

基于大数据的《多元统计分析》教学研究

程丽娟

岭南师范学院数学与统计学院, 广东 湛江
Email: chenglijuan666@126.com

收稿日期: 2020年11月13日; 录用日期: 2020年12月23日; 发布日期: 2020年12月30日

摘要

大数据时代, 数据的爆炸式增长给多元统计分析课程教学带来了前所未有的机遇和挑战, 通过优化课程体系, 强化统计软件的辅助性, 充分发挥网络资源的优势和推进以赛促学等方法, 加强大数据时代数据挖掘和多元统计分析的理论教学, 增强真实数据分析与应用的实践实训, 以培养出更适应市场需求和社会发展的统计学人才。

关键词

多元统计分析, 大数据, 以赛促学

The Teaching Research of “Multivariate Statistical Analysis” Based on Big Data

Lijuan Cheng

School of Mathematics and Statistics, Lingnan Normal University, Zhanjiang Guangdong
Email: chenglijuan666@126.com

Received: Nov. 13th, 2020; accepted: Dec. 23rd, 2020; published: Dec. 30th, 2020

Abstract

In the era of big data, the explosive growth of data brings unprecedented opportunities and challenges to the course teaching of multivariate statistical analysis. By optimizing the course system, strengthening the auxiliary of statistical software, giving full play to the advantage of network resources and promoting learning by competition, the theory teaching of data mining and multivariate statistical analysis in the era of big data is strengthened, and the practical training of real data analysis and application is strengthened, so as to cultivate statistics talents more suitable for market demand and social development.

Keywords

Multivariate Statistical Analysis, Big Data, To Promote Learning

Copyright © 2020 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

大数据时代, 各种各样的数据爆炸式增长, 面对海量的数据, 如何在满足精度的条件下, 尽量确保数据的准确性和完整性, 合理的进行数据处理和分析, 是我们解决的主要问题, 也给《多元统计分析》课程的教学带来了前所未有的机遇和挑战。

2. 大数据时代, 《多元统计分析》教学的挑战和机遇

大数据时代, 互联网和云技术的广泛应用, 给我们的工作和生活带来极大的便利, 2015年5月, 李克强提出: 大数据产业是中国推动互联网+战略的重要支撑。数据规模和数据模式的改变, 使得我们可以获得的数据不断的累积和扩充, 可以进行量化的事物越来越多。多元统计分析是使用统计的方法对多维数据进行分析, 数据的多样性和复杂性给多元统计分析带来了更大的挑战和机遇。

2.1. 多元统计分析教学的机遇

大数据时代, 互联网等相关技术的迅猛发展, 越来越多的课程资源都能从网络平台直接获取, 给多元统计分析课程教学带来了前所未有的机遇。一是, 整合网络资源, 不断地增强知识的融合, 提高教学效率。在线学习逐渐成为大家获取知识的新的途径, 面对海量的学习资源, 通过名师课堂的学习, 不仅可以提高学习效率还能提高上课效果。通过名师讲座的学习, 不仅增强对学科前沿的了解, 还能不断提高自身的专业修养。通过和不同学校的学者交流和沟通, 碰撞出不同的思维火花。通过给学生整合网络资源, 不断地增强知识的融合, 提高教学效率。二是, “翻转课堂”式教学, 提高教学质量。当在线学习在大数据时代逐渐成为学生获得知识的主要途径之时, 传统的课堂教学则会起到辅助教学的作用, “翻转课堂”式教学组织形式将改变以往老师课堂讲授, 学生课余练习的教学模式。学生自主学习遇到的困难通过课堂讲解答疑解惑共同探究进行解决, 不仅提高学生的自学能力, 从很大程度上提高教学质量。三是, 创造了个性化的学习环境, 提高学生学习的积极性。教育环境的改变。大数据为学生创造了个性化的学习环境, 为教师创建了丰富的教学平台, 根据学生学习情况和个人特点, 教师开展有针对性的教学指导, 因材施教, 不仅能提高学生的学习能力, 还能提高学生的学习积极性, 变要我学为我要学, 更好的促进学生的自我发展。

2.2. 多元统计分析教学的挑战

传统的多元统计分析教学中, 学生主要通过老师在课堂上讲解的多元统计分析的理论知识, 以及上机实验课中, 上机实现老师给的程序。大数据时代, 信息的爆炸式增长给多元统计分析课程教学带来了前所未有的挑战[1], 具体表现在: 第一, 教学模式的变革。大数据时代, 随着互联网技术的飞速发展, 学生学习知识不再局限于课堂教学, 网上在线学习日益成为获取知识的重要途径。面对海量的学习资源, 如何从中快速筛选出合适的教学资源, 又能提高上课的效果, 是多元统计分析教学面临的重大挑战。第

二，教学观念的转变。当在线学习在大数据时代逐渐成为学生获得知识的主要途径之时，传统的课堂教学则会起到辅助教学的作用。如何合理的调整这种转变，这是大数据时代，多元统计分析教学面临的另一挑战。第三，教育环境的改变。大数据为学生创造了个性化的学习环境，为教师创建了丰富的教学平台，如何根据学生学习情况和个人特点，开展有针对性的教学指导，这也是多元统计分析教学即将面临的又一挑战。

