

大数据时代下《应用回归分析》课程的教学改革探索

高润诚, 陈旭东, 田瑞琴

杭州师范大学数学学院, 浙江 杭州

收稿日期: 2024年4月8日; 录用日期: 2024年5月10日; 发布日期: 2024年5月21日

摘要

在当今大数据时代,《应用回归分析》作为各个高校统计学专业的一门必修课,其重要性不言而喻。另外,目前大多数高校都把这门课建设得比较成熟完善,但在当今大数据的浪潮下,其仍存在一些需要我們进一步探索、创新、改革的空间。故本文通过深入研究《应用回归分析》课程在当今大数据时代下可能存在的一些不足和仍需改进的地方,来进一步探索如何在当今日新月异的时代使《应用回归分析》焕发新的生机和活力,帮助人们更好地学习回归分析。

关键词

大数据, 应用回归分析, 教学探索与改革

Exploration on Teaching Reform of Applied Regression Analysis Course in the Era of Big Data

Runcheng Gao, Xudong Chen, Ruiqin Tian

College of Mathematics, Hangzhou Normal University, Hangzhou Zhejiang

Received: Apr. 8th, 2024; accepted: May 10th, 2024; published: May 21st, 2024

Abstract

In the era of big data, the importance of Applied Regression Analysis as a compulsory course for statistics majors in various universities is self-evident. Moreover, although most universities have now developed and perfected this course, there are still some areas that need further exploration,

innovation, and reform in the current wave of big data. Therefore, this paper conducts in-depth research on the possible shortcomings and areas that still need improvement in the course of Applied Regression Analysis in the current era of big data, in order to further explore how to make Applied Regression Analysis radiate new vitality and vigour in today's fast-changing era, and help people learn regression analysis better.

Keywords

Big Data, Applied Regression Analysis, Teaching Exploration and Reform

Copyright © 2024 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

1855年,“回归”的思想在高尔顿的笔下流出,此时距离1969年互联网的诞生,相差也不过百余年。不仅如此,从高斯提出最小二乘法算起,回归分析已经有200多年的历史[1]。回顾这200多年的历史,我们难以找到没有它的领域,其在社会的广泛运用,日益彰显它的庞大的生命力与活力,这种生命力没有道理不能与当今大数据相结合。不仅如此,在当今大数据时代,数据类型繁多,数据量庞大,价值密度低,同时其要求人们对于数据的处理要更加快速,时效要更高,这些无疑是对当今统计学的理论和方法提出了新的挑战 and 困难[2]。回归分析可以为我们提供一些思路和方法来解决我们所面临的难题。回归分析作为统计学一个十分重要的分支,可以帮助我们确认变量之间的相互关系,其方法不仅包括lasso回归、稀疏回归,同时还有回归树等方法,运用范围十分广泛,在当今大数据时代显得更加重要[3]。

大数据所产生的巨大的生产力引发了各行各业的重视,这不仅为统计学专业、数据科学与大数据技术专业等多个专业的人才提供了更多的就业土壤、就业途径以及实现人生价值的机会,同时也对以上专业人才的培育和专业素质提出了更高的条件[4][5]。如何解决当今大数据时代提出的问题与挑战,这个不仅是我们的困境,同时也是我们的机遇,本课程作为大数据应用的基础课程,更应该主动进行教学改革、锐意进取,来帮助学生们更好面对和解决大数据时代所带来的挑战,为社会提供更多高质量的大数据人才。

时至今日,许多学者们注意到《应用回归分析》所具有的巨大潜力,他们围绕《应用回归分析》课程教学改革的科学性、规范性,合理性等问题提出了许多宝贵有益建议,这些建议对《应用回归分析》课程的教学质量改善提供了强而有力的智力支持和理论支撑,但是纵观这些建议,对教学内容、教学方法和考核方式等方面的探讨并不是非常多,对于这些方面的前沿性、针对性、建设性意见也并不多[6]。针对这种大数据环境的情况[7],本文笔者于所在学校统计专业发放了调查问卷,在调查问卷的统计结果的基础上,对《应用回归分析》课程的教学内容、教学方法、考核方式等方面的教学现状进行研究并提出教学改革的思路[8]。

