为73.85分，属中等水平，说明消费者对于放心消费“双承诺”活动有更高的期待。

就影响消费者满意度指数的四个前因变量得分而言，最优化

政策、执行机构、执行环境和消费者认可度满意度指数分别为75.32、72.66、73.19和75.13分，其中最优化政策得分最高，执行机构得分最低，说明消费者对放心消费“双承诺”活动持肯定态度，而对执行机构的满意度相对较低（图4）。

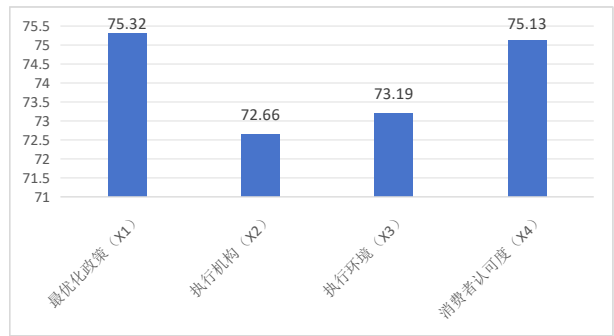


图4 江门市消费者潜变量满意度指数比较分析

2. 不同类别消费者满意度分析

本次调查中，消费者类别主要指身份类别（江门市居民、外地游客）和人口统计学类别（性别、年龄、学历和职业），运用卡方分析法进行显著性检验。就身份类别而言，样本中江门市居民322人，外地游客78人。针对这两种身份与消费者满意度三个观察变量进行卡方分析，结果如表6所示。由于V11、V12、V13的渐进显著性（双侧）均小于0.05，因此江门市居民与外地游客在消费者满意度三个观察变量上存在显著差异，其消费者满意度平均值如表7所示。可见，在消费者满意度三个维度上，江门市居民的满意度均高于外地游客。这主要是江门市居民的满意度是基于本地重复消费的体验，而外地游客是基于不同城市间消费的比较体验。

表6 不同身份消费者与消费者满意度卡方检验

	V11卡方检验			V12卡方检验			V13卡方检验		
	值	自由度	渐进显著性(双侧)	值	自由度	渐进显著性(双侧)	值	自由度	渐进显著性(双侧)
皮尔逊卡方	15.802 ^a	4	0.003	20.384 ^a	4	0	19.410 ^a	4	0.001
似然比	12.529	4	0.014	18.917	4	0.001	18.313	4	0.001
线性关联	6.834	1	0.009	3.732	1	0.053	4.109	1	0.043
有效个案数	400			400			400		
a2 个单元格 (20.0%) 的期望计数小于 5。最小期望计数为 1.76。									

表7 不同身份消费者满意度平均值比较

身份	V11 满意度平均值	V12 满意度平均值	V13 满意度平均值
江门市居民	4.07	4.03	4.04
外地游客	3.77	3.81	3.81

同理，进行性别、年龄、学历和职业等人口统计学特征与消费者满意度三个观察变量的卡方检验发现（数据表略），只有性别与消费者满意度三个观察变量的渐进显著性（双侧）全部小于0.05，即不同性别的消费者满意度存在显著差异。进一步分析发现（表8），男性消费者的消费者满意度总体上高于女性消费者，这可能是由于女性消费者的消费行为较男性更细致、挑剔所致。

表8 不同性别消费者满意度平均值比较

性别	V11 满意度平均值	V12 满意度平均值	V13 满意度平均值
男	4.03	4.03	3.99
女	3.99	3.95	4.00

（五）经营主体与消费者满意度指数比较分析

图5可见，无论是潜变量得分还是满意度指数（SI）得分，经营主体均高于消费者，说明经营主体对于“双承诺”活动的肯定程度高于消费者者。这可能是由于经营主体在该活动中获得的收益更显现，而消费者在该活动中获得的收益不明显，毕竟消费者在消费过程中遇到退货、赔付等情况极少。

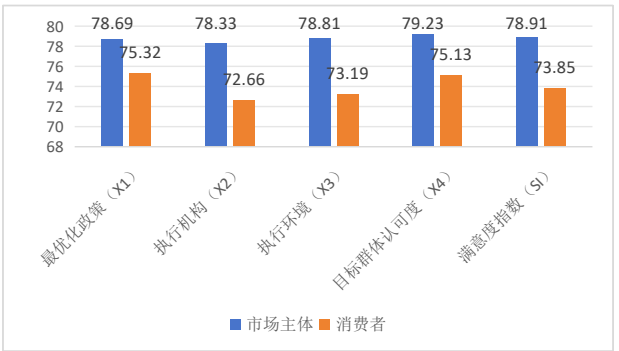


图5 经营主体与消费者满意度指数比较分析

四、结论与建议

（一）主要结论

上文的分析得到以下结论：

1. 江门全市经营主体满意度指数为78.91分，属良好水平；在四个前因变量中，执行机构得分最低。进一步分析发现，执行机构潜变量下“‘双承诺’活动应用程序便利”得分最低，因此主管部门需进一步简化“双承诺”活动应用程序，提高其便利性。
2. 江门全市消费者满意度指数为73.85分，属中等水平，明显低于经营主体满意度指数；在四个前因变量中，执行机构得分最低。进一步分析发现，执行机构潜变量下“商家能够主动告知‘双承诺’活动内容”得分最低，因此商家实施“双承诺”活动的主动性需要加强。
3. 不同企业称号的经营主体满意度存在差异，其中高新技术企业和专精特新中小企业的经营主体满意度低于“其他”类别经营主体，说明消委会组织需加强对前两类经营主体的活动组织与指导。
4. 不同身份消费者满意度存在差异，江门市居民的消费者满意度高于外地游客，说明江门市需要为外地游客营造更好的消费环境。
5. 不同性别消费者满意度存在差异，男性消费者满意度高于女性消费者，说明女性消费场所或针对女性的消费活动需进一步加强“双承诺”活动的推广。

（二）对策建议

针对以上分析结论，提出以下改进建议。

1. 优化服务，加强放心消费“双承诺”活动推广与指导

进一步简化申请程序，提高便利性，确保各类经营主体能够迅速而高效地参与到这一活动中来。通过线上线下的多种宣传渠道，努力提高经营主体和消费者对“双承诺”活动的认知度和参与度，让更多的经营主体和消费者了解并加入到这一活动中。

2. 加强激励，提高商家开展放心消费“双承诺”活动积极性

为了激发商家的积极性，相关部门可设立奖励机制，对那些积极参与并认真执行“双承诺”活动的商家给予物质奖励或荣誉认证。同时，我们将定期评选和表彰表现优秀的“双承诺”商家，以此激励更多的商家加入到这一活动中来。

3. 分析特点，为不同类别经营主体提供差异化服务

针对不同类别的经营主体，应分析其特点，提供差别化的服务。对于高新技术企业和专精特新中小企业，要提供更加精准的活动解读和个性化服务指导。根据经营主体的行业特点和需求，制定相应的支持策略和服务套餐，以满足他们的具体需求。

4. 聚焦群体，优化重点群体的消费服务

为了优化重点群体的消费服务，应重点提升旅游景点及周边商圈的服务质量，以提升外地游客的消费体验。对于女性消费者，要加强市场调研，了解她们的消费习惯和需求，提供更加贴心的服务和产品。同时，还应关注性别差异，促进性别平等的消费环境。通过市场调研和数据分析，深入了解男性和女性消费者满意度不同的原因，并制定相应的改进措施。加强对女性消费场所和活动的推广，确保性别平等在消费活动中得到充分体现。

5. 改善管理，持续开展监测与评估

要持续监测与评估放心消费“双承诺”活动的效果。为此，要建立满意度指数的持续监测机制，定期收集和分析经营主体和消费者的反馈。根据监测结果，我们将及时调整和优化活动策略，确保活动效果与社会需求保持一致。

参考文献

[1] 广东省市场监督管理局. 广东突出“六个强化”全面推进放心消费行动 提振消费信心 [EB/OL]. [2024-08-13]. https://amr.gd.gov.cn/zwdt/xwfbt/content/post_4477020.html.
[2] 汤万金等. 顾客满意测评理论与应用 [M]. 北京：中国计量出版社，2009：36-37.
[3] ACSI. Health insurance benchmarks[CP/OL]. [2019-02-02]. https://www.theacsi.org/index.php?option=com_content&view=article&id=147&catid=&Itemid=212&i=Health+Insurance.
[4] 吴建南，张萌，黄加伟. 基于ACSI的公众满意度测评模型与指标体系研究 [J]. 广州大学学报（社会科学版），2007，(6)1：13-17.
[5] 张新安，田澎，张列平. 上海顾客满意指数测评模型研究 [J]. 工业工程与管理，2002，(05)：29-33.
[6] 朱国玮，郑培. 服务型政府公共满意度测评理论与实践 [M]. 北京：科学出版社，2010：23-31.
[7] 王润良. 基于SEM的政务服务公众满意度指数模型研究 [J]. 五邑大学学报（社会科学版），2022，24(4)：46-52，91.
[8] 张兆阳等. 我国智慧医保政策执行困境及推进策略：基于史密斯政策执行过程模型视角 [J]. 中国医院，2023，27(01)：3-6.
[9] 王润良等. 基于SEM的医疗保险服务满意度指数模型研究——以江门市为例 [J]. 统计与管理，2024，(8)：19-31.
[10] 于本海等. 城乡居民基本医疗保险满意度及影响因素研究——以上海市为例 [J]. 价格理论与实践，2019，(7)：104-107.
[11] 吕慎，田峰，张乾坤. 基于结构方程的快递业收派员工作满意度评估 [J]. 深圳大学学报理工版，2020，37(3)：323-330.

基于高维协方差矩阵估计的投资组合优化策略比较与分析

孙章爽¹, 张婷², 万宇雷³, 王国强^{1*}

1. 上海工程技术大学 数理与统计学院, 上海 201620

2. 菏泽职业学院 信息工程系, 山东 菏泽 274000

3. 上海金仕达软件科技股份有限公司 金仕达研究院, 上海 201203

DOI:10.61369/ASDS.2025070019

摘 要 : 在金融领域, 协方差矩阵的精确估计对于优化投资组合至关重要。研究旨在综合比较无条件协方差矩阵估计和条件协方差矩阵估计在投资组合优化中的表现, 并基于不同维度的股票和商品两种大类资产的数据进行投资组合优化的实证分析。结果显示: 对于低维情形, 无条件协方差矩阵估计在组合收益和组合风险偏差方面均具有突出表现; 对于高维情形, 条件协方差矩阵估计的表现更为有效。

关 键 词 : 协方差矩阵估计; 条件协方差; 压缩估计; 稀疏估计; 投资组合优化

Comparison and Analysis of Portfolio Strategies Based on High-Dimensional Covariance Matrices Estimation

Sun Zhangshuang¹, Zhang Ting², Wan Yulei³, Wang Guoqiang^{1*}

1. School of Mathematics, Physics and Statistics, Shanghai University of Engineering Science, Shanghai 201620

2. Department of Information Engineering, Heze Vocational College, Heze, Shandong 274000

3. Kingstar Research Institute, Shanghai Kingstar Software Technology Co., Ltd., Shanghai 201203

Abstract : In the financial field, the accuracy of covariance matrix estimation is crucial for the portfolio optimization. This paper aims to comprehensively compare the performance of unconditional covariance matrix estimation and conditional covariance matrix estimation in portfolio, with an empirical analysis of portfolio optimization based on data from two broad asset categories: stocks and commodities, across different dimensions. The results indicate that for low-dimensional scenarios, unconditional covariance matrix estimation exhibits outstanding performance in terms of portfolio returns and portfolio risk deviation. However, for high-dimensional scenarios, conditional covariance matrix estimation demonstrates more effective performance.

