

思政元素融入多元统计分析课程教学中的探索 ——以降维方法章节教学为例

朱笑雨*, 李 龙

苏州科技大学数学科学学院, 江苏 苏州

收稿日期: 2023年8月31日; 录用日期: 2023年10月3日; 发布日期: 2023年10月13日

摘 要

《多元统计分析》是统计学专业的一门专业必修课程, 在自然科学和社会科学领域都得到了越来越广泛的应用, 是一种重要且实用的数据分析方法。根据多元统计分析的课程特点, 文章将探讨如何在降维方法章节教学中融入思政元素, 以立德树人为根本目标, 把思想政治工作贯穿教育教学全过程, 实现全程育人、全方位育人, 在多元统计分析的课堂里积极提升学生的学科知识、品德修养、创新精神以及实践能力, 积极践行社会主义核心价值观。

关键词

课程思政, 多元分析, 统计专业, 降维方法

Exploration of Integrating Ideological and Political Elements into the Teaching of Multiple Statistical Analysis Course

—Taking the Chapter Teaching of Dimensionality Reduction Methods as an Example

Xiaoyu Zhu*, Long Li

School of Mathematical Sciences, Suzhou University of Science and Technology, Suzhou Jiangsu

Received: Aug. 31st, 2023; accepted: Oct. 3rd, 2023; published: Oct. 13th, 2023

*通讯作者。

Abstract

Multivariate Statistical Analysis is a compulsory course in the field of statistics, which has been increasingly applied in both natural and social sciences. It is an important and practical method for processing multivariate data. Based on the characteristics of the course of multivariate statistical analysis, this article integrates ideological and political elements into the teaching of the dimensionality reduction methods section, with the fundamental goal of cultivating morality and cultivating people. It integrates ideological and political work throughout the entire process of education and teaching, achieving full and all-round education. In the classroom of multivariate statistical analysis, students' abilities in subject knowledge, moral cultivation, innovative spirit, and practical ability are actively improved, and socialist core values are actively practiced.

Keywords

Ideological and Political Education, Multivariate Analysis, Statistics Major, Dimensionality Reduction Methods

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

统计学是关于收集、整理、分析以及解释数据的科学,其目的是探索数据内在的数量规律性,为相关决策提供依据和参考。统计学理论与方法已被广泛应用于自然科学、工程技术和人文社科领域。同时该学科也强调学生的动手操作能力、解决实际问题的能力[1]。因此,本专业志在培养德、智、体、美、劳全面发展,具有良好的职业道德和科学文化素养,扎实的统计思想,能够把具体行业领域的特点与所学统计学知识相结合,根据社会与科技发展的需要知识更新的应用型人才[2]。

“多元统计分析”是一门统计学专业必修课程,系统讲述统计学中多元分布的基本理论和常用的多元数据分析方法,介绍每种多元统计方法的实际背景、统计思想、统计模型,并结合统计软件解决实际问题[3]。而在工农业生产、科学研究、人文社会科学及经济管理中常常存在着这样一类问题,其涉及的变量庞杂,变量之间的关系复杂多样。但是现实中各种条件的限制又使得我们无法面面俱到,这就需要抓住主要矛盾,对数据进行降维,对问题进行提炼[4]。因此,基于现实问题的需要,对于几种降维方法的掌握是十分有必要的。多元统计分析课程中包含的常见降维方法有聚类分析、主成分分析以及因子分析,三种降维方法都考虑到了数据集的相关性或样品的相似性,都是通过某种方式在一定程度上实现数据结构的简化。具体而已,聚类分析偏向类别的划分,能从指标集和样品集两个方面实现降维。主成分分析和因子分析虽然都通过降维的方式简化了数据结构,但主成分侧重于对数据信息的压缩和精确使用,因子分析侧重于对数据结构的提炼和解释。

降维方法与其他多元统计分析方法相比,由于其理论的完整性,方法的实用性,目前已成功应用于经济管理、生物医学、人工智能、企业决策、电子商务、教育测量、地质生态等多个领域。例如,在企业决策与分析中,通过聚类分析可以将客户群体划分为不同的类别,企业据此可以开展更有针对性的营销活动。在电商平台中,聚类分析也可以让相似的用户或商品聚类在一起,其推荐系统便可以提供更加

