

大数据背景下应用统计专业硕士案例分析课程的教学模式构建与实践探索

褚彦彤, 姜博丹, 姜永琦, 郭云飞*
延边大学 理学院, 吉林 延吉 133000
DOI:10.61369/ASDS.2026030005

摘要 : 针对大数据背景下应用统计专业硕士案例分析课程中存在的教学问题, 研究其课程的独特性, 提出教学改革策略。通过“基于共形分位数回归的空气质量预测”这一真实案例的教学实践, 展示了从数据处理、模型构建到结果转化的完整教学过程, 并基于多元评价结果验证了该教学模式在提升学生建模能力、解释能力与决策支持能力方面的显著成效。教学实践表明, 该模式不仅能有效解决传统教学中案例真实性不足、重工具轻思维、结果应用弱、考核单一等问题, 而且为应用统计专业课程建设提供可复制、可推广的范式, 对推动教学方法创新与人才培养质量提升具有重要实践意义。

关键词 : 应用统计; 案例分析; 教学改革; 分位数回归; 共形预测

Construction and Practical Exploration of Teaching Mode for The Case Analysis Course of Master of Applied Statistics under The Background of Big Data

Zhuo Yantong, Jiang Bodan, Jiang Yongqi, Guo Yunfei*
College of Science, Yanbian University, Yanji, Jilin 133000

Abstract : The core objective of the Applied Statistics major is to cultivate the ability of students to solve practical problems. As a key carrier connecting statistical theory with practical applications, the case analysis course plays a significant role in teaching optimization. This paper first analyzes the uniqueness of the case analysis course in the Applied Statistics professional degree, and points out the existing pain points in current teaching. In response to these issues, it proposes teaching reform strategies. Through the teaching practice of the real case “Air Quality Prediction Based on Conformal Quantile Regression”, it demonstrates the complete teaching process from data processing and model construction to result transformation. In addition, based on the results of multi-dimensional evaluation, it verifies the significant effectiveness of this model in enhancing modeling ability, interpretation ability and decision support ability. Finally, the research points out that this case-driven teaching model provides a replicable and scalable paradigm for the construction of Applied Statistics courses. It has important practical significance for promoting teaching method innovation and improving the quality of talent cultivation.

Keywords : applied Statistics; case study; teaching reform; quantile regression; conformal prediction

引言

大数据时代的到来使得数据规模与复杂度呈现爆炸式增长, 与此同时, 统计学的应用也迎来了相应的机遇与挑战。在此背景下, 应用统计专业研究生的教育被赋予了培养高层次、具备数据驱动思维和实践能力的应用型人才的使命。近些年应用统计专业硕士的人才培养问题成为研究热点。张波^[1]、张彩虹等对应用统计专业硕士人才培养模式进行了研究, 完善了应用统计专业硕士的培养方案、教学计划等主要问题^[1,2]。这类人才不仅需要扎实的统计学理论功底, 更需要能够在复杂多变的实际场景中灵活运用模型与算法解决问题。因此, 教学模式的优化与课程体系的完善显得尤为关键。优化统计专业的教学体系, 已成为学界的核心议题之一。例如, 大数据

基金项目: 吉林省教育科学“十四五”规划项目(GH21036)。

作者简介:

褚彦彤, 第一作者, 延边大学, 硕士研究生, 研究方向为大数据与机器学习;

姜博丹, 第二作者, 延边大学, 硕士研究生, 研究方向为大数据与机器学习;

姜永琦, 第三作者, 延边大学, 硕士研究生, 研究方向为多目标整数优化。

通讯作者: 郭云飞, 延边大学, 副教授, 博士, 研究方向为非参数统计、模型平均、模型选择。

背景下, 统计学教学改革的核心在于推动理论教学与实践应用的深度融合^[5]; 为深化课程改革, 统计学专业充分挖掘学科竞赛的育人价值, 使其有效赋能于专业实践教学^[4]; 依托实践项目激活统计学课程, 是实现高素质复合型人才培养目标的有效途径^[6]。在此背景下, 案例分析课程作为连接统计学理论与现实应用的桥梁, 其战略地位日益凸显。近年来, 国内外教育工作者对应用统计案例分析课程进行了广泛的理论探讨与实践创新^[6]。案例分析课程在应用统计专业研究生培养中具有承上启下的作用。一方面, 它是学生将所学统计方法与工具运用于真实问题的重要桥梁; 另一方面, 它又是学生培养数据驱动思维、提升实践能力和科研素养的核心环节。与传统课堂以理论讲授为主不同, 案例分析课程强调数据获取、建模、分析与解释的完整过程, 更加契合大数据背景下人才培养的现实需求。在国际上, 案例驱动与项目导向型教学方法已在统计教育和数据科学教育中广泛应用。许多高校在硕士阶段引入真实企业案例或科研项目, 突出跨学科整合能力和创新性问题解决能力的培养。与之相比, 国内统计类课程也逐步引入成果导向教育(OBE)^[7]理念和BOPPPS教学模式, 并在本科层面积累了一定经验。然而, 在研究生教育阶段, 尤其是应用统计专业学位的案例课程建设中, 仍存在体系化不足、实践环节不够深入等问题。如何增强案例课程的针对性与系统性, 使其更好地服务于应用型人才培养, 已成为亟待解决的重要议题。