Keywords : covariance matrix estimation; conditional covariance; shrinkage estimation; sparse estimation; portfolio optimization

引言

协方差矩阵估计是多元统计分析中的重要问题之一, 在生物学、化学和金融学等领域中有着重要的应用价值^[1-3]。在当今大数据时代, 许多实际问题的数据呈现出高维特征, 这导致传统的样本协方差矩阵估计往往表现出奇异性, 特别是在金融领域的马科维茨均值方差模型等实际问题中, 传统的估计方法不再适用。因此, 高维情形下的协方差矩阵估计成为金融统计领域的热门研究问题。

协方差矩阵估计的研究主要涉及以下三种估计。一是基于收缩模型的协方差矩阵估计, 其主要思想是引入一个目标矩阵, 以减小估计量与真实协方差矩阵之间的误差, 包括线性收缩估计与非线性收缩估计。例如, Ledoit 和 Wolf 提出以单位矩阵为目标的线性收缩估计^[4]。在缺失真实协方差矩阵的情况下, Ledoit 和 Wolf 提出了仅对特征值进行收缩的非线性收缩估计^[5], 更多基于收缩模型的协方差

基金项目: 国家自然科学基金资助项目 (12171307, 11971302)。

作者简介:

孙章爽, 上海工程技术大学数理与统计学院, 硕士研究生, 研究方向为高维协方差矩阵估计及其应用;

张婷, 菏泽职业学院信息工程系, 硕士, 副教授, 研究方向为计算机应用技术;

万宇雷, 上海金仕达软件科技股份有限公司金仕达研究院, 博士, 研究方向为金融科技、机器学习与人工智能、金融统计与风险管理、大模型应用等。

通讯作者: 王国强, 上海工程技术大学数理与统计学院, 博士, 教授, 硕士生导师, 研究方向为最优化理论与算法、高维数据统计推断、金融统计与风险管理、综合评估、智能运维、机器学习与人工智能等。邮箱: guoq_wang@hotmail.com。

矩阵估计研究可参考^[6, 7]。二是基于因子模型的协方差矩阵估计,其主要思想是通过因子模型对数据进行降维处理,从而得到可行的估计量。例如 Fan 提出基于严格因子模型的协方差矩阵估计^[8], Fan 提出主成分正交补阈值 (Principal Orthogonal complement Thresholding, POET) 估计器^[9], 杨小卜构造了样本协方差矩阵与因子收缩模型协方差矩阵的凸组合^[9]。最近, Ruan 等提出了基于投影技术的半参数因子模型进行高维协方差矩阵估计^[11]。更多基于因子模型的协方差矩阵估计研究可参考^[12]。三是基于稀疏方法的协方差矩阵估计,其主要思想是引入惩罚函数或阈值算子,以构建满足稀疏性假设的协方差矩阵估计。Rothman 提出基于广义阈值算法的协方差矩阵估计^[13]。更多基于稀疏方法的协方差矩阵估计研究可参考^[14-16]。

在金融领域,协方差矩阵估计的精度直接影响到投资组合优化的效果。特别地,多元广义自回归条件异方差 (Generalized Autoregressive Conditional Heteroskedasticity, GARCH) 模型常被用来预测具有时变性质的条件协方差矩阵。Engle 提出的动态条件相关 (Dynamic Conditional Correlation, DCC) 模型是最常用的多元 GARCH 模型之一^[17]。国内外学者构建出系列高效的拓展 DCC 模型。例如基于区间数据的 DCC-OHLC 模型^[18]、基于动态条件角相关的 DCC 模型及其扩展^[19, 20]等。特别地,刘丽萍将 POET 估计器与 DCC 模型充分结合,构建了基于主成分正交补门限方法的 DCC 模型^[20]。更多条件协方差矩阵估计的研究可参考^[21]。

本文旨在综合比较多种高维协方差矩阵估计方法在投资组合优化中的应用效果。具体而言,本文基于我国股票和商品两种大类资产的数据进行协方差矩阵估计的实证分析,进而完成对市场风险的度量。进一步,以此为基础构建全局风险最小的投资组合优化,并综合对比不同协方差矩阵估计方法在投资组合优化中的表现。相对于其他风险预测或投资组合优化分析的文献,本文的创新之处在于以下两个方面:首先,本文同时对比了无条件协方差矩阵估计和条件协方差矩阵估计在投资组合优化中的估计效果;其次,本文选用多种维度和多种类别资产的资产指数,以识别高维协方差矩阵估计所对应的有效应用场景。

一、研究设计

(一) 协方差矩阵估计

本文利用多种方法对收益率协方差矩阵进行预测与对比分析。由于无条件协方差矩阵估计不包含时间信息,因此我们使用滑动窗口算法对每个时期的协方差矩阵进行估计。假设 21 个交易日为一个月,固定的滑动时间周期为 T (设置为 252 个交易日),则对于第 m 月的无条件协方差矩阵,可采用该月第一天前 T 日的资产收益率来预测该月的协方差矩阵,并以该月的 21 天收益率作为未来回报收益。

下面,简要介绍常见的无条件协方差矩阵估计和条件协方差矩阵估计,详细可参考对应的参考文献。

表 1: 常见的无条件协方差矩阵估计

Table 1: Common unconditional covariance matrix estimation

方法	缩写	协方差矩阵估计公式	文献
单位矩阵为目标的线性收缩估计	I-LW	$\hat{\Sigma} = \rho_1 I + \rho_2 S$	[4]
二次逆收缩估计	QIS	$\hat{\Sigma} = \frac{u_i^T \Sigma^2 u_i}{u_i^T \Sigma u_i}$	[23]
图套索估计	Glasso	$\hat{\Sigma}^{-1} = \arg \max_{\Sigma} (\log \Sigma^{-1} - tr(\Sigma^{-1} S) - \rho \Sigma^{-1} _1)$	[24]
主成分正交补阈值估计	POET	$\hat{\Sigma}_K = \sum_{i=1}^K \hat{\lambda}_i \hat{\xi}_i \hat{\xi}_i' + \hat{R}_K^*$ $\hat{R}_K^* = (\hat{r}_{ij}^*)_{p \times p}, \hat{r}_{ij}^* = \begin{cases} \hat{r}_{ii}, & i = j \\ s_{ij}(\hat{r}_{ij}) I(\hat{r}_{ij} \geq \tau_{ij}), & i \neq j \end{cases}$	[9]

其中, ρ_1, ρ_2 为未知参数, S 为样本协方差矩阵; ρ 为惩罚参数, $(\cdot)^{-1}$ 表示将对角元素变为 0、其他元素保持不变的矩阵算子; $\hat{\lambda}_i$ 和 $\hat{\xi}_i$ 为重排后的特征值与对应特征向量, $s_{ij}(\cdot)$ 为广义收缩函数, $\tau_{ij} > 0$ 为阈值参数。

对于条件协方差矩阵估计,假设 N 个资产的收益率 r_t 向量均满足一定的均值方程,即

$$r_t = \mu + \varepsilon_t, \quad (1)$$

其中, μ 为平均收益率, ε_t 为随机变量,满足 $E[\varepsilon_t] = 0$ 。则最终的条件协方差矩阵可通过

$$H_t = D_t R_t D_t \quad (2)$$

计算得到,其中 $D_t = \text{diag}(\sqrt{\sigma_{1t}^2}, \dots, \sqrt{\sigma_{Nt}^2})$, σ_i^2 为单个资产的方差。

表 2: 常见的条件协方差矩阵估计

Table 2: Common conditional covariance matrix estimation

方法	缩写	协方差矩阵估计公式	文献
动态条件相关模型	DCC	$\begin{cases} Q_t = (1 - \theta_1 - \theta_2)S + \theta_1 \varepsilon_{t-1} \varepsilon_{t-1}' + \theta_2 Q_{t-1} \\ R_t = \text{diag}(Q_t)^{-1} Q_t \text{diag}(Q_t)^{-1} \end{cases}$	[17]
使用 OHLC 价格的动态条件相关模型	DCC-OHLC	$R_t = (1 - \theta_1 - \theta_2)R + \theta_1 \Phi_{t-1} + \theta_2 R_{t-1}$	[18]
动态条件角相关模型	DCAC	$R_t = (1 - \theta_1 - \theta_2)R + \theta_1 \Gamma_{t-1} + \theta_2 R_{t-1}$	[19]
主成分正交补门限动态条件相关模型	poetDCC	$\begin{cases} \tilde{Q}_t = (1 - \theta_1 - \theta_2)S + \theta_1 \varepsilon_{t-1} \varepsilon_{t-1}' + \theta_2 \tilde{Q}_{t-1} \\ \tilde{R}_t = \text{diag}(\tilde{Q}_t)^{-1} \tilde{Q}_t \text{diag}(\tilde{Q}_t)^{-1} \end{cases}$	[20]

其中, ε_t 为标准化残差, Q_t 为估计过程中引入的辅助矩阵, \tilde{Q}_t 表示模型额外对矩阵 Q_t 进行了谱分解与正交补门限方法的应用; θ_1, θ_2 为相关性参数,满足 $0 < \theta_1, \theta_2 < 1$ 且 $\theta_1 + \theta_2 < 1$; Φ_t 为 Popov 相关矩阵, Γ_t 为角相关矩阵。值得注意的是, DCC 模型、poetDCC 模型与 DCAC 模型的单资产方差 σ_i^2 均满足 GARCH (1, 1) 过程,即

$$\sigma_t^2 = \omega_0 + \sum_{i=1}^q \alpha_i r_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2. \quad (3)$$

而 DCC-OHLC 使用 Range-GARCH 模型:

$$\sigma_t^2 = \omega_0 + \sum_{i=1}^q \alpha_i h_{pt-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2, \quad (4)$$

进行单变量方差估计, 其中 $h_{pt}^2 = [\ln(H_t / L_t)]^2 / (4 \ln 2)$ 是基于最高价 H_t 与最低价 L_t 的区间估计。