精准的服务。另外,在基因序列研究中,由于能够采集到的样本变量个数往往远大于样本个数,面对如此高维度的信息提取,主成分分析常常被用于基因组序列表达的研究中。在人脸识别中,主成分分析也可以有效地提取数据中的关键特征,帮助使用者强化准确性。在多变量分析中,因子分析以善于挖掘数据背后潜在的影响因素著称。在教育评价与测量中,因子分析能够帮助相关人员准确分析与把握学生的各项能力发展情况。在企业调研与决策领域,利用因子分析,可以通过各种形式收集来的消费者信息,评估企业的服务质量与品牌价值。

相对已有文献,目前关注于对降维方法章节的思政教学研究较少,可参考的优秀案例并不多[4]。因此基于思政教学的需要,本文将针对多元统计分析课程教学中涉及到的三种常见降维方法,结合统计学专业学生善于用数据说话的思维方式[5],深入探讨其中的思政元素,尝试为高校进一步推进统计学专业课程思政建设提供一定的参考。

2. 课程内容与降维方法介绍

《多元统计分析》为统计学专业第5学期开设的专业必修课程,属于统计学必修技能,也是开展现代统计学研究的基础。本课程一共3学分48学时,其中理论讲授40学时,上机操作8学时。所用教材是朱建平教授编写的《应用多元统计分析》(第三版)[3]。该教材的特点在于与时俱进,构建与培养目标相适应的教学体系内容,比较适合没有太多基础的本科生,学生可以通过学习本书中的理论知识,同时与计算机结合,培养自己的动手能力,逐步掌握使用多元统计分析方法进行数据分析的技能。降维问题是多元统计分析中的重要问题之一,聚类分析、主成分分析以及因子分析则是几种常见的具体降维方法[6]。下面将简要梳理聚类分析、主成分分析以及因子分析的基本原理和方法。

聚类分析用距离衡量样品(或变量)的相似程度,并通过量化分类的方式来表达数据的结构。聚类分析通常分为 Q 型聚类和 R 型聚类, Q 型聚类的研究对象是样品,而 R 型聚类的研究对象是变量本身。具体方法包括系统聚类(根据计算出的距离的大小,从小到大将样品相继进行聚类,直到每个样品(或变量)被分配到合适的类中)和 K -均值聚类(每个样品聚集到其最近中心的类中去)。当考虑与研究对象有关的多个指标时,问题会变得复杂,并且指标间的高度相关关系可能会造成信息的重叠。主成分分析通过线性组合的方式,用较少的不相关的主成分代替原变量,利用这些主成分来揭示变量之间的内在关系,简化数据结构。假设变量之间存在一定的相关性,将相关度较高的变量分为一组,用不可观测的潜在因子代表每组,它能反映数据的大部分信息。因子分析通过研究原变量间的内在依赖关系,提炼出少数几个能反映众多指标共同作用的主要因子,达到了降维的目的[7]。

为了达到良好的课堂教学效果,在融入思政元素的过程中,我们将根据统计学专业的思维特点,用数据说话,注重严密性、逻辑性、艺术性[8]。在此基础上,在思政元素选取上要保证一定的热点性。这里需要特别注意的是如果每一节课所包含的思政元素过多,便会淡化专业知识的教学,甚至可能招致学生的抵触心理,很容易达不到思政教育的根本目标。下面我们将重点聚焦思政元素如何更好地融入聚类分析、主成分分析以及因子分析的教育教学。

3. 章节思政教学思路与实践探索

3.1. 在章节教学目标方面体现思政目标

应用型、技能型人才的培养,不仅直接关乎经济社会发展,更是关乎国家安全命脉的重大问题。对此,本课程以“能力导向”为课程建设目标、以“知识应用”为课程评价标准,为国家培养具备初步研发能力的应用型人才。应用型人才培养的关键是要将知识转化为能力。因此在降维方法章节的教学过程

中将系统梳理聚类分析、主成分分析以及因子分析的基本原理和方法, 挖掘各种方法背后所蕴含的统计思想, 总结不同降维方法在理论上的区别和联系。其具体教学目标有包括三方面:

1) 知识目标。聚类分析: 掌握熟练掌握距离判别法, 了解贝叶斯判别法, 掌握费希尔判别法, 了解相似性量度背后的本质。主成分分析: 掌握主成分的几何意, 数学推导以及性质, 包括主成分方法应用中应注意的问题。因子分析: 熟练掌握因子分析模型, 理解因子载荷矩阵求解方法, 最后包括主理解主成分分析与因子分析的异同。

2) 能力目标。能够根据应用领域和数据的特点选用合适的降维方法进行分析和推断, 能够运用聚类分析、主成分分析以及因子分析的知识 and 思想方法解决实际问题, 能够熟练应用 Excel、SPSS、Python、R 语言等软件进行数据处理, 正确利用数学、统计学的知识和方法分析解释实际问题。