2. 《应用回归课程》课程教学现状及问题

2.1. 《应用回归分析》的教学内容

回归分析是统计学中一个极其重要的分支,它在自然科学、社会与管理科学和社会经济等领域有着极其普遍的应用。在本科阶段,教学内容以一元和多元线性回归的参数估计、显著性检验及其应用,逐

步回归、岭回归、主成分回归与偏最小二乘、曲线回归、多项式回归、非线性回归和违背回归模型基本假设的异方差、自相关和异常值与多重共线性的诊断与处理方法为主。在大数据时代的背景下，将其产生的数据与书本所提供的传统数据相比较，首先互联网平台所产生的数据样本量更大，从大量数据中挖掘重要信息的难度更大。其次，互联网平台所产生的数据具有非结构化的特征，这要求学生在收集数据后，对数据进行清洗和预处理。其中可能包括处理缺失值、异常值，以及对数据进行归一化、标准化等操作，以此来确保数据的可靠性和一致性。在这样的背景下，仅学习书本上对传统数据的简单处理与回归应用可能会使同学将来就业时产生巨大“落差”，缺失处理新型大数据的能力。最后，《应用回归分析》课程于统计专业的同学而言，与将来期望工作的相关性较强，为此培养同学们数据处理和分析的能力非常重要。先前发放的调查问卷统计结果显示，针对书本案例是否需要改进的问题，总计 209 份调查问卷中(见图 1)，70 份认为需要改进，占总问卷数量的 33%。而在课上是否需要引入泰迪杯、挑战杯等竞赛例题的问题上，154 份认为需要引进，占总问卷的 73%。由此可见，同学们希望在书本中引入更加新颖且贴合实际运用的实例。

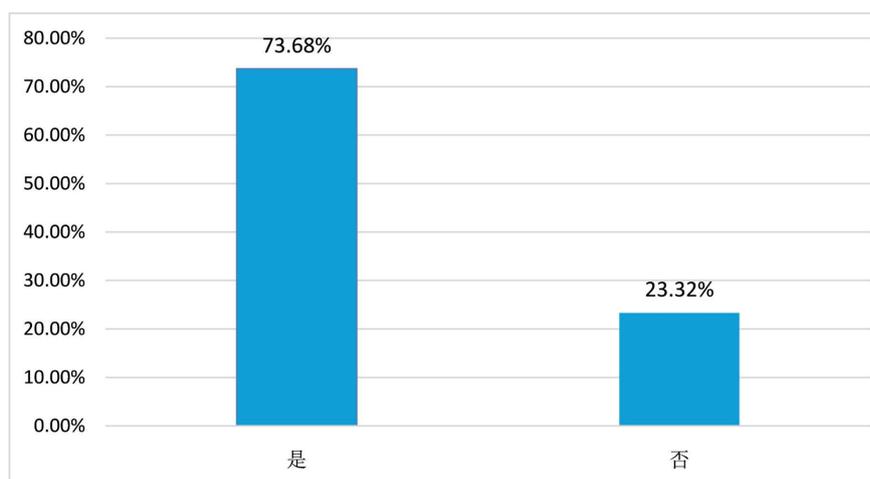


Figure 1. About whether the cases in Applied Regression Analysis need to be improved
图 1. 关于《应用回归分析》案例是否改进