(二) 投资组合优化模型

为了评估上述方法在协方差矩阵估计中的效果及其差异, 本文采用了不设定预期收益的全局最小风险 (Global Minimum Variance, GMV) 策略^[25]来避免预期收益过高的影响, 具体优化模型如下:

$$\begin{aligned} \min \quad & \omega' \Sigma \omega \\ \text{s.t.} \quad & \sum_{i=1}^N \omega_i = 1, \end{aligned} \quad (5)$$

其中, N 为资产数量, ω 为资产组合权重, Σ 使用 2.1 节所述估计进行预测与替换。

(三) 预测效果评估

为了比较各协方差矩阵估计在上述 GMV 策略中的性能差异, 本文使用样本外组合收益和样本外组合风险进行评估。

1. 样本外组合收益评估

对于组合收益, 本文采用 AV, SD, IR 和 SR 进行评估, 详细可参考^[26]。假设第 i 月的平均收益分别为 r_i , 在样本外预测中共包含 n 个月, 则样本外年化平均组合收益为:

$$AV = 252 \times \frac{1}{n} \sum_{i=1}^n r_i. \quad (6)$$

类似地, 假设这 n 个月间的平均收益为 \bar{r} , 则样本外年化标准差为:

$$SD = \sqrt{252} \times \frac{1}{n-1} \sqrt{\sum_{i=1}^n (r_i - \bar{r})^2}. \quad (7)$$

若以 0 为基准, 则可以计算样本外信息比率 $IR = AV / SD$ 。若以无风险利率 r_f 作为基准, 则可以计算样本外夏普比率 $SR = (AV - r_f) / SD$ 。

2. 样本外组合风险评估

对于组合风险偏差, 假设 ω 为预测资产权重, Σ 为已实现资产风险, $\hat{\Sigma}$ 为预测资产风险, 则可采用组合风险的平方追踪误差 (Square Tracking Error, STE), 即

$$STE = (\sqrt{\hat{\omega}' \hat{\Sigma} \omega} - \sqrt{\omega' \Sigma \omega})^2 \quad (8)$$

进行评估。STE 越接近于 0, 则代表预测的风险变化越接近于真实风险变化且组合风险越小。

(四) 样本选取

为了对比分析不同协方差矩阵估计的预测效果, 本文选择了不同维度的两种大类资产来进行检验, 包括以上证 50、中证

100 和沪深 300 指数为代表的股票, 以及以南华商品综合指数 (NHCI) 为代表的商品。由于不同公司创立时间存在差异, 因此本文选择数据覆盖范围最广的时间段 2014 年 1 月至 2024 年 7 月的所有日线数据进行分析。表 3 展示了不同维度指数在样本区间日收益率的描述性统计结果。

表 3: 描述性统计分析
Table 3: Descriptive statistical analysis

	指数 维度	均值 $\times 10^{-4}$	最小值	最大值	标准差	偏度	超额峰 度
NHCI	20	4.4843	-0.1700	0.1700	0.0153	0.0178	3.8272
上证 50	50	2.8099	-0.6461	0.1013	0.0222	-0.5309	16.8943
中证 100	100	5.0338	-0.5170	0.2001	0.0249	-0.4105	12.6600
沪深 300	300	4.0650	-0.5170	0.2001	0.0243	-0.2159	10.8349

进一步, 计算大类资产指数间的相关系数, 如图 1- 图 3 展示。

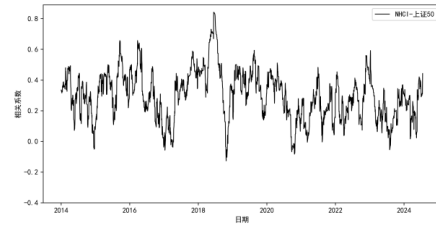


图 1: NHCI- 上证 50 资产间动态相关性

Fig. 1: Dynamic correlation between NHCI-SSE 50 assets

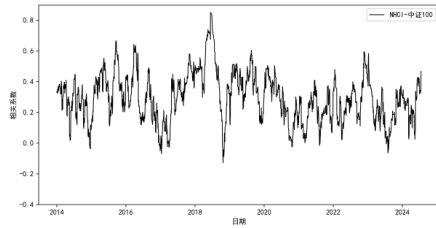


图 2: NHCI- 中证 100 资产间动态相关性

Fig. 2: Dynamic correlation between NHCI-CSI 100 assets

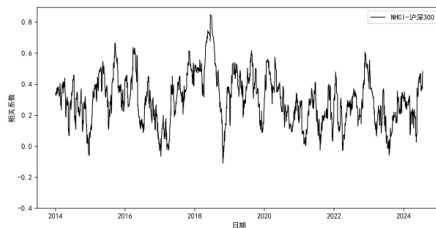


图 3: NHCI- 沪深 300 资产间动态相关性

Fig. 3: Dynamic correlation between NHCI-HS 300 assets

图 4 呈现了不同维度指数成分间的相关性热力图。可以得到, 各资产间相关系数均不超过 0.95, 即无高度相关资产, 符合传统协方差矩阵估计和动态条件相关模型的限制, 能够使用上述方法进行协方差矩阵估计。

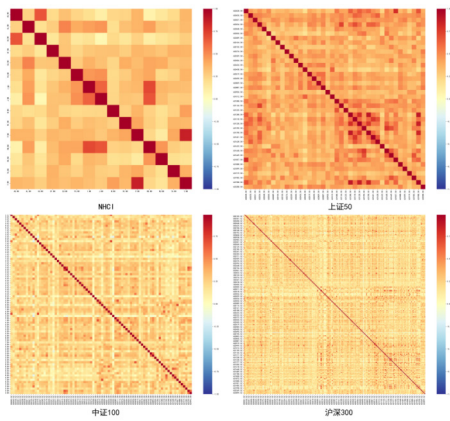


图4：不同指数成分股间相关性

Fig. 4: The correlation among the constituent stocks of different indices

二、实证检验

（一）投资收益评估

表4汇总了等权重估计策略和8种协方差矩阵估计在投资组合

表4：不同估计器在投资组合优化中的收益预测表现

Table 4: The performance of different estimators in the return prediction of portfolio optimization

		等权重	I-LW	QIS	Gllasso	POET	DCC	poetDCC	DCAC	DCC-OHLC
NHCI	AV	3.82894	3.85053	3.82631	3.82364	3.82924	3.71015	3.75442	3.41183	3.26705
	SD	4.50922	4.50374	4.58795	4.50662	4.59185	5.89029	4.78411	4.11373	4.18988
	IR	0.84914	0.85496	0.83399	0.84845	0.83392	0.62988	0.78477	0.82938	0.77975
	SR	0.31341	0.31859	0.30746	0.31242	0.30784	0.21976	0.27983	0.24215	0.20319
上证50	AV	5.55536	5.54165	5.55263	5.50682	5.83244	5.38105	5.37941	5.24610	4.30779
	SD	5.56154	6.09980	6.32674	6.16631	10.23962	6.19460	5.95298	5.12468	5.40368
	IR	0.99889	0.90850	0.87765	0.89305	0.56960	0.86867	0.90365	1.02369	0.79720
	SR	0.56453	0.51247	0.49582	0.50129	0.33368	0.47870	0.49785	0.55231	0.35015
中证100	AV	6.19891	6.16588	6.13778	6.04795	6.45002	6.55631	5.88006	5.40376	4.88705
	SD	7.01976	6.18041	6.32899	6.18823	7.51895	3.45990	5.67271	4.66914	5.49440
	IR	0.88307	0.99765	0.96979	0.97733	0.85783	1.89494	1.03655	1.15733	0.88946
	SR	0.53894	0.60679	0.58810	0.58696	0.53655	1.19674	0.61071	0.63996	0.44980
沪深300	AV	3.35751	3.01501	3.01005	2.97209	2.87005	3.36201	3.11137	2.86376	2.67106
	SD	5.99815	4.77110	4.76993	5.43716	5.56813	4.83269	4.67697	3.51101	3.37101
	IR	0.55976	0.63193	0.63105	0.54663	0.51544	0.69568	0.66525	0.81565	0.79236
	SR	0.15702	0.12562	0.12461	0.10233	0.08160	0.19582	0.14874	0.12762	0.07575

（二）投资风险评估

表5展示了不同协方差矩阵估计在投资组合优化中的STE表现。由此可知，在维度较低时，无条件协方差矩阵估计表现出较好性能。例如，对于NHCI指数，线性收缩I-LW估计呈现出最

优化中的预测表现，粗体代表该指标下的最优值，斜体代表在等权重策略为最优时剩余方法中的次优值。可以看到，在不同情况下，协方差矩阵估计的表现呈现显著差异。从指数角度来看，无条件协方差矩阵估计在维度较小、低峰值时更具有优势，其中线性收缩I-LW估计预测效果最好，得到了最高的AV，IR和SR，其AV相对于条件协方差矩阵估计平均提升8.90%，SR平均提升34.86%。随着指数维度提升，条件协方差矩阵估计逐渐展现其优势。对于中证100和沪深300，DCC模型得到最大的AV值和SR值，且AV相对于无条件协方差矩阵估计平均提高5.74%和13.32%，SR值平均提升106.48%和80.41%。同时，对于预测收益波动性的SD指标，条件协方差矩阵估计总是得到最低值。

综上，本文可以得到如下结论：第一，在低维且分布平坦的数据集中，无条件协方差矩阵估计在收益预测方面呈现出显著优势；第二，随着维度增长，条件协方差矩阵估计在收益预测上逐渐展现其优越性，代表超额收益的夏普比率SR也保持相同趋势；第三，条件协方差矩阵估计在预测收益的稳定性方面具有比较稳定的突出表现。

低的STE，其相对于条件协方差矩阵估计平均降低99.51%。随着资产维度提升，条件协方差矩阵估计凸显其优势，其中DCC模型表现出强劲的性能。相对于无条件协方差矩阵估计，DCC模型在上证50和中证100中组合风险平均降低62.68%和92.55%。

表5：不同估计器在投资组合优化中的风险预测表现

Table 5: The performance of different estimators in risk prediction of portfolio optimization

	等权重	I-LW	QIS	Gllasso	POET	DCC	poetDCC	DCAC	DCC-OHLC
NHCI	0.12213	0.00005	0.00006	0.00006	0.00006	0.01050	0.00054	0.03318	0.00009
上证50	0.00515	0.00076	0.00079	0.00089	0.00394	0.00063	0.00194	0.00767	0.00092
中证100	0.00099	0.00024	0.00024	0.00040	0.00179	0.00005	0.00076	0.02190	0.00056
沪深300	0.00005	0.00048	0.00035	0.00051	0.00096	0.00054	0.00071	0.01737	0.00038

由上，不难发现协方差矩阵估计在投资组合优化中呈现出两方面的差异，一是无条件协方差矩阵估计在低维情形下对于组合收益与组合风险偏差具有较大优势；二是条件协方差矩阵估计在高维数据集中更具有亮眼的表现。