3) 价值目标。在章节知识的传授中, 注重培养学生良好科学素质与科学精神, 激发学术报国的家国情怀和使命担当。能够自觉践行社会主义核心价值观, 增强思想认同、政治认同、理论认同和情感认同。

其中, 价值目标融入思政元素体现在如下三方面:

1) 引导学生践行社会主义核心价值观, 爱党爱国, 遵纪守法, 能够在数据收集、数据分析、数据处理实践中理解并遵守相关领域的职业道德规范。

2) 通过丰富的案例教学分析(例如, 如何通过各种降维方法对各地的科技创新指标进行评价和建议), 引导学生了解与社会经济统计、医药卫生统计、生物统计或工业统计等某一领域的基本知识, 培养学生良好的科学素质与科学精神, 实现专业创新持续发展。

3) 依托自改革开放以来, 中国积极融入全球化进程, 加入世界贸易组织后, 在各项指标数据上所取得的成就(例如, 经济、民生、外贸和科研等方面的数据), 激发学生爱国主义情怀, 内化社会主义制度自信, 激发学术报国的家国情怀和使命担当。

3.2. 在章节教学内容方面融入思政元素

目前多元统计分析课程教学中使用的教材普遍是分章节讲授的, 这样比较容易显得课程体系完整, 有利于培养学生运用课程知识解决实际问题。我们在各大章节模块的教学过程中, 结合专业教学内容, 引入大量的思政案例分析, 案例中分别体现为科学的探索精神, 崇高的人生理想, 忠诚统计, 乐于奉献, 实事求是, 善抓主要矛盾, 勿以恶小而为之, 勿以善小而不为等思政元素。章节模块教学的具体思政元素详见表 1。

Table 1. Introduction of ideological and political elements in the modular teaching mode

表 1. 章节模块教学模式思政元素的切入

模块	教学内容	思政案例	思政元素
降维方法之一: 聚类分析	1) 相似性度量。	1) 随着改革开放的深入发展, 我国城市经济群快速发展, 其背后所代表的聚集学习和自主学习的意识和能力, 能够自我规划、自我管理, 学会学习, 学会发	1) 物以类聚, 人以群分, 引导学生终生经济群快速发展, 其背后所代表的聚集学习和自主学习的意识和能力, 能够自我规划、自我管理, 学会学习, 学会发重要力量。聚集经济效应, 又称为: 规模, 不断缩小与优秀人群的差距。模经济效应, 简称聚集经济, 集聚效应、群聚效应等。 2) 使学生意识到在目前的大数据时代, 应该相信科学, 脚踏实地, 培养自己执着专注、精益求精、一丝不苟、追求卓越的工匠精神。实现个人精神的锲而不舍、金石可镂。
	2) 系统聚类分析法。	2) K 均值聚类的核心是通过不断迭代寻找最优的均值中心, 从而到达最优的聚	
	3) K 均值聚类分析。	2) K 均值聚类的核心是通过不断迭代寻找最优的均值中心, 从而到达最优的聚	
	4) 有序样品的聚类分析法。	2) K 均值聚类的核心是通过不断迭代寻找最优的均值中心, 从而到达最优的聚	