2.2. 《应用回归分析》的教学方法

传统的教学方法主要采用的是“回归分析方法理论讲解 + 书本案例”的模式，学生以听讲和上机课通过程序进行回归方法代码实现为主，教师则以 PPT，黑板讲课为主，在课堂上鲜有学生能够与教师互动，教学方式单一甚至讲课效果往往不太好。R 语言成为课上同学进行实验的主要工具，在上机课上需要验证的实验较多，但是由自己设计，包含多个知识点的实验过少，在理论课上教师 PPT 演示较多，但能让学生主动发现并解决问题的机会较少。并且由于课程的设置，即便是能用于实践的上机课往往也会变为理论课，使上机课更显得有些名存实亡。这种“重理论而轻实践”的教学方法容易使学生在面对日常生活的实际问题时产生无力感，不利于学生将理论付诸于实践。本质上《应用回归分析》课程不应该忽视如何将理论反作用于实践并且指导实践，这种传统的教学方法固然有利于培养学生的归纳推理思维，但过于注重反而又不利于学生的发散性思维培养。通过调查问卷(见图 2)有 1/3 学生认为该课程的教学方法需要改进。不仅如此，有 78.47% 的学生认为《应用回归分析》的主要目的在于实际应用和提高专业技能，这反映出学生更主要期待这门课可以为自己的未来提供帮助。但是传统教学局限于基本回归模型的介绍和一些常见情况下对回归模型的处理，未能进一步指引学生如何深入学习，引导其掌握更多前沿的

回归模型及其应用方法，在面对日益复杂的实际问题下，这容易导致学生面对大数据时无从下手，不利于学生在当今大数据时代下迅速掌握理解并灵活运用前沿的回归模型和应用方法。

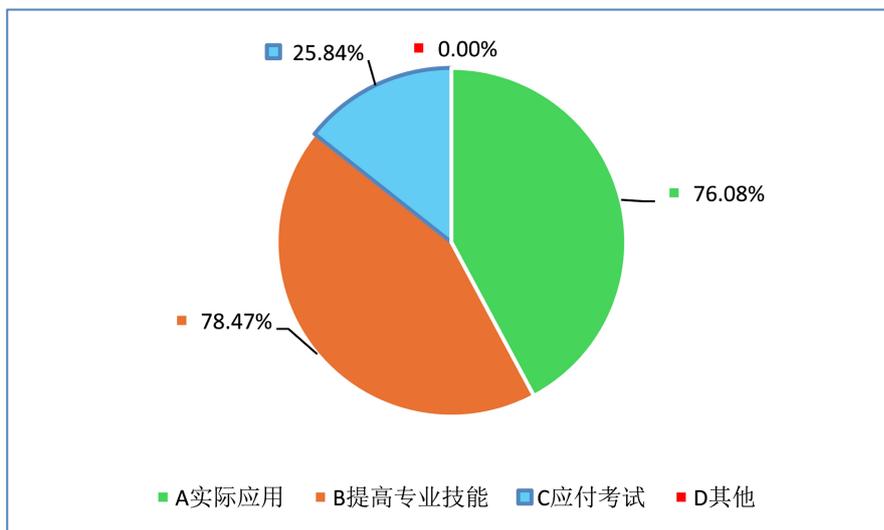


Figure 2. The main purposes of Applied Regression Analysis course
图 2. 《应用回归分析》课程主要目的

2.3. 《应用回归分析》的考核方式

目前《应用回归分析》的考核方式仍为传统的闭卷笔试考核，显然面对这种实际应用性很强的课程，仅单纯采用闭卷形式难以反映出学生对于该课程的掌握程度。不仅如此，对于实践考核的环节又不够重视，容易使实践考核环节流于形式，不利于培养学生主动实践的积极性。调查问卷的具体数据表明(见图 3)，有 58.85% 的人们希望将小组合作，课程论文引入平时成绩。由此可见许多同学希望该课程的考核方式不应该拘泥于传统的闭卷考试。在具体的考核方式上，与当今大数据时代背景相结合，引入课程论文、小组报告能显著提升同学们的学习自主性和积极性。