（三）稳健性检验

为了验证上述结果的稳健性，本文分别采用50天和150天作

为滑动时间周期 T ，使用相同指数进行验证，得到的研究结果与上述分析一致，说明了结果的稳健性，具体结果如

表6- 表9所示。特别地，随着滑动时间周期减少，DCAC 模型在沪深300指数中呈现更集中的优势；同时，随着滑动时间周期减少，可代表数据集维度信息的浓度比 N/T 增加，DCC-OHLC 模型在 NHCI 中获得更低的 STE，最终呈现仍与原结论保持一致。

表6：组合收益预测稳定性检验（ $T=50$ ）
Table 6: Stability test of portfolio return prediction ($T=50$)

		等权重	I-LW	QIS	Gllasso	POET	DCC	poetDCC	DCAC	DCC-OHLC
NHCI	AV	4.05645	4.08493	4.08812	4.07778	4.12882	4.01431	4.01323	3.85316	3.84385
	SD	5.00450	4.66697	4.74811	4.69699	4.90506	6.87263	5.26804	5.05354	5.13216
	IR	0.81056	0.87528	0.86100	0.86817	0.84175	0.58410	0.76181	0.76247	0.74897
	SR	0.32786	0.35767	0.35223	0.35386	0.34926	0.23261	0.30325	0.28445	0.27828
上证50	AV	5.69840	5.79437	5.79567	5.75936	5.79683	5.74946	5.69737	5.57430	5.25689
	SD	6.55726	5.91509	5.74654	6.02382	6.98647	6.78131	7.05989	5.93142	6.36033
	IR	0.86902	0.97959	1.00855	0.95610	0.82972	0.84784	0.80701	0.93979	0.82651
	SR	0.50062	0.57120	0.58818	0.55507	0.48396	0.49161	0.46483	0.53252	0.44671
中证100	AV	6.23216	6.27769	6.31985	6.23944	6.27517	6.27241	6.00119	5.77376	5.47845
	SD	5.80007	5.63015	6.88436	5.57696	5.89607	2.89026	5.73981	4.93186	5.32168
	IR	1.07450	1.11501	0.91800	1.11879	1.06430	2.17019	1.04554	1.17071	1.02946
	SR	0.65800	0.68595	0.56711	0.68563	0.65458	1.33438	0.62467	0.68089	0.57552
沪深300	AV	3.17462	3.17071	3.22526	2.93947	3.00517	2.65814	3.20546	3.44204	3.17894
	SD	4.61773	4.02921	4.08423	4.75465	4.38141	5.10566	5.37155	3.51332	3.88470
	IR	0.68748	0.78693	0.78969	0.61823	0.68589	0.52063	0.59675	0.97971	0.81832
	SR	0.16435	0.18739	0.19822	0.11016	0.13454	0.04749	0.14703	0.29213	0.19648

表7：组合风险偏差预测稳定性检验（ $T=50$ ）
Table 7: Stability test of portfolio risk deviation prediction ($T=50$)

	等权重	I-LW	QIS	Gllasso	POET	DCC	poetDCC	DCAC	DCC-OHLC
NHCI	0.10306	0.00018	0.00020	0.00024	0.00028	0.01010	0.00059	0.03389	0.00007
上证50	0.00483	0.00120	0.00130	0.00137	0.00328	0.00070	0.00211	0.00813	0.00087
中证100	0.00082	0.00057	0.00509	0.00064	0.00133	0.00005	0.00079	0.02201	0.00062
沪深300	0.00004	0.00061	0.00125	0.00059	0.00086	0.00059	0.00071	0.01550	0.00036

表8：组合收益预测稳定性检验（ $T=150$ ）
Table 8: Stability test of portfolio return prediction ($T=150$)

		等权重	I-LW	QIS	Gllasso	POET	DCC	poetDCC	DCAC	DCC-OHLC
NHCI	AV	3.95321	3.92085	3.88901	3.88931	3.89221	3.84141	3.90572	3.57597	3.47681
	SD	4.46090	4.54572	4.60386	4.57761	4.65415	5.85242	4.88533	4.22581	4.25964
	IR	0.88619	0.86254	0.84473	0.84964	0.83629	0.65638	0.79948	0.84622	0.81622
	SR	0.34467	0.33112	0.32002	0.32192	0.31725	0.24361	0.30500	0.27457	0.24911
上证50	AV	5.63342	5.71814	5.74577	5.63352	5.99712	5.66276	5.54032	5.41972	4.62547
	SD	5.56690	6.02514	6.27589	6.48502	7.89466	5.48152	5.93485	5.17605	5.55745
	IR	1.01195	0.94905	0.91553	0.86870	0.75964	1.03306	0.93352	1.04708	0.83230
	SR	0.57801	0.54811	0.53061	0.49619	0.45365	0.59237	0.52649	0.58037	0.39762
中证100	AV	6.15572	6.22662	6.22611	6.17706	6.29077	6.63589	5.96055	5.56262	5.17472
	SD	6.14951	6.12871	6.09755	5.99338	7.34056	4.06715	5.90069	4.87056	5.36791
	IR	1.00101	1.01598	1.02108	1.03065	0.85699	1.63158	1.01014	1.14209	0.96401
	SR	0.60818	0.62182	0.62491	0.62759	0.52790	1.03763	0.60075	0.64611	0.51399
沪深300	AV	3.16179	3.00623	3.21916	2.90387	2.88240	3.06583	3.06143	2.93288	2.75381
	SD	5.03951	3.90996	7.09539	4.85133	4.58604	4.78131	4.64535	2.72713	3.21297
	IR	0.62740	0.76887	0.45370	0.59857	0.62852	0.64121	0.65903	1.07544	0.85709
	SR	0.14805	0.15103	0.11324	0.10063	0.10177	0.13597	0.13901	0.18964	0.10523

表9：组合风险偏差预测稳定性检验（T=150）
Table 9: Stability test of portfolio risk deviation prediction (T=150)

	等权重	I-LW	QIS	Glasso	POET	DCC	poetDCC	DCAC	DCC-OHLC
NHCI	0.11114	0.00010	0.00010	0.00012	0.00011	0.01039	0.00054	0.03333	0.00009
上证50	0.00469	0.00095	0.00098	0.00109	0.00420	0.00068	0.00199	0.00775	0.00093
中证100	0.00090	0.00037	0.00036	0.00049	0.00173	0.00005	0.00076	0.02184	0.00060
沪深300	0.00004	0.00056	0.00337	0.00056	0.00089	0.00056	0.00072	0.01706	0.00035

三、结束语

本文综合比较了无条件协方差矩阵估计和条件协方差矩阵估计在投资组合优化中的效果。研究发现，两类协方差矩阵估计在不同维度的数据集下表现出显著差异。对于低维情形，无条件协方差矩阵估计在组合收益与组合风险偏差方面均展现其明显优势。对于高维情形，条件协方差矩阵估计在上述两方面具有突出

性能。在两种不同维度情形下，条件协方差矩阵估计均拥有更稳定的组合收益预测。特别地，研究结果显示主成分方法与阈值化方法带来的稀疏性使无条件协方差矩阵估计在维度提升时仍可保留部分优势，且揭示了 DCAC 模型在尖峰情形下的超额收益能力得到提升。

在未来的工作中，可以考虑更高效的算法应用于协方差矩阵的估计中，以降低条件协方差矩阵估计面对高维数据时的计算压力。

参考文献

[1]ENGEL J, BUYDENS L, BLANCHET L. An overview of large-dimensional covariance and precision matrix estimators with applications in chemometrics[J]. Journal of Chemometrics, 2017, 31(4): e2880.

[2]SUN R, MA T, LIU S, et al. Improved covariance matrix estimation for portfolio risk measurement: a review[J]. Journal of Risk and Financial Management, 2019, 12(1): 48.

[3]LI D. Estimation of large dynamic covariance matrices: A selective review[J]. Econometrics and Statistics, 2024, 29: 16–30.

[4]LEDOIT O, WOLF M.A well-conditioned estimator for large-dimensional covariance matrices[J]. Journal of Multivariate Analysis, 2004, 88(2): 365–411.

[5]LEDOIT O, WOLF M. Nonlinear shrinkage estimation of large-dimensional covariance matrices[J]. The Annals of Statistics, 2012, 40(2): 1024–1060.

[6]ZHANG Y, TAO J Y, YIN Z X, et al. Improved large covariance matrix estimation based on efficient convex combination and its application in portfolio optimization[J]. Mathematics, 2022, 10(22): 4282.

[7]LEDOIT O, WOLF M. The power of (non-)linear shrinking: A review and guide to covariance matrix estimation[J]. Journal of Financial Econometrics, 2020, 20(1): 187–218.

[8]FAN J Q, FAN Y Y, LV J C. High dimensional covariance matrix estimation using a factor model[J]. Journal of Econometrics, 2008, 147(1): 186–197.

[9]FAN J Q, LIAO Y, MARTINA M. Large covariance estimation by thresholding principal orthogonal complements[J]. Journal of the Royal Statistical Society Series B: Statistical Methodology, 2013, 75(4): 603–680.

[10] 杨小卜. 基于因子收缩方法的高维协方差估计 [J]. 数学的实践与认识, 2022, 52(10): 94–103.

[11]RUAN Y W, ZHANG X F, LIU Y J. A semi-parametric factor-GARCH model for high dimensional covariance matrix estimation[J]. Journal of the Korean Statistical Society, 2025: 1–32.

[12]FAN J Q, LIAO Y, LIU H. An overview of the estimation of large covariance and precision matrices[J]. The Econometrics Journal, 2016, 19(1): 1–32.

[13]ROTHMAN A J, LEVINA E, ZHU J.Generalized thresholding of large covariance matrices[J]. Journal of the American Statistical Association, 2009, 104(485): 177–186.

[14]XIAO Y H, LI P L, LU S. Sparse estimation of high-dimensional inverse covariance matrices with explicit eigenvalue constraints[J]. Journal of the Operations Research Society of China, 2021, 9: 543–568.

[15] 宋鹏, 胡永宏. 基于已实现协方差矩阵的高维金融资产投资组合应用 [J]. 统计与信息论坛, 2017, 32(08): 63–70.

[16] 王鑫, 孔令臣, 王力群. 高维协方差矩阵的估计问题 [J/OL]. 运筹学学报, 1–14[2025-05-30]. <http://kns.cnki.net/kcms/detail/31.1732.O1.20240202.1134.006.html>.

[17]ENGLE R. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models[J]. Journal of Business and Economic Statistics, 2002, 20(3): 339–350.

[18]FISZEDER P, FALDZINSKI M, MOLNAR P. Modeling and forecasting dynamic conditional correlations with opening, high, low, and closing prices[J]. Journal of Empirical Finance, 2023, 70: 308–321.

[19]JARJOUR R, CHAN K S.Dynamic conditional angular correlation[J]. Journal of Econometrics, 2020, 216(1): 137–150.

[20]SUN Z S, GAO X Y, LUO K Y, et al. Enhancing high-dimensional dynamic conditional angular correlation model based on garch family models: Comparative performance analysis for portfolio optimization[J]. Finance Research Letters, 2025, 75: 106808.