因此, 本文立足于应用统计专业学位案例分析课程的独特性, 结合当前教学实践中存在的痛点, 探讨在大数据背景下研究生教学模式优化的路径。通过对案例课程特征与不足的剖析, 可以为课程设计与教学改革提供理论依据和实践启示。

一、应用统计专业学位案例分析课程的独特性与教学痛点

(一) 课程独特性

与工商管理、法学等领域的案例课程不同, 应用统计专业的案例分析课程具有鲜明的“数据中心”和“模型驱动”特征, 其独特性主要体现在以下几个方面:

1. 数据中心性

应用统计案例分析通常涉及大规模、多维度的数据, 要求学生在复杂的数据环境中开展清洗、处理与挖掘。这种“以数据为核心”的特征, 不仅体现了学科本质, 也对学生的实践能力和数据敏感性提出了更高要求。

2. 模型驱动性

案例分析并非停留在现象描述, 而是强调基于统计模型与算法对问题进行系统建模与深入解释。模型的建立与验证是案例讨论的核心环节, 使学生在方法选择与逻辑推理中得到全面训练。

3. 跨学科性与应用导向

应用统计的案例往往源自金融、医疗、环境、互联网等多领域, 呈现高度跨学科特征。这类案例促使学生将统计方法与领域知识相结合, 形成面向实际问题的综合性解决方案。

4. 实践导向与能力培养

案例课程不仅关注学生对方法的掌握, 更强调从数据获取、建模分析到结果解释与呈现的全流程训练。这一过程有效提升了学生的数据思维、沟通能力与科研素养, 契合了大数据背景下应用型人才培养的需求。

这些特征不仅彰显了课程区别于其他学科案例教学的鲜明定位, 也契合了大数据时代对应用型统计人才的培养要求。然而, 正因其复杂性与高要求, 该类课程在具体实施过程中仍面临诸多挑战, 亟需在教学内容、方法和评价体系等方面进行进一步探索与优化。

(二) 教学痛点

尽管案例课程在培养研究生实践能力方面发挥着重要作用,

但在实际教学中仍存在一些突出问题。

1. 案例素材缺乏真实性与复杂性

许多课程采用的案例数据集经过大量清理与简化(如鸢尾花数据集、波士顿房价数据集), 虽然便于课堂操作, 却忽略了数据在实际工作场景中的混乱性与不完整性。学生缺少与真实数据打交道的经验, 容易在今后的职业实践中出现适应性不足。

2. 教学内容偏重工具操作

当前教学中, 学生往往被要求熟练掌握R、Python等工具, 但对模型选择的理论逻辑、假设条件及统计推断的本质理解不足。这种偏重“工具技能”的倾向, 导致学生在统计思维和问题抽象方面的能力培养受到削弱。

3. 模型结果的解释与应用不足

在建模过程中, 学生往往热衷于使用复杂方法, 但缺乏对结果进行深入解读和转化的能力, 难以提出切实可行的对策。这种“重方法、轻解释”的现象削弱了案例教学与实际决策之间的联系。

4. 考核方式单一, 评价维度不足

目前案例课程的考核多集中于书面作业, 缺乏对学生在数据处理、建模探索、团队协作和成果展示等方面能力的综合评价。这种单一的评价方式难以全面反映学生的学习成效。

综上所述, 应用统计案例分析课程既具有鲜明的学科特征, 又在教学实践中暴露出素材真实性不足、思维训练缺位、结果应用乏力等问题。这些痛点制约了课程在人才培养中的作用, 也为后续教学模式的优化与改革指明了方向。

二、教学改革策略

(一) 提升案例素材的真实性与复杂性

增加来自金融、医疗、环境等领域的真实数据, 避免过度清洗与简化。在教学中可引入开放数据平台和行业合作案例, 或引导学生自主在公开数据集中寻找未经清理的原始数据, 再根据数据特征选择合适的应用场景开展分析与建模操作, 这不仅能够强