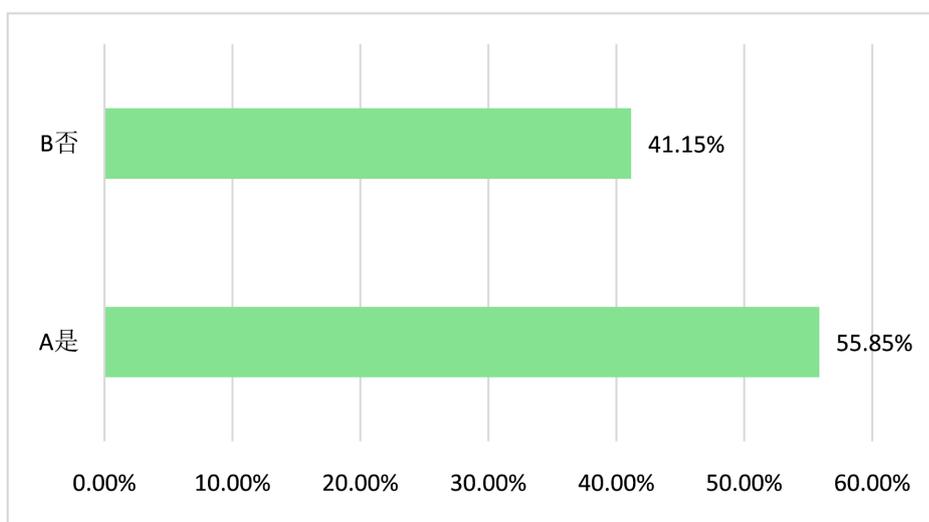


Figure 3. Whether group cooperation should be brought into Applied Regression Analysis
图 3. 《应用回归分析》是否引入小组合作等

3. 《应用回归课程》教学改革的思路

大数据时代同时也是互联网时代，能否与互联网时代进行深度融合，能否将互联网运用到教学改革上来，能否适应现在的互联网新环境，将信息化与教学改革相融合，这不仅是衡量优秀教师的重要标准，同时也是衡量教学改革是否有效的重要标准。教学活动的构成要素 1) 受教者(学生); 2) 施教者(教师); 3) 教学目的; 4) 教学内容; 5) 教学方法; 6) 教学环境; 7) 教学设备。其中最重要的无疑是学生，老师，教学内容。而完成教学任务则需要一定的教学方法，倘若没有使用适当的教学方法，则会难以实现教学目标和任务，进而影响到整个教学系统功能的实现。考核则是作为一种检验教学模式是否合理有效且科学的一种重要凭借，可以想象，在不同的考核下，同样的教学方法和教学内容也会产生不同的考核结果。所以本文将从教学内容，教学方法，考核方式等方面对《应用回归课程》提供教学改革思路。

3.1. 优化教学内容

《应用回归分析》课程应当秉持以“精简、拓展、应用”为原则。《应用回归分析》的教学内容主要是以一元或者多元线性回归，回归诊断，回归变量的选择，参数估计方法的改进，以及非线性回归等为内容。在目前大数据时代海量的数据下，这些传统的回归方法难以承担从大数据中发现和挖掘重要信息的责任，这就需要在该课程中引入一些新的拓展内容，拓展新的回归方法，例如数据预处理，决策树，MCMC 等方法。这更有助于在大数据时代下实践应用。同时，《应用回归分析》的案例也存在老旧，简单等问题，这些案例与实际生活脱离，难以吸引学生们的兴趣，也无法反映出大数据时代下数据类型繁多、数据量庞大、价值密度低等特点。我们应当使用一些目前较为新颖、贴近生活的案例来激发学生的学习兴趣。因此，我们可以将《应用回归分析》与各类赛事平台(例如正大杯全国市场研究竞赛，泰迪杯数据挖掘大赛，全国大学生数学建模大赛等全国性竞赛)相结合，这些赛事平台不仅题目比较贴近实际问题，例如保险反诈预测，银行客户认购金融产品预测，而且在具体赛事案例上，《应用回归分析》课程所涉及的回归概念，回归方法都可以为解决具有庞大数据量的具体实践问题提供必不可少的工具。