Continued

<p>1) 主成分的几何意义及数学推导。</p> <p>降维方法之二: 2) 主成分分析的性质。</p> <p>3) 主成分方法应用中应注意的问题。</p>	<p>1) 当今国际形势复杂, 国际贸易环境乱象丛生, 因此我国需要不断提升对外开放水平, 加快从贸易大国向贸易强国的转变。为了响应这一号召, 促进全国各地外贸水平的发展, 各级政府针对各类外贸方面的数据, 探究外贸竞争力的影响因素, 对各地区的外贸竞争水平进行综合评价, 为更好地发展外贸水平给出建议。</p> <p>2) 随着云计算和大数据技术的逐步成熟, 近几年全国各地公安机关结合人工智能技术中的人脸识别深入挖掘与分析案件, 优化了现代安全工作效能, 取得了许多成功案例, 凸显了数据分析在行业中的重要地位。</p> <p>1) 介绍外贸竞争力的影响因素分析时, 强调主成分分析可以在该领域发挥显著的作用, 以此进一步激发学生学术报国的家国情怀和使命担当。</p> <p>2) 人脸识别中的关键技术便是主成分分析与因子分析, 在此背景下讲解章节知识背后的原理, 更好地使学生认识到要善于抓住问题的主要矛盾和矛盾的主要方面, 理论可以应用于实践, 同时也要注重量的积累, 促成质的飞跃。</p>
<p>降维方法之三: 因子分析</p> <p>1) 因子分析模型。</p> <p>2) 因子载荷矩阵求解。</p> <p>3) 公因子重要性的分析。</p>	<p>1) 2020 至 2022 年是全世界抗击新冠肺炎疫情的三年, 期间局部地区时有爆发疫情传染事件, 部分媒体账号借机杜撰夸大疫情蔓延误导信息, 严重干扰政府与社会疫情防控大局, 但好在中国政府能够实事求是, 及时公布疫情, 最终取得抗疫的重大胜利。作为新时代的大学生我们更应该透过现象看本质, 不被表象所迷惑。这也是因子分析所蕴含的哲学思想。</p> <p>2) 课堂上, 以近年来不断凸显的大学生就业压力为主题, 布置学生以小组为团队, 通过问卷调查对我校及周边大学的学生就压力情况和个人未来发展规划进行研究分析, 综合运用相关分析、方差分析、因子分析等多种统计学方法研究大学生压力状况和就业规划等影响因素, 并分析各因素之间的关系, 从而探寻其面临的主要问题及成因, 尽可能帮助该群体提供有用的信息及参考建议, 也希望能够帮助自己更好的规划自己的未来, 平稳步入社会。</p> <p>3) 邀请证券基金领域从业者对行业发展最新状况作专题报告, 通过与学生课堂的互动, 将统计领域的先进方法应用于投资实践完整地展现在大家面前, 加强沉浸式“大思政课”的交互感。</p> <p>1) 通过抗疫时期的大量不实谣言的例子讲述, 告诫学生要善于抓住事件背后的根本逻辑, 理解其中的前因后果, 而不是被谣言的表象、影响了判断。这也是因子分析的思维方式。</p> <p>2) 在问卷调查环节中, 通过小组分工协作, 培养团队意识和责任心, 能够积极承担责任, 发挥出个人在团队中的作用。同时强调学生的动手操作能力、解决实际问题的能力和扎根基层的务实精神。</p> <p>3) 目前量化投资, 量化金融以及量化基金等领域发展迅速, 对统计人才的渴望相当强烈。通过面对面与从业者的交流, 以此让学生顺应时代与专业发展的需求进行自我规划, 树立崇高的人生理想, 并为此努力奋斗, 实现个人价值。</p>

3.3. 在章节考核方面突出思政要求

我们以培养学生能力为导向, 以统计方法应用为课程考核标准, 根据教学内容和思政内容建立了包含专业知识与思政育人两方面的考核体系。

1) 专业内容考核。一方面, 三种降维方法各适用于解决什么样的问题? 其中, 聚类分析适用于解决

对指标集和样品集无监督地划分类别的问题。主成分分析适用于解决对指标集高度相关的数据进行处理或者综合评价类问题。因子分析适用于解决对指标集高度相关的数据进行解释或者综合评价类问题。最后结合以上结论, 面对实际问题需要进行结构简化时, 我们要充分考虑到数据的特征和研究目的, 分清不同方法的区别和联系, 结合不同方法的优点, 取长补短, 选择合适的方法, 发挥结构简化方法的最大效益。另一方面, 三种降维方法各自的优缺点是什么? 聚类分析的优点在于思想直观, 结论简洁, 方法多样。缺点在于受数据性质的限制和影响比较大, 有时分类的结果不够准确, 效果不好。主成分分析的优点在于便捷地压缩变量高度相关的数据信息, 并且能量化压缩的信息量, 准确计算出结果。缺点在于, 受数据集指标的限制, 有时提取信息的效果不佳, 并且模型线性的表达可能不完全符合实际情况。因子分析的优点在于能较好地提炼并解释相关性背后的原因, 能量化因子的解释力度。缺点在于不能准确地计算出因子的值, 要求指标集高度相关, 否则效果不佳。

2) 思政元素考核。在三种降维方法讲授完成后, 教师提供几个主题, 比如: 大学生就业压力相关的主题, 大学生的手机使用情况及其影响或者“宅”生活下大学生消费情况调查分析等, 针对某一具体主题, 学生按3人一组进行问卷调查实践, 首先设计出合理的调查问卷, 问卷可以包含5道量表题, 通过问卷调查的方式收集并分析数据, 将统一设计的问卷发放给调查对象填写, 收集并整合好最终数据。之后通过描述性统计、显著性检验、聚类分析、主成分分析、因子分析和信效度分析等统计方法研究所选的主题问题。通过适用的统计分析方法, 深入挖掘数据背后隐藏的信息, 提升数据价值, 得出可靠的结论。整个过程中, 学生须归纳出分析的结论, 制定出良好的对策以及展现出必要的思政元素, 最终形成一份完整富有知识性与思政性的抽样调查报告。同时可结合当下时事新闻, 对思政元素进行论述。最终, 教师根据各小组报告表现, 特别是在实践调查过程中所体现出的思政素养进行综合评价。