3. 大数据时代下多元统计分析教学研究

多元统计分析是传统统计学发展的一个重要分支，它研究多个变量的统计总体，它可以把多个变量对一个或多个变量的作用程度大小线性地表示出来，反映事物多变量间的相互关系，因此广泛应用于数据处理和挖掘等领域。大数据时代，数据的多样化给多元统计分析教学带来了新的机遇和挑战。

3.1. 优化多元统计分析课程体系

多元统计分析是进行科学研究的一项重要工具，在自然科学，社会科学等领域有广泛的应用。传统多元统计分析课程教学模式中重视理论教学，轻视实践教学，没有重视软件的学习和应用，在知识的讲解中，选取的数据是处理后的数据，数据的实用性较弱。

1) 既重视理论教学，又重视实践教学

传统的多元统计分析教学主要是通过基础理论的学习让学生熟悉相关理论知识的思想和推导过程等数学方法，通过例题给出相关知识的应用，而忽视了例题的数据是经过处理的理想型数据。大数据时代，由于数据的多样性和复杂性，需要先对数据进行预处理之后才能运用理论知识进行分析和处理，因此，在进行多元统计分析学习的时候，不只要告诉学生理论知识来源的重要性，还要提高他们实际分析问题的能力[2]。因此，在构建统计专业多元统计分析课程体系时，既要强调理论来源的重要性，还要注重案例教学的实用性，提高学生实际分析问题的能力。

多元统计分析是一门应用性很强的专业课程，学会多元统计分析理论的来源是统计专业区别于计算机和金融等其他专业进行数据分析的优势，大数据时代，面对海量的数据，只使用理论对问题进行求解时，不仅花费很多的时间，而且得不到准确的结果，因此，理论联系实践，使用软件对数据进行分析，提高实际分析问题的能力可以使统计专业的学生在大数据时代具有更强的就业优势。

2) 与机器学习相结合，提高教学的实用性

大数据时代，数据的多样化使得数据处理的方法也是多样的。在神经网络等机器学习方法刚出现时，准确度较低，而且只能进行二分类，因此，对多维数据进行分类时，使用多元统计分析中的判别分析方法效果更好。在教学时，可以以多元统计分析的方法为主，机器学习的方法为辅，通过对比，给出判别分析在分类中的优势和劣势。在机器学习发展趋于完善时，机器学习算法不仅可以进行多分类，而且可以通过调参等手段提高分类效率，在教学时，可以分别使用机器学习和判别分析两种方法处理问题，通过对结果进行比较来判断哪种方法更优。通过比较既增强了学生处理问题的严谨性，也提高了学习的趣味性，增强学生学习多元统计分析的主动性。

3) 根据校企合作项目，增强学生处理问题的能力

借助于和统计局联合或者校企联合的培养模式，根据他们的需求，不断修改多元统计课程体系。如在人口普查中，主要注重数据的真实可靠性，可以专门设置一节课从数据收集和处理各方面解决人口普查的数据分析问题。通过建立相应的资源共享平台，和合作单位形成稳定的合作关系。让学生参与到合作单位的具体项目中，通过对不同项目的数据分析过程，增加学生对多元统计分析理论知识的理解，并掌握如何将学到的理论知识应用于解决实际问题的能力。

4) 根据市场需求, 提高学生的竞争力

大数据时代, 数据的多样化使得数据处理的多样化, 也给毕业生找工作带来新的挑战。如何在竞争日益激烈的人才大战中脱颖而出, 这对学生统计知识的学习提出更高的要求, 也给多元统计的教学提出更大的挑战。作为一门实用性很强的课程, 在教学中, 除了学习多元统计的思想方法之外, 还有根据市场的需求, 掌握大数据下, 多元统计分析的新的需求, 不断地更新多样化的数据处理方法, 掌握新的软件需求。总之, 可以通过不断地修改教学内容, 培养更适应市场需求和社会发展的统计学人才。

3.2. 强化统计软件的辅助性

大数据时代, 数据量较大, 如果使用传统的多元统计分析方法处理, 运算量较大, 运算效果较差, 使用统计软件不仅节约时间, 而且可视化效果较好[3]-[8]。适合多元统计分析课程的统计软件较多, 如 SAS、R、SPSS 和 Python 等软件, 这些软件中 SPSS 是相对简单的统计分析软件, 不需要学生具有较强的编程能力, 但是只能使用系统自带的统计分析方法进行数据处理, 对于需要使用新的统计模型解决问题的情况处理不了, 其他三个软件对编程要求较高, 但可以根据不同的统计模型需求进行编程, 从而可以使用多元统计分析方法实际问题。因此, 在多元统计分析课程教学时可以要求学生掌握 SPSS 软件和其他三种软件中的一种, 以适应不同的数据分析要求。