3.2. 改进教学方法

大数据时代的背景下，传统黑板 + PPT 的讲课教学方式在调动同学的上课积极性和学习自主性上显然有所不足，为此在课程中引入慕课、微课、翻转学堂等新型教学方式尤为重要。慕课、微课等依托新媒体的网课方式可以为无法跟上老师教学节奏、课上打小差的同学提供再次学习的机会，为课上听懂的同学提供一条复习的途径[9]。同时老师也可以根据网课平台提供的观看数据了解同学们对哪个知识点依旧存在较大的问题，以便在之后的课程上有针对性地复习。翻转课堂则是给同学们一个“当老师”的机会，在对教学内容自我学习并充分理解的前提下，同学为同学们讲解知识点，使得在增添教学方式多样化的同时，也提高同学们的自主学习能力。不仅如此，还可以制作一些关键课程的微课来帮助学生学习[10]。此外，调查报告显示，同学们对将小组合作、课程论文引入平时成绩较为支持。可以通过小组论文展示的方式，让同学们对一组数据进行处理和分析，此方式不仅可以锻炼同学们的数据处理、数据分析的能力，了解数据分析的流程，同时可以提高同学们的实践能力、团队协作能力、心理素质等，使同学们在参与泰迪杯、挑战杯等相关竞赛时更加得心应手。

3.3. 拓展考核方式

调查报告显示，同学们学习《应用回归分析》课程的主要目的是实际应用和提高专业技能，正如《应用回归分析》的课程名称所述，本课程注重“应用”。传统笔试证明等闭卷考核方式并不受同学所青睐，但掌握“回归”的相关证明又尤为重要，因此拓展考核方式是当务之急。考查学生学习成果时，需适当

地增加学生利用理论知识解决实际问题的能力和基本专业素质考核的权重[9]。在期末考核的基础上,可以通过增加平时成绩的分数权重,将实践报告、小组合作展示、竞赛名次、随堂测验等作为平时成绩的重要指标,这不仅可以实现考核形式多元化,还可以尽量避免同学考完就忘,缺乏实践能力。

4. 结束语

在当今大数据蓬勃发展的时代背景下,《应用回归分析》课程的重要程度不言而喻,为此推动课程改革迫在眉睫[11]。本文根据调查报告结果,对《应用回归分析》的教学内容、教学方法、考核方式等方面就教学现状和改革方式提出了一些个人见解。问题的关键在于如何在改革中体现出以“学”为中心[12]。由于数据的样本量不足、个人看法过于主观等因素,本文对于《应用回归分析》的改革建议上依旧存在一定的局限性。

参考文献

- [1] 何晓群,刘文卿. 应用回归分析[M]. 北京: 中国人民大学出版社, 2011.
- [2] 田茂再. 大数据时代统计学重构研究中的几个热点问题[J]. 统计研究, 2015, 32(5): 3-12.
- [3] 刘冬. 《应用回归分析》课程教学改革探索——以赤峰学院为例[J]. 赤峰学院学报(自然科学版), 2022, 38(9): 97-100
- [4] 韩静舒. 大数据背景下“统计学”课程教学方法研究[J]. 教育教学论坛, 2019(9): 178-179.
- [5] 周曙东, 李静. 适应大数据时代的统计学教学改革探讨[J]. 中国统计, 2018(10): 59-61
- [6] 罗敬, 胡军浩. 《应用回归分析》课程的教学改革探索——以中南民族大学为例[J]. 中国新通信, 2020, 22(22): 197-198.
- [7] 何兴海. 大数据时代科学统计问题研究[J]. 统计与管理, 2017(12): 6-7.
- [8] 邓丽. 对分课堂在“应用回归分析”课程教学中的应用[J]. 西部资质教育, 2022, 8(5): 166-168.
- [9] 王化琨, 李春艳, 陈莉莉, 周亚晶. 《应用回归分析》课程教学内容与教学方法的改革与实践[J]. 黑龙江教育(理论与实践), 2015(Z1): 81-82.
- [10] 郭洪伟. 微课对《统计学》教学效果提升的实践[J]. 教育教学论坛, 2015(38): 124-125.
- [11] 魏云舒. 大数据情境下高校教学质量评估改进路径探析——基于 20 篇《本科教学质量报告》的研究[D]: [硕士学位论文]. 北京: 对外经济贸易大学, 2020.
- [12] 王静, 程从华. 统计学专业主干课程的教学思考[J]. 高等数学研究, 2014, 17(2), 57-59.