[21] 刘丽萍, 马丹, 白万平. 大维数据的动态条件协方差阵的估计及其应用 [J]. 统计研究, 2015, 32(6): 105–112.

[22] 刘进. 条件协方差矩阵的估计方法研究综述 [J]. 统计与决策, 2019, 35(23): 23–27.

[23]LEDOIT O, WOLF M. Shrinkage estimation of large covariance matrices: Keep it simple, statistician?[J]. Journal of Multivariate Analysis, 2021, 186: 104796.

[24]FRIEDMAN J, HASTIE T, TIBSHIRANI R. Sparse inverse covariance estimation with the graphical lasso[J]. Biostatistics, 2008, 9(3): 432–441.

[25]ALEXANDER K, MEMMEL C. Estimating the global minimum variance portfolio[J]. Schmalenbach Business Review, 2006, 58: 332–348.

[26]DE NARD G, ENGLE R F, LEDOIT O, et al. Large dynamic covariance matrices: Enhancements based on intraday data[J]. Journal of Banking and Finance, 2022, 138: 106426.

基于 T 检验的现场展示教学模式影响规律探索 ——以统计学专业实验案例教学为例

车金星, 陈锦文, 曹寒问, 董炜, 张毓华, 万冰蓉, 曾康

江西水利电力大学 理学院, 江西 南昌 330099

DOI:10.61369/ASDS.2025070020

摘 要 : 实践教学作为培养学生创新能力和实践能力的关键环节, 面临着新的要求和挑战。现场展示教学模式不仅可以提升学生的学业成绩, 而且能够调动学生的组织和实践能力。为了认证现场展示教学模式的有效性, 将 T 检验引入到教育教学数据分析之中, 设计了基于配对样本 T 检验的显著差异性实验方案, 利用回归分析定量研究了现场展示答辩对统计模式实验的影响, 验证了学业成绩的提升显著性及其三个因素对学生学业成绩的影响规律。结果显示, 现场展示的实验组表现明显优于其对照组表现, 找出影响学生学业成绩的主要因素, 为进一步的教育教学改革提供了一定的理论依据。

关 键 词 : 现场展示教学模式; 配对样本 T 检验; 回归分析; 显著差异性

Exploration of Influence Patterns of Field Demonstration Teaching Mode Based on T-test — Taking Experimental Case Teaching of Statistics as an Example

Che Jinxing, Chen Jinwen, Cao Hanwen, Dong Wei, Zhang Yuhua, Wan Bingrong, Zeng Kang

Jiangxi University of Water Resources and Electric Power, School of Science, Nanchang, Jiangxi 330099

Abstract : Practical teaching, as a key link in cultivating students' innovative and practical abilities, faces new requirements and challenges. On-site display teaching mode can not only improve students' academic performance (assessment scores), but also mobilize their organizational and practical abilities. In order to verify the effectiveness of the on-site display teaching mode, T-test was introduced into the analysis of educational and teaching data. A significant difference experimental scheme based on paired sample T-test was designed. Regression analysis was used to quantitatively study the impact of on-site display defense on the statistical mode experiment, verifying the significance of academic performance improvement and the impact of three factors on students' academic performance. The results showed that the experimental group in the on-site display group performed significantly better than their control group. The main factors affecting students' academic performance were identified, providing a theoretical basis for further educational and teaching reforms.

Keywords : on-site display teaching mode; paired sample T-test; regression analysis; significant differences

引言

教育统计学领域在分析学生成绩的影响因素方面, 长期以来取得了一系列研究成果。早期研究, 如马文珠等人在2009年的工作, 通过比较传统与模式教学法, 发现后者能显著提高学生的分析能力和理论理解^[1]。随后, 张喜娟和丁钊鹏在2010年利用因子分析法, 从多因素问卷调查中提取出影响学生成绩的主要因素, 为教育改革提供了理论支持^[2]。进入2012年, 肖艳群采用决策树 C4.5 算法, 研究高职学生成绩的影响因素^[3]。同年, 任兆林对不同学期课程设置的影响进行了统计分析, 发现学期期间学生成绩存在显著差异^[4]。2016年, 李桂荣和李向辉基于河南省中职学校数据, 使用多层模型分析了个体和学校因素对学业成绩的影响, 突出了家庭和个体因素的重要

基金项目: 江西省高等学校教学改革研究课题“基于案例教学的《统计套刊》课程思政教学改革与实践”(编号: JXJG-23-18-21); 江西“双干计划”项目“复杂数据的智能集成预测建模及其应用研究”(编号: jxsq2019201064)。

作者简介:

车金星(第一作者), 男, 汉族, 江西九江人, 理学博士, 教授, 博士研究生导师; 研究方向: 不确定性人工智能与综合能源系统、能源与环境统计;

陈锦文, 男, 汉族, 江西宜春人, 应用统计学本科生; 研究方向预测模型与人工智能; 曹寒问, 女, 汉族, 湖南醴陵人, 副教授; 研究方向为数理统计与人工智能;

董炜, 男, 汉族, 浙江东阳人, 应用统计硕士研究生; 主要研究方向为时间序列预测、多模态特征融合与不确定性建模;

张毓华, 男, 汉族, 江西赣县人, 理学博士, 副教授, 硕士研究生导师; 研究方向为能源经济统计;

万冰蓉, 女, 汉族, 江西吉安人, 副教授; 研究方向为应用统计;

曾康, 男, 汉族, 江西抚州人, 经济学博士, 讲师; 研究方向为宏观经济统计分析。

性^[5]。金昆萌在2017年利用贝叶斯网络分析继续教育学生成绩的影响因素，通过后验概率值确定因素间的关系^[6]。王帅和莫惠玲探索案例教学在保证授课效果、提升学生素质的教学改革成效^[7]。这些研究不仅标志着教育统计学在方法论上的演进，从传统的统计分析到现代的数据挖掘技术，而且反映了教育领域对于提高教学质量和学生成绩的持续关注和探索。

近期，随着 AI 赋能教育的快速发展，研究人员通过应用先进的数据挖掘技术和机器学习算法，不断探索和改进教育效果的预测和分析方法。Zhang 等人 (2024) 提出了一个集成先进教育数据挖掘模块的学习管理系统 (LMS)，通过分析学生表现记录，以规划有效的教学策略^[8]。同样，Changoa 等人 (2021) 发现使用集成方法和选择最佳属性的方法在预测学生最终表现时最为有效^[9]。针对大语言模型的兴起，Zhang 等人发现学生更倾向于使用 ChatGPT，激发学生自主提问在教育领域具有较大潜力^[10]。陈超辉等人 (2025) 利用柯氏模型评估方法，通过分析启发式教学模式在大气物理学中的应用效果，预测学生在该课程中的表现^[11]。这些研究表明，采用调动学生自主学习的教学模式，可以显著提高教育预测的准确性和教学策略的有效性。

本研究的目的在于评估现场展示答辩对学生学业成绩的影响，探索现场展示的现场核验、提问、理论回顾对学生学习效果的潜在影响，并进一步理解这些教学方法在统计模式实验课程中的应用效果。本研究的创新之处在于引入现场展示答辩机制，考察其对学业成绩的提升作用；通过提问环节，进一步探索互动式教学对学生学习效果的影响，利用回归和 T 检验验证现场展示答辩教学方法的有效性。通过对比不同教学模式下的学生表现，为教育工作者在做出教学决策提供支持。同时，本研究的发现也将为统计模式实验教学方法的改进提供理论依据和实践指导，从而推动教育改革的深入发展，提高教学质量和学生的学习成效。

一、数据样本与研究方法

(一) 数据样本

在多次实践教学中，我们发现现场展示教学模式可有效提高教学质量和学生的学习成效。为了从统计上分析其有效性，本研究选取江西水利电力大学统计学专业二年级2个班级的学生作为研究对象，他们的8次实验成绩构成了样本数据见表1。研究设计中，第1次实验作为对照组，采用传统的授课方式；后7次实验作为对照组，分别考虑统计模式实验现场展示教学模式的三个因素（现场核验、提问、理论回顾）。

表1：应用统计学专业《智能优化算法（实验）》
二年级学生的8次实验成绩

传统授课教学模式	现场展示教学模式1 (因素：现场核验)			现场展示教学模式2 (因素：现场核验 + 提问)				现场展示教学模式3 (因素：现场核验 + 提问 + 理论回顾)			
	成绩1	成绩2	成绩3	成绩4	成绩5	成绩6	成绩7	成绩8	成绩9	成绩10	成绩11
	96	96	98	99	99	98	95	99	96	97	98
	86	90	96	98	92	95	92	97	94	96	97

	85	94	98	97	99	99	95	98	93	95	96
	92	93	99	99	99	99	99	99	92	94	95

为了进一步了解数据的总体情况，表2对不同教学模式的数据做了描述性分析。从表2的描述性分析结果来看，现场展示教学模式1、教学模式2和教学模式3相较于传统授课教学模式，在学生实验成绩的平均值上表现出了更高的水平，显示出现场展示教学模式在提高学生成绩和促进学生参与方面显示出积极的趋势，但不同模式间的效果存在差异，这可能与教学内容的呈现方式、学生的互动机会以及教学资源的有效利用有关。

表2：学生在4组教学模式实验成绩描述性统计分析表

	N	最小值	最大值	平均值	标准差
传统授课教学模式	85	60	98	87.55	6.06

现场展示教学模式1	85	60	99	93.72	5.45
现场展示教学模式2	85	60	99	95.08	6.13
现场展示教学模式3	85	40	99	94.12	9.22

为了更清晰了解不同模式下的成绩变化，图1绘制了不同模式的成绩平均图。从图1看出，学生的成绩随着实验次数的增加而提高，现场展示教学模式1和现场展示教学模式2的频数随着成绩的增加而增加，呈现出正偏态分布的趋势。传统授课教学模式的频数分布则相对均匀，没有明显的偏态。这些趋势可能反映出现场展示模式可能促进了学生间的多样化表现，而传统授课模式则可能在一定程度上限制了学生表现的多样性。此外，图2还做了不同模式成绩的箱线图。从图2中看出，学生在4组模式的成绩变化呈现一定的上升趋势。现场展示教学模式1和模式2的成绩更为集中，但存在一定数量的异常值。传统授课模式的成绩分布较为分散，异常值数量较少，但成绩波动较大。教学模式3虽然异常值数量较少，但成绩分布适中，且有一些成绩特别低的学生。这些趋势反映了不同教学模式对学生成绩分布和一致性的影响。