3.4. 在章节目标达成度方面注重思政育人

课程目标达成度评价是判断各章节思政教学成果的重要环节之一。为了保障结果的客观性与时效性, 我们将在每章节结束, 借助思政元素教学评价表(见表2), 对学生最新评价和反馈进行分析。旨在实时获取学生对“多元统计分析”课程所授章节评价和反馈, 其中, “达成”代表有显著作用, “未达成”代表没有完成目标, “弃权”表示放弃。

Table 2. Teaching evaluation of ideological and political elements on cluster analysis, principal component analysis and factor analysis

表 2. 聚类分析, 主成分分析以及因子分析章节思政元素目标达成度评价表

课程授课情况	达成	未达成	弃权
自觉践行社会主义核心价值观, 增强思想认同、政治认同、理论认同和情感认同			
诚信、守法, 具有职业操守, 有高度的责任感			
增强了爱国主义情感, 增强民族自豪感			
注重量的积累, 促成质的飞跃			
忠诚统计, 乐于奉献; 实事求是, 不出假数; 依法统计, 严守秘密; 公正透明, 服务社会			
树立崇高的人生理想, 并为此努力奋斗, 实现个人价值			
相信科学, 脚踏实地, 经得起诱惑, 不轻信不盲从			
爱国情怀和工匠精神, 乐于奉献, 爱岗敬业, 鞠躬尽瘁			
具有良好的团队合作能力			
能够顺应时代与专业发展的需求进行自我规划, 具有职业发展的意识和能力			

通过章节目标达成度评价以及上一节的章节考核, 我们发现, 在融入思政教学后, 育人成效显著。首先, 学生的思政素养得到了很大的提高, 以及增强了社会责任感。其次, 培养了学生专业知识的思辨性和创新性, 最后, 实际动手操作能力和小组团结协作能力都有所加强。在面对实际问题时, 可以对数据进行结构简化, 同时结合不同方法的优点, 取长补短, 最终发挥思政育人的最大效益。

4. 结语

现如今, 大数据迅猛发展, 无论是何种领域、何种产业, 数据分析对于各行各业高效能, 科学决策都具有举足轻重的作用。而聚类分析、主成分分析以及因子分析都是多元统计分析课程中重要的降维方法, 其旨在培养学生善于抓住主要矛盾, 对数据进行简化, 对问题进行提炼。该章节的内容与实际生活联系紧密, 在授课过程中教师可以很好地结合时事热点进行思政元素的融入。思政教育是当代教师的使命, 教师需要具有终生学习和自主学习的意识和能力, 能够自我规划、自我管理, 学会学习, 学会发展, 不断提升自身的思想政治素养。在教书育人的过程中, 培养高度的社会责任感和良好的职业道德, 积极践行社会主义核心思想价值观。

基金项目

苏州科技大学数学科学学院课程思政示范课程建设项目(MKCSZ202001); 苏州科技大学数学科学学院课程思政示范课程建设项目(MKCSZ202003)。

参考文献

- [1] 刘美芬. 课程思政视域下《统计学》教学改革探索[J]. 科技资讯, 2021, 7(4): 154-156.
- [2] 龚一鸣. 课程思政的知与行[J]. 中国大学教学, 2021(5): 77-84.
- [3] 朱建平. 应用多元统计分析[M]. 北京: 科学出版社, 2016.
- [4] 董小刚, 郭志元, 袁晓惠, 王纯杰. 《多元统计分析》的课程思政教学模式改革探索[J]. 教育现代化, 2019, 6(85): 148-149.
- [5] 林海明. 主成分分析综合评价应该注意的问题[J]. 统计研究, 2013, 30(8): 25-31.
- [6] 林海明. 因子分析综合评价应该注意的问题[J]. 数理统计与管理, 2019, 38(6): 1037-1047.
- [7] 周文强, 阙思思, 曾德芳, 等. 基于典型相关分析的长江航运与水环境间关系的实证研究[J]. 长江流域资源与环境, 2020, 29(3): 1183-1191.
- [8] 许瑞芳, 张宜萱. 沉浸式“大思政课”的价值意蕴及建构理路[J]. 思想理论教育导刊, 2021(11): 83-88.