Journal of Education and Educational Research
ISSN: 2957-9465 | Vol. 2, No. 1, 2023

Research on the integrated and innovative teaching mode of probability and statistics courses under the background of big data

Chenmin Ni

School of International Business, Zhejiang Yuexiu University Shaoxing, 312000, Zhejiang Province, China

Abstract: Big data has become a buzzword in many industries, forums and economic exchanges, and its development is interrelated and permeated with mathematics, computers and mathematical statistics courses. Probability theory and statistics, as tools for big data analysis, often belong to different teaching departments, and teaching and research are separated, this paper is based on the Department of Economics and Statistics doing innovative reforms in the discipline foundation, course class connotation and talent training direction to achieve two This paper is based on an innovative reform of the Department of Economics and Statistics in terms of the foundations of the subject, the content of the course and the direction of personnel training. Vertically, realize the theoretical content integration of "high school probability statistics university probability theory and mathematical statistics university statistics big data and statistics"; In the horizontal expansion of the teaching process, grasp the context of "theoretical system interesting cases mathematical modeling ideas project-based embedded teaching competition participation guidance theory and practice review and strengthening"; Form a student-centered and output-oriented integrated and innovative teaching model.

Keywords: Big data; Probability theory; Statistics; Discipline construction; Teaching model.

1. Introduction

In 2015, the State Council officially issued the Action Outline for Promoting the Development of Big Data. In 2016, the country will issue the "13th Five-Year Plan" for Big Data, which marks that China has entered the era of big data. However, the development of big data has brought a certain impact on statistics. Some people even think that statistics will be eliminated and vigorously pursue big data. Comparing the Baidu heat index of statistics and big data (see Figure 1), it can be seen that the heat of statistics and big data was reversed at the end of 2012.

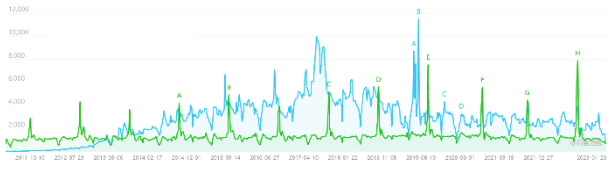


Fig. 1 Comparison of Baidu's popularity trend between big data and statistics

At the end of 2008, "big data" was recognized by some well-known computer researchers in the United States. In the white paper "Big Data Computing: Creating a Revolutionary Breakthrough in the Business, Science and Social Fields" published by the Computing Community Alliance, it was proposed that big data is really important for new purposes and new