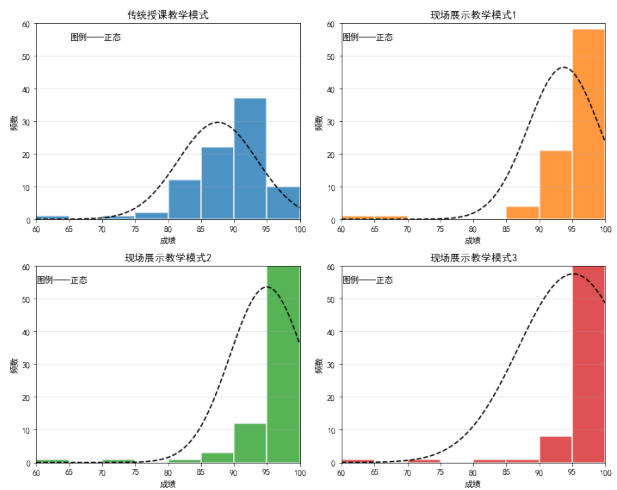


图1：学生在4组教学模式下成绩条形图

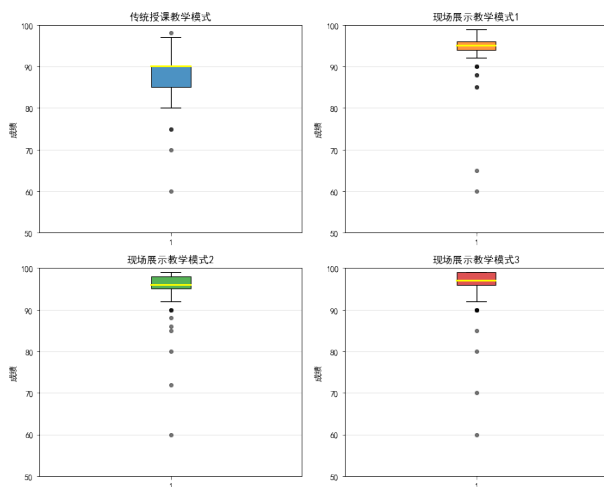


图2: 学生在4组教学模式下成绩箱线图

（二）斯皮尔曼相关系数

斯皮尔曼相关系数（Spearman's rank correlation coefficient）是一种衡量两个变量单调关系的非参数方法，它通过将数据转换为等级，然后计算这些等级之间的相关性。斯皮尔曼相关系数利用以下公式来计算两个变量之间的等级相关性：

$$\rho = 1 - \frac{6 \sum d^2}{n(n^2 - 1)} \quad (1)$$

其中， $\sum d^2$ 是每对观测值等级差平方的总和， n 是样本量。 ρ 表示斯皮尔曼相关系数衡量等级之间的差异程度：等级差越大， $\sum d^2$ 就越大，相关系数 ρ 就越接近于 0，表明两个变量之间的相关性越弱；相反，如果等级差很小， \sum 就很小，相关系数 ρ 就越接近于 1 或 -1，表明两个变量之间存在较强的正向或负向相关关系。

（三）配对样本 T 检验

配对样本 T 检验（Paired two-sample T-test）是一种统计方法，用于评估两组相关样本的均值差异是否显著。检验的核心在于计算差值的均值（ \bar{D} ）和标准差（ s_D ），然后使用以下公式来构建 T 统计量：

$$T = \frac{\bar{D}}{s_D / \sqrt{n}} \quad (2)$$

其中， \bar{D} 是差值的样本均值， s_D 是差值的标准差， n 是配对样本的数量。该统计量遵循自由度为 $n-1$ 的 T 分布。如果计算出的 T 统计量对应的 p 值小于预定的显著性水平，则拒绝原假设，认为两组相关样本的均值存在统计学上的显著差异。

（四）回归分析

回归分析是一种统计方法，它基于一个或多个自变量来预测因变量。这种方法的核心是建立一个数学模型，描述变量之间的线性关系。在多元线性回归中，模型形式为 $Y = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + \dots +$ 。这里，每个 $a_i, i=1,2,\dots,n$ 表示相应自变量对因变量的影响程度。

二、探究不同展现方式对统计实验的影响

（一）不同成绩差异性分析

为了探究不同展示教学模式与传统教学模式的差异，本文建立以下假设。

H_0 ：现场展示教学模式 1, 2, 3 与传统教学模式没有显著性差异

H_1 ：现场展示教学模式 1, 2, 3 与传统教学模式有显著性差异

在本研究中，配对样本 T 检验被用来评估不同教学方法对统计模式实验中学业成绩的影响。分别做了两次配对样本 T 检验具体操作如下：

第一次配对样本 T 检验比较了传统教学模式与现场展示教学模式 1、2、3 之间的差异，检验结果见表 3。结果显示，实验成绩 1 与其他三次实验成绩之间存在显著性差异，说明现场展示答辩和提问环节对提升学生的实验成绩有显著影响。

表3：第一次 T 检验结果表

模式比较	p 值	T 值	显著性
传统授课教学模式 vs 现场展示教学模式 1	8.15e-17	-10.427	是
传统授课教学模式 vs 现场展示教学模式 2	2.03e-22	-13.322	是
传统授课教学模式 vs 现场展示教学模式 3	1.505123e-12	-8.302	是

通过表 3 得出，传统授课教学模式与三种现场展示教学模式之间存在显著性差异。具体来说，现场展示教学模式在统计上显著优于传统授课教学模式。我们发现所有 p 值都非常小，远小于设定的 0.05。这意味着在所有三种比较中，传统授课教学模式与现场展示教学模式之间存在显著性差异。T 值分别为 -10.427、-13.322 和 -8.302。这些值的绝对值都较大，表明传统授课教学模式与现场展示教学模式之间的差异非常显著。

此外，为了探究不同展示教学模式之间的差异，本文建立以下假设。

H_0 ：现场展示教学模式 1, 2, 3 之间没有显著性差异

H_2 ：现场展示教学模式 1, 2, 3 之间有显著性差异

第二次配对样本 T 检验比较了现场展示教学模式 1、2、3 之间的差异，检验结果见表 4。在现场展示教学模式 1 与现场展示教学模式 2 的比较中，p 值为 1.07e-05 这个值远小于 0.05 的常用显著性水平，表明可以拒绝原假设，即认为两种教学模式之间存在统计学上的显著差异。T 值为 -4.684，因此，得出结论，现场展示教学模式 1 在效果上与现场展示教学模式 2 有显著的不同。另一方面，在现场展示教学模式 2 与现场展示教学模式 3 的比较中，p 值为 8.38e-01，这个值高于 0.05 的显著性水平，这意味着我们没有足够的证据拒绝原假设，即认为两种教学模式之间没有统计学上的显著差异。因此，我们可以得出结论，现场展示教学模式 2 与现场展示教学模式 3 在效果上没有显著的不同。

表4：第二次 T 检验结果表

模式比较	p 值	T 值	显著性
现场展示教学模式1vs 现场展示教学模式2	1.07e-05	-4.684	是
现场展示教学模式2vs 现场展示教学模式3	8.38e-01	-0.205	否

综合两次配对样本 T 检验分析的结果，得出结论：现场展示答辩和提问环节显著提升了学生的实验成绩，这表明实践教学方法的创新对于提高学生的实际操作能力和理解深度具有重要作用。这种教学方法可能通过增强学生的参与度、促进知识的内化以及提供即时反馈来提高学习效果。然而，理论考试复习对于现场展示答辩成绩的影响不显著，这可能意味着理论与实践之间存在一定的脱节。这种现象可能源于学生在理论知识学习过程中缺乏足够的实践应用机会，或者理论与实践教学之间缺乏有效的衔接机制。

（二）探讨不同展示方法的相关性分析和回归分析

在本次教学中，我们首先设计了4个教学模式，每个模式都针对不同的教学目标。这些模式指导了考试安排和成绩评估。模式1是为了评估学生的写作技能，安排了第1次实验来测试并得出成绩。模式2旨在评估学生的理解能力，安排了第二次实验来得出相应的成绩。模式3关注学生的临场发挥能力，通过第3次和第4次实验的成绩来评估这一能力，并计算平均值。最后，模式4探讨理论学习如何转化为实验技能，通过第5次至第8次实验的成绩来评估，并计算这4次成绩的平均值。这样，每个模式都通过特定的实验成绩来实现其教学和评估目标。8次实验成绩分布偏左且比正态分布更尖锐，因此不适合使用要求数据正态分布的皮尔逊相关系数。相反，斯皮尔曼相关系数是一种非参数方法，它基于数据的等级而非实际值，适合非正态分布数据。在本研究中，使用斯皮尔曼相关系数能更准确地评估变量间的相关性，因为它不受数据分布形态的影响。图3绘制了其斯皮尔曼相关系数热力图。

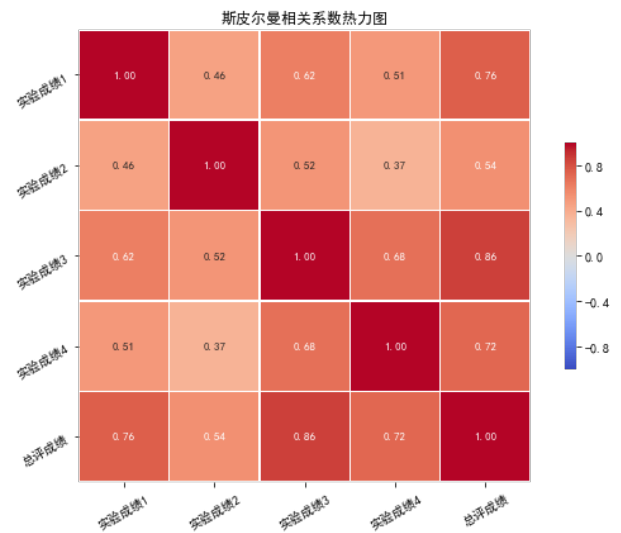


图3：斯皮尔曼相关系数热力图

从图3中得出，所有教学模式与总评成绩之间的相关系数均接近或超过0.70，表明这些教学模式与总评成绩之间存在高度的正

相关性。这一结果强调了各教学模式在综合评价中的重要性，也说明了总评成绩能够较好地反映各教学模式的整体表现。特别是现场展示教学模式2与总评成绩的相关系数高达0.87，这表明在实验考试中的表现与总评成绩密切相关。传统授课教学模式与总评成绩的相关系数为0.76，显示出传统授课模式与总评成绩的相关性较低，可能在提高学生总评成绩方面效果有限。现场展示教学模式在提升学生总评成绩方面表现更为有效，尤其是现场展示教学模式2和模式3。

为了探究不同教学方法对学生实验成绩的具体影响，本实验使用总评成绩 (y) 作为因变量，而将4次模式 (x_1, x_2, x_3, x_4) 作为自变量进行回归分析。

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4$$

考试1记录模式1，它代表学生的表达情况。考试2记录模式2，它代表学生的理解情况。考试3或4记录模式3，它代表学生的临场发挥情况。考试5或6或7或8记录模式4，它代表理论考试对实验考试的影响。本文选取变量进行回归分析得出表5。