在课堂上对进行理论知识的讲解过程中, 通过软件的实现, 可以使学生理论联系实际, 更好的提高学生的理论应用能力。在学习软件时, 不仅要掌握该软件的开发环境和编程语法习惯, 还要学会养成注释的好习惯, 以便以后查找使用。在使用软件进行数据分析时, 不只会调用工具箱的函数或者使用从网上搜到的程序, 还要根据具体的需求, 编出更适合的程序, 提高数据分析的质量。通过和老师同学的沟通和交流, 不断地提高编程的效率。在确定掌握一门语言之后, 通过参与老师的项目或者参加数据分析的相关比赛提高学生使用软件和多元统计方法分析和处理多元数据的能力。

3.3. 充分发挥网络资源的优势

教师充分利用网络资源, 对爱课程、中国大学 Mooc 网和腾讯云课堂等在线课程网站上的优质资源, 根据学生的基础进行筛选择优推荐与新课相关视频, 引导学生进行学习。上课前一周在 QQ 群中发放学习资料, 鼓励学生做好课堂预习, 将预习时遇到问题通过 QQ 群反馈给老师, 在课堂上, 老师可以通过引导的形式进行解决, 从而提高课堂效率。在教学中, 充分体现素质教育, 突出教学思想, 贯彻由浅入深、循序渐进、融会贯通的教学原则与直观形象的教学方法。既注重多元统计分析的基础概念、基本理论和方法的阐述, 又注重学生基本运算能力的训练和分析问题、解决问题能力的培养。

在互动区域和学生进行实时的互动, 便于老师及时调整教学方法。讨论过程中, 及时发现学生的问题, 加以适度引导; 当学生遇到问题时, 给予恰当的启发; 引导学生利用网络上的相关训练题, 进行自测性训练, 并对自己的学习活动和成效进行反思, 以改善自己的学习方式方法。通过 QQ 的交流和作业的批改, 不断地进行教学反思, 提高教学质量。

3.4. 推进以赛促学

大数据时代, 数据的多样化也衍生出了很多比赛, 数学建模、泰迪杯数据挖掘比赛和正大杯市场分析大赛等。

有计划地组织学生参加数学建模、泰迪杯数据挖掘比赛和正大杯市场分析大赛等比赛, 并针对性地进行赛前培训。竞赛结束后, 对赛题进行深入分析与创新性研究, 并对优秀竞赛成果进行固化, 形成实践教学资源, 运用到多元统计分析的实践教学中, 通过对赛题的分析和优秀论文的讲解, 使得学生理论联系实际, 学会使用多元统计分析解决实际问题的能力。

通过培训和参加比赛等途径,提升统计专业学生的实践能力和创新创业能力,从而增强学生的职业技能和就业竞争力。通过培训和实训项目的参与,提升统计专业教师的专业技术水平和实践指导能力,从而促进教师自身的发展。通过这些比赛,推进以赛促学、以赛促教、以赛促研,结合高校师生的特点和企业的技术优势和资源进行竞赛指导和后续研发等方面的合作。

大数据时代,给多元统计分析教学带来巨大的机遇和挑战,通过既重视理论教学,又重视实践教学;与机器学习相结合,提高教学的实用性;根据校企合作项目,增强学生处理问题的能力和根据市场需求,提高学生的竞争力等优化课程体系。要求学生掌握 SPSS 软件和 SAS、R 及 Python 三种软件中的一种,以适应不同的数据分析要求来强化统计软件的辅助性。根据教师充分利用网络资源,对爱课程、中国大学 Mooc 网和腾讯云课堂等在线课程网站上的优质资源,根据学生的基础进行筛选择优推荐与新课相关视频,引导学生以充分发挥网络资源的优势。有计划地组织学生参加数学建模、泰迪杯数据挖掘比赛和正大杯市场分析大赛等比赛,推进以赛促学等方法,来促进学生在大数据时代更好的学习多元统计分析课程。增强数据挖掘与大数据分析和多元统计分析的理论教学,加强真实数据分析与应用的实践实训,以培养出更适应市场需求和社会发展的统计学人才。

基金项目

教育部协同育人项目(项目编号:201901104052),岭南师范学院高等教育教学改革项目(项目编号:LSJG1708,LSJGMS1829)。

参考文献

- [1] 刘党社,陈岩.大数据背景下基于 OBE 理念的《多元统计分析》课程教学设计研究[J].豫章师范学院学报,2020,35(2):50-54.
- [2] 朱辉.大数据时代多元统计分析课程教学模式构建[J].统计与管理,2016(12):11-15.
- [3] 张莉,李拂晓.大数据时代下高等教育改革探索[J].科技风,2020(29):24-25.
- [4] 钟云燕.基于培养学生实操能力的实验教学改革探索[J].教育教学论坛,2019(46):106-107.
- [5] 马华,陈振.应用型本科院校程序设计课程实验教学改革研究[J].计算机教育,2011(4):19-22.
- [6] 唐汉卫,张姜坤.大数据教育应用的限度[J].华东师范大学学报(教育科学版),2020,38(10):60-68.
- [7] 吴喜之.多元统计分析——R 与 Python 的实现[M].北京:中国人民大学出版社,2019.
- [8] 任雪松,于秀林.多元统计分析[M].北京:中国统计出版社,2011.