化学生的动手能力与问题敏感度，还能够帮助他们在未来科研与职业实践中更好地应对复杂多变的真实数据环境。

（二）平衡工具操作与统计思维训练

在课程设计中加强统计建模原理、假设设定及推断逻辑的讲解，使工具的使用服务于统计思维的培养。通过结合经典案例与跨学科实例，帮助学生理解方法背后的逻辑与价值，避免“唯工具化”的倾向。

（三）强化模型结果的解释与应用

在教学环节中增加对结果解释与政策应用的训练。要求学生复现经典案例建模并撰写报告，阐明实际意义与可行性，从而推动“重方法”向“重方法与重应用并重”转变。

（四）多元化考核方式与综合评价

针对考核方式单一的问题，应建立“过程性+结果性”相结合的多维度评价体系。在传统书面作业的基础上，增加数据获取与处理、团队合作建模、成果展示与交流等环节的考核，以全面反映学生在知识掌握、技能应用与综合素质方面的学习成效。

教学改革的核心在于以学生发展为本，推动课程、方法、评价、师资与技术的协同创新。通过上述策略，可以有效破解现有教学中的痛点与难题，促进学生语言能力与综合素质的全面提升。这不仅回应了新时代教育发展的需求，也为学科建设与人才培养提供了切实路径。

三、教学实践示例

将“基于共形分位数回归的空气质量预测”案例融入应用统计专业硕士的课程教学中。

（一）案例背景与教学目标设计

本案例以北京市2018-2025年空气质量监测数据为基础，引导学生直面真实环境数据中的典型问题如季节性波动、极端值、缺失值，旨在培养以下能力。

1. 高级统计方法应用能力：掌握分位数回归、共形预测等方法，解决非正态分布数据的建模挑战。

2. 行业场景的理解与问题转化能力：应用统计专业案例分析课程的核心价值在于服务行业，案例分析课程通过引入不同领域（如金融、医疗、环保、互联网）的真实案例，打破纯技术视角，培养学生的业务思维。

3. 基于统计结论的决策支撑与落地建议能力：引导学生将模型分析结果深度拆解、落地转化为具体可落地的环保政策建议。最终输出成果需以实用价值为核心，确保转化后的建议具备针对性、可操作性，能够直接为环保决策制定提供明确参考与行动支撑。

（二）教学内容与过程

1. 数据处理

从专业数据平台获取北京市空气质量相关真实数据，在AQI相关数据研究中，因监测系统不稳定、外界环境突变或人为操作失误等，原始数据易出现空值，给后续建模分析造成阻碍。三次样条插值凭借其适配环境数据连续性与局部波动性的特性，成为

高效的数据预处理手段，为模型搭建筑牢根基。三次样条插值核心思路为先将完整连续监测区间，按已有数据点划分为若干子区间，再在每个子区间内，用三次多项式函数构建插值曲线。相较于简单均值补缺、线性插值，三次样条插值优势显著，能精准拟合AQI局部非线性波动，避免线性插值“强制线性化”误差；保障数据光滑性，使补缺后数据曲线贴合真实空气质量变化趋势，为分位数回归、时序预测等后续模型提供可靠输入；还可适配温度、湿度等多环境因子空值补缺，为多变量耦合建模奠定基础。三次样条插值以“分段三次拟合+全局光滑约束”，凭借其科学的公式定义与约束条件，成为环境统计预处理中连接原始监测与深度分析的关键技术，助力AQI预测及环境模型构建。

2. 分位数回归建模：从原理到实践

在分位数回归建模的教学中，首先应阐释其核心机制——是一种用于估计因变量Y在不同分位点上如何受自变量X影响的统计方法^[9]。它比传统的最小二乘回归更灵活，能够分析整个条件分布，适用于非对称数据、异方差数据和极端值分析，进而引导学生对比不同机器学习方法上的表现，此案例分析教学分别构建了基于随机森林、LightGBM与神经网络的分位数回归模型^[9]，并以点预测与区间预测相结合的双重评价机制对模型性能进行对比。通过可视化工具即不同模型的不同分位数预测曲线图直观展示各模型对空气质量相关数据波动范围的捕捉能力，结果显示，在北京市空气质量指数（AQI）预测中，基于0.6-0.75分位数的点预测表现更优。在此基础上，引导学生将此法举一反三，思考其在医疗数据中用于识别异常生理指标，或在金融领域中计算风险价值等广阔应用场景，以深化对方法通用性的理解。