表5: 回归分析结果表						
模型 B		未标准化系数		标准化系数	t	显著性
		标准错误	Beta			
	(常量)	-1.75	2.58		-0.68	0.50
	实验成绩1	0.14	0.03	0.12	4.99	0.00
	实验成绩2	0.27	0.04	0.22	7.60	0.00
	实验成绩4	0.19	0.03	0.23	6.08	0.00
	实验成绩5	0.43	0.03	0.59	13.98	0.00

通过表5发现实验成绩对学生的总评成绩有显著的正向影响。尽管模型中的常量项不显著，表明在没有任何实验成绩信息的情况下，无法准确预测总评成绩，但实验成绩1、实验成绩2、实验成绩4和实验成绩5的系数都显著且为正，说明这些实验成绩都对总评成绩有积极的贡献。其中，实验成绩5的影响最为显著，Beta系数高达0.59，表明它在所有实验成绩中对总评成绩的提升贡献最大。实验成绩1、实验成绩2、实验成绩4虽然影响相对较小，但它们的Beta系数分别为0.12、0.22和0.23，也都显示出显著的正向效应。为此，将四种模式下不同实验成绩进行组合，表6展示了8组分组结果。

表6: 分组结果表

灵敏度 分析组 别	实验1	实验2	实验3	实验4	实验5	实验6	实验7	实验8
组1	✓	✓	✓		✓			
组2	✓	✓	✓			✓		
组3	✓	✓	✓				✓	
组4	✓	✓	✓					✓
组5	✓	✓		✓	✓			
组6	✓	✓		✓		✓		
组7	✓	✓		✓			✓	
组8	✓	✓		✓				✓

针对以上8组分组情况，本文进行灵敏度分析。图4绘制了不同组合的平均方差和确定系数图。从图4中可以看出不同考试组合的MSE变化不大，说明模型对不同自变量组合的敏感度较低，模

型较为稳定。且不同考试组合的 R-squared 值普遍较高，说明模型整体具有良好的解释能力。

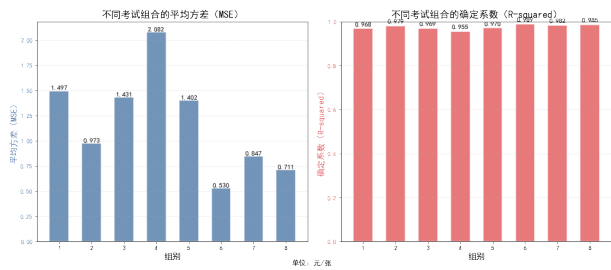


图 4：不同考试组合的均方根误差和确定系数

三、结语

本研究旨在探讨现场展示答辩教学方法对学生在统计案例实验中学业成绩的影响，并通过实证数据分析了该教学方法的有效性。以下是根据研究结果得出的详细结论：

1. 本研究结果显示，与传统教学方法相比，现场展示答辩教

学方法显著提高了学生的实验成绩。且这一差异在统计上是显著的。这表明现场展示答辩通过增加学生的参与度和即时反馈，促进了学生对实验内容的深入理解。

2. 提问环节被认为可以增强学生的批判性思维，本研究发现该环节对提升实验成绩的影响显著。这说明提问代表学生能够迅速适应、灵活应对和创造性解决问题，这些能力对于取得更好的实验结果具有重要作用。

3. 针对理论与实践成绩的关系，本研究进一步分析了理论考试复习对实践成绩的影响，结果显示，实验成绩3和4的平均值与实验成绩5、6、7、8之间的比较并未发现显著差异，P 值均高于0.05。这表明理论知识的掌握和实践技能的提升可能是由不同的学习机制所驱动，且理论考试复习并不直接转化为实践环节分数的提高。

综上所述，本研究不仅验证了现场展示答辩教学方法在提高学生实践成绩方面的有效性，同时也指出了理论与实践教学之间可能存在的独立性。这些发现对于教育工作者在教学实践中做出更有效的教学决策具有重要的指导意义。

参考文献

[1] 马文珠, 雕明河, 张莉, 等. “模式教学法”对学生针灸治疗学成绩影响的比较分析 [J]. 中医教育, 2009, 28(03):16.

[2] 张喜娟, 丁钊鹏. 因子分析在学生成绩影响因素调查结果中的应用 [J]. 北京联合大学学报 (自然科学版), 2010, 24(01):78.

[3] 肖艳群. 决策树在高职学生成绩影响因素分析中的应用 [J]. 宁波职业技术学院学报, 2012, 16(05):57.

[4] 任兆林. 不同学期课程设置对学生成绩影响的统计分析 [J]. 劳动保障世界 (理论版), 2013, (12):32.

[5] 李桂荣, 李向辉. 中职学生学业成绩影响因素分析——基于河南省的经验研究 [J]. 教育经济评论, 2016, 1(02):71.

[6] 金昆萌. 基于贝叶斯网络的继续教育学生成绩影响分析 [J]. 内江科技, 2017, 38(05):143.

[7] 王帅, 莫惠玲. 新工科背景下软件工程课程案例教学改革实践 [J]. 高教学刊, 2025, 11(04):129-133.

[8] Zhang X, Obaidat M S, Lee V C S, et al. An Effective Learning Management System for Revealing Student Performance Attributes[C]//2024 International Conference on Computer, Information and Telecommunication Systems (CITS). IEEE, 2024:3.

[9] Chango, Wilson et al. “Multi-source and multimodal data fusion for predicting academic performance in blended learning university courses.” Comput. Electr. Eng. 89 (2021): 5.

[10] Zhang, Mengmeng and Xiantong Yang. “Google or ChatGPT: Who is the Better Helper for University Students.” ArXiv abs/2405.00341 (2024): 8.

[11] 陈超辉, 田罗庚, 陈雄, 等. 基于柯氏模型评估启发式教学模式在大气物理学中的应用效果 [J]. 高教学刊, 2025, 11(16):120-123.

关于构建促进高效能社会治理体系的 政务数据目录的研究

谷剑芳

河南省政务大数据中心, 河南 郑州 450000

DOI:10.61369/ASDS.2025070021

摘要 : 随着信息技术的快速发展和数字化转型的深入推进, 政务数据治理在政府决策、管理和服务中的重要性日益凸显。政务数据作为国家治理体系的重要基础资源, 其数据质量和价值直接关系到政府效能的提升和社会经济的发展。构建高效能社会治理体系, 亟须建立与之相适应的全量编目、同步更新、全网互认的政务数据目录体系。通过编制数据目录, 全面梳理政务数据, 实现统一标准、规范管理、动态更新和安全共享, 确保数据资源可见、可管、可用、可信, 为跨层级、跨地域、跨系统、跨部门、跨业务的协同治理提供坚实的数据基础。本文围绕政务数据目录的构建与实践, 探讨其在促进高效能社会治理中的关键作用与实施路径。

关键词 : 政务数据; 目录治理; 目录要素; 社会治理

Research on the Construction of Government Data Directory for Building a High-Performance Social Governance System

Gu Jianfang

Henan Provincial Government Big Data Center, Zhengzhou, Henan 450000

Abstract : With the rapid advancement of information technology and the deepening digital transformation, the significance of government data governance in governmental decision-making, management, and service delivery has become increasingly prominent. As a foundational resource for the national governance system, the quality and value of government data directly impact the enhancement of government efficacy and the development of the socio-economy. To build a high-performance social governance system, it is imperative to establish a corresponding government data directory system that features comprehensive cataloging, synchronized updates, and universal recognition across networks. Through the development of data catalogs, government data can be systematically inventoried, enabling unified standards, standardized management, dynamic updates, and secure sharing. This ensures data resources are visible, manageable, usable, and trustworthy, thereby laying a solid data foundation for collaborative governance that spans multiple levels, regions, systems, departments, and domains. This paper focuses on the construction and practice of government data catalogs, exploring their pivotal roles and implementation pathways in facilitating high-performance social governance.

Keywords : government data; catalog governance; catalog elements; social governance

引言

随着信息技术的快速发展和数字化转型的深入推进, 政务数据治理的重要性日益凸显。政务数据在政府决策、管理和服务中发挥着重要作用, 其数据质量和价值直接关系到政府效能的提升和社会经济的发展。党的二十届三中全会的决议明确指出“建设和运营国家数据基础设施, 促进数据共享”。数据目录是数字政府建设的重要基础性工作, 对促进政务数据共享开放、系统互通、协同治理、质量提升等有着重要意义。国务院宣布2025年8月1日《政务数据共享条例》正式生效, 政务数据工作在法规层面获得有力支撑, 建立起对政府部门数据责任的刚性约束。政务数据管理部门在加强目录治理、促进数据共享方面也将获得政策背书, 并对目录治理、数据共享工作中不作为、慢作为的部门实施有力监督问责。政务数据目录治理不仅是一项技术工程, 更是推进国家治理体系和治理能力现代化的重要抓手。

一、高效能社会治理体系所需的政务数据管理体系

（一）基本概念

政务数据目录，是指通过对政务数据依据规范的要素描述，按照一定分类方法进行排序和编码的一组信息，用于描述政务数据的特征，以便对政务数据进行检索、定位与获取。政务数据目录由目录要素构成，目录要素用于描述数据特征，包括目录编码、目录名称、信息项、业务分类、数据类型、数据层级、数据敏感级别、共享类型、使用要求、来源系统、来源系统编码、系统所属分类、数据区域范围、数据时间范围、政务数据提供方、更新周期等几十项信息。《政务数据共享条例》从政务数据实行目录管理的基本原则出发，明确了政务数据目录要素构成，确定了“有条件共享”“无条件共享”“不予共享”的三种共享属性，明确政务数据目录实行全国统一管理，强化政府部门履行数据共享的职责和数据目录编制、维护、更新、发布等管理流程。

（二）主要目标

一是全量编制。依据部门职责和权责清单，实现政务数据资源的全面梳理和编目，覆盖所有政务数据资源，形成完整的“数据－目录”关联关系。二是全网互认。建立全国统一的目录标准和互认机制，统一目录要素标准和编码规则，构建全国政务数据目录“一本账”，推动政务数据在跨层级、跨地区、跨业务、跨部门、跨系统流动过程。三是同步更新。建立目录与数据的动态关联机制，确保目录随业务变化实时更新。

（三）政务数据目录治理对高效能社会治理的意义

1. 构建国家数据基础设施，形成全国政务数据“一本账”管理框架。破除数据流通壁垒，建立目录与数据的“同步更新”机制，促进跨区域数据共享。
2. 优化公共服务，实现“数据多跑路，群众少跑腿”。通过编制全量目录，促进数据供给，支撑精准施策。
3. 促进协同治理，打破部门壁垒实现跨域协作。强化风险防控，通过数据关联分析提升预警能力。