3. 共形预测：提升决策可靠性

共形分位数回归（CQR）在教学中可作为连接统计理论与实际应用的典型范例，其核心在于先用QR拟合上下分位数，再通过校准数据集计算残差校正，该方法不依赖于数据分布形态的假设，却能在有限样本下提供严格的频率覆盖保证^[10]。针对时间序列中常见的分布漂移问题，本案例进一步引入JS散度加权机制，通过动态衡量历史数据与当前数据分布的差异，调整校准权重，使模型具备对季节、气象等外部条件变化的适应能力，这一设计可引导学生理解不同变量对模型泛化性的影响及应对策略。实证结果显示，经共形校准后，QRNN模型的覆盖率由0.928提升至0.954，而区间宽度仅增加0.007，生动体现了统计学习中所追求的有效性-精度平衡，具体模型对比结果如表1和表2所示。借此引导学生展开跨领域思考，例如，在医疗诊断中如何利用CQR构建具有覆盖保证的疾病风险区间，或在金融风险管理中如何通过分布自适应校准提升风险价值的预测稳健性，从而深化学生对不确定性量化方法通用性与可迁移性的理解。

表1：三种分位数回归的区间预测评价结果

	90%区间预测			95%区间预测		
	PICP	PINAW	IS	PICP	PINAW	IS
QRF	0.989	0.015	6.105	0.994	0.021	8.374
LGBM-QR	0.981	0.077	20.901	0.822	0.123	33.785
QRNN	0.919	0.021	8.349	0.928	0.023	9.169

表2: 基于共形分位数回归的空气质量区间预测评价结果

	90% 区间预测			95% 区间预测		
	PICP	PINAW	IS	PICP	PINAW	IS
CQR-RF	0.920	0.013	18.290	0.9809	0.031	19.294
CQR-LGBM	0.906	0.022	21.138	0.945	0.061	40.894
CQR-NN	0.901	0.017	12.760	0.954	0.030	18.731

四、教学效果评价

基于“基于共形分位数回归的空气质量分析与预测”真实案例的教学实践,通过对应用统计专业50名学生的多元评价显示,该教学模式取得了显著成效。采用过程性评价(30%)与结果性评价(70%)相结合的考核体系,学生在项目计划书、小组协作和平时作业等过程性环节表现优异(平均得分25.8/30),尤其在最终技术报告(40%)和成果答辩(30%)等结果性评价中展现出扎实的技术功底和良好的实践能力(平均得分60.6/70)。85%的学生能够严谨实现合理利用统计学方法并撰写规范的技术报告,

80%的学生在答辩中有效展示了数据分析结果并将其转化为切实可行的环境决策建议。结果表明,这种以真实数据驱动的案例教学模式显著提升了学生的统计建模能力、结果解读能力和跨学科应用能力,为应用统计人才培养提供了成功范例。

五、结论

本文基于真实的案例教学实践,深入探索了应用统计专业课程的教学创新路径。课程创新性地将分位数回归、共形预测等统计学方法与真实环境数据相结合,开发出“理论教学-模型构建-结果转化”的三阶段教学模式。通过多元评估体系,有效解决了传统教学中的问题。这一教学探索不仅为应用统计专业课程建设提供了新型的教研范式,更对推动整个应用统计专业领域教学方法的迭代创新、助力学科人才培养质量升级具有重要的启示价值与实践意义。未来研究需在课程的系统性设计、真实性平衡、师资队伍建设以及与人工智能技术的融合等方面进一步深化。

参考文献

- [1] 张波. 中美专业硕士学位研究生教育对比分析[J]. 潍坊工程职业学院学报, 2015, 28(5): 17-20.
- [2] 张彩虹, 徐仁洪, 李频. 完善我国高校应用统计专业硕士培养体系的建议——基于中美应用统计专业硕士培养体系的比较[J]. 中国林业教育, 2017, (3): 44-47.
- [3] 李春阳, 沈煜焜. 大数据技术在高校统计学教学中的应用思考[J]. 科技视界, 2023(7): 91-93.
- [4] 张萌. 以学科竞赛为载体的统计学课程实践教学改革[J]. 大学, 2022(23): 169-172.
- [5] 刘明月, 杨孝海, 徐立焕. 基于实践式项目教学的统计学课程教学改革[J]. 学园, 2022(2): 19-22.
- [6] 杨利军, 杨晓慧, 郑晨. 案例教学在应用统计专业研究生培养中的应用研究[J]. 大学教育, 2023, (22): 50-54.
- [7] 丁倩, 陈霞, 汪军. 基于OBE理念的“应用统计与优化设计”课程教学改革[J]. 纺织服装教育, 2021, 36(05): 470-473+486.
- [8] Koenker R, Bassett Jr G. Regression quantiles[J]. Econometrica: journal of the Econometric Society, 1978: 33-50.
- [9] 段昶旭. 基于分位数回归神经网络的学术论文影响力预测研究[D]. 哈尔滨工程大学, 2023.
- [10] 万瑾璟. 基于分位数回归与共形预测的动态不确定性量化融合框架[J]. 统计与管理, 2025, 40(06): 23-36.