（四）构建高效的政务数据目录的目标

1. 全量覆盖与标准统一：数据目录应全面涵盖各级政府部门在履职过程中产生、采集和管理的所有政务数据资源，做到“应编尽编”。遵循统一的元数据标准规范，确保不同来源、不同层级的数据目录能够以一致、规范的方式被理解。
2. 动态更新与权威发布：数据目录是一个动态更新的体系，要建立完善的目录维护和更新机制，确保目录内容能及时反映政务数据资源的实际变化。经过审核确认的权威目录版本，应在统一平台上集中发布，作为数据资源查询和使用的唯一可信来源，避免“数出多门”。
3. 全网互认与高效检索：坚持统一标准，实现全国范围内数据目录的互认互通。目录管理系统提供强大的、基于元数据的检索功能，支持用户快速、精准地查找到所需的数据资源及其提供方。
4. 清晰界定权责与安全可控：目录要明确标注每项数据的提供部门及其联系方式，准确标识数据的敏感级别、安全分类以及具体的共享条件和访问控制要求，确保数据在可见、可管的基础

上实现安全、可信。

二、国家与省级政务数据目录管理的政策演进

（一）国家层面政策体系

2022年6月，国家提出“建立全国标准统一、动态管理的政务数据目录，实行‘一数一源一标准’，实现数据资源清单化管理”^[1]，要求“编制公共数据开放目录及相关责任清单，构建统一规范、互联互通、安全可控的国家公共数据开放平台，分类分级开放公共数据，有序推动公共数据资源开发利用”。

2022年10月，国家提出“数据目录一体化”^[2]工作任务，按照应编尽编的原则，推动各地区各部门全量编制、全面汇聚、动态更新、同源发布，形成全省政务数据目录一体化治理格局，实现全国政务数据“一本账”管理。国家以试点形式在内蒙古、浙江、安徽、山东、广东、海南、重庆、工信部、生态环境部、农业农村部开展一体化数据目录编制，探索新的政务数据目录要素标准、编目流程和治理规则^[3]。

2024年4月，国家全面开展数据直达基层工作^[4]，明确了各省数据目录向国家平台汇聚、级联的目录要素标准和规范。同年12月，国家全面启动政务目录治理工作，以统一的数据目录标准和认领、编制流程，明确“全量覆盖、全网互认、同步更新”的目标。

2025年5月，国务院颁布的《政务数据共享条例》以“目录管理”的专门章节，进一步强化政务数据实行目录管理。

（二）河南省政务数据管理政策

2023年河南省开展数字政府建设^[5]，提出建设标准统一、及时更新、管理高效的数据目录，以目录为抓手推进数据治理、共享、开放，要求制定标准统一、全量覆盖、动态管理的政务数据和公共数据目录，实行数据分类分级和“一数一源一标准”、健全省市一体化数据共享交换体系、整合各部门数据交换通道，实现数据目录统一管理、编制公共数据开放目录、责任清单。明确了“梳理各级政务部门履职过程中产生、采集、管理的数据资源和结果数据资源，统一目录标准、编目流程，编制全量覆盖、互联互通的政务数据目录，建立目录同步更新管理机制，构建分级管理的全省一体化数据目录体系”和“统一政务数据开放政策和操作规程，分级发布数据开放目录，按年度制定政务数据开放重点清单”。

三、河南省政务数据目录治理实践

从2017年至今，河南省政务数据目录治理历经整合共享、专项编目、数据普查三阶段，持续应对标准化与动态化挑战。2025年出台《河南省政务数据目录治理技术规范（试行）》，实现要素标准、编制流程、省市平台对接“三统一”，为全域目录治理提供标准化施工图。

（一）主要成效

1. 目录实用性提升。结合河南省政务数据目录标准，共明确

61项目录要素。其中,16项要素为系统自动生成,33项要素为选择项,12项是录入项,录入项占20%,大大降低了业务人员编目的复杂性。

2. 管理效率提升。一是建立与省直部门数据资源库的数据采集链路,自动采集、更新各业务系统数据库表信息,为业务人员编制目录提供数据来源,建立“系统一数据一目录”的关联关系,为动态更新目录提供技术支持。二是市级数据目录通过省市政务数据平台级联,以API接口形式实时向省一体化政务大数据平台上传目录信息并在门户展示,形成一体化联动的全省数据资源“一本账”。

3. 目录质量提升。一是持续推进目录治理。包括规范了政务数据编目、资源挂载的流程;督促各级政务部门及时维护、更新、补全数据目录。二是编制《政务数据分类分级指南》《政务数据目录编制规范》《政务数据交换技术规范》等工程标准,在数据目录的实践中逐步丰富、完善标准体系。

(二) 存在不足

1. 统一的标准体系尚未完全建立。一是存量的目录待规范。各地各部门历史目录的格式、内容和质量参差不齐,难以实现跨部门和跨系统的数据共享和交换。需要随着数据供出的增加,持续推动存量目录的标准化。二是工作机制待完善。省、市、县目录要素内容缺乏规范衔接,没有形成上下联动、规范编制的工作局面,没有实现省内目录体系同标准、全要素、同质量管理。

2. 数据定级存在挑战。不同部门在开展数据定级时,对核心、重要和一般三个级别的国家数据分级标准^[6]的理解有偏差,或是受到特定项目需求的影响,导致数据级别的划分出现随意性,产生了定级过高、过低和冲突等问题。比如,已在互联网公开的信息,定义为3级;个人身份证号和手机号属于个人敏感信息,却被定为2级;同一个系统,身份证号在不同的目录,有的是2级、有的是4级。不合理定级将造成数据在管理、运维、共享、开放中,数据处置方法的混乱,增加管理使用成本。

3. 目录治理职责尚未落实。不同的政府部门,对于数据目录的理解存在显著的差异。一些部门认为,数据目录的编制应该是在共享的基础上进行的,即在数据共享之后再编制目录。一些部门将数据编目视为一项纯粹的技术性工作,而忽视了数据编目中需要业务管理和决策支持。一些部门缺乏对数据目录进行常态化管理的意识,这导致了数据目录的通报不及时、更新不频繁,缺乏提醒机制,全要素编制工作得不到充分执行,全量编目的难度因此增加。

四、推进政务数据目录治理的路径探讨

(一) 强化数据工作机制

《政务数据共享条例》第二十条“政务数据共享主管部门应当建立政务数据共享供需对接机制,明确工作流程。”第二十八条“政务数据共享主管部门应当建立健全政务数据共享争议解决处理机制。”省级层面应当落实条例中关于建立有关工作机制、细化本地业务流程的要求,制定本地区数据条例或政务数据管理

细则,从实际操作层面明确、强化数据目录管理的细节,明确数据提供途径和周期、数据质量、数据异议处理、供需对接响应等刚性约束,压实政务部门数据共享责任,让条例与现实情况衔接起来,真正实现“共享为原则,不共享为例外”。

(二) 规范开展目录治理

首先,提升整体规划,制定全省统一的数据资源系统构建方案,确立省市数据收集、处理的技术角色,从全省数据共享交换、目录管理、数据资源、数据服务等多个业务,强化业务的一体化协作、联动。其次,编制数据目录治理的技术规范,包括数据开放、共享、分类、分级等要素标准,明确“自上而下”的工作模式,即由省级行业单位统一编制数据目录,市县负责认领目录要素,分级编制扩展目录要素。最后,要加强数据资产全生命周期的管理^[7],从确保国有资产保值增值的层面,将数据目录管理系统与资产管理系统相衔接,为实施数据资产登记管理打下坚实基础。

(三) 科学数据分类分级

应按照分类分级保护的原则进行数据防护管理,分级标准应当根据行业特点和数据特点去把握。对于核心数据,应当采取更加严格的安全保护措施,采取加密存储、加密传输、访问控制等技术手段。对于重要数据,应当建立专门的的安全管理制度,定期评估和更新数据的分级情况。数据分类分级应当动态管理,定期评估和更新数据的分类分级标准,确保分级体系与实际数据特点和安全需求保持一致。核心数据应当至少采取加密存储、加密传输、访问控制安全保护措施。数据分级体系应当定期评估和更新,确保符合法律法规的要求和实际业务需求。

(四) 加强数据安全防护

在推进数据资源体系建设的同时,必须重视数据安全和隐私保护。为此,应建立严格的数据安全管理制度,确保数据在采集、存储、传输、处理和使用过程中的安全。同时,加强数据安全技术的研发和应用,采用先进的加密技术、访问控制和审计机制,防止数据泄露和滥用。此外,还需定期对数据安全进行风险评估和漏洞检测,及时发现并解决安全隐患。为了确保数据资源体系建设的顺利进行,需要建立完善的政策法规体系,明确数据资源的权属、管理、使用和保护等方面的规定,为数据资源的合理利用和有效管理提供法律保障。

(五) 加快数据共享开放

应建立开放共享机制,鼓励跨层级、跨地区、跨部门、跨系统、跨业务的数据资源流动,制定激励政策,鼓励更多部门愿意主动共享、开放数据。发挥政务数据平台的作用,进一步强化国、省、市政务数据平台级联,提高全省一体化共享交换运行效率,促进数据直达基层,保障“高效办成一件事”。当前,国家数据正在开展公共数据资源登记工作,有些省份将数据共享开放属性与数据资源登记、编目相结合,在全面登记、编制,形成数据资源“一本账”的基础上,摸清数据底数、摸清可共享、可开放的底数,促进数据开放共享。为保障数据开放共享,要建立监督评估机制,定期对数据共享、开放、运营的效果进行评估,及时发现问题并进行整改,确保工作的持续推进。

五、总结

数据目录是数据治理的基础，通过编目、校核、共享等手段，逐步梳理出具备潜在价值的数据资源，推进资产化管理，挖掘数据价值，促进数实融合。河南省政务数据目录治理在规范制定、数据质量、资源共享等方面取得初步成效，但在制度建设、常态化治理和科学定级等方面仍有不足。未来需从工作机制、目

录治理、共享开放、分类分级管理等方面，系统性提升政务数据目录治理能力。需从完善本地政务数据管理制度和保障机制促进部门落实目录和数据治理责任，以社会治理需求驱动目录治理，运用区块链、AI 等技术提升目录管理智能化水平等方面，持续深化政务数据目录实践。只有构建起科学规范、高效运行的目录体系，才能充分释放政务数据价值，为高效能社会治理提供坚实的数据基础。

参考文献

-
- [1] 国务院.《关于加强数字政府建设的指导意见》[Z]. 2022.
 - [2] 国务院办公厅.《全国一体化政务大数据体系建设指南》[Z]. 2022.
 - [3] 国务院办公厅.《全国一体化政务大数据体系政务数据目录编制和治理规范》[Z]. 2022.
 - [4] 国务院办公厅.《全国一体化政务数据共享数据直达工作规范》[Z]. 2024.
 - [5] 河南省政府.《数字政府建设实施方案（2023-2025 年）》[Z]. 2023.
 - [6] 国家标准化管理委员会.《GB/T 43697-2024 数据安全技术 数据分类分级规则》[S]. 2024.
 - [7] 财政部.《关于加强行政事业单位数据资产管理的通知》（财资〔2024〕1 号）[Z]. 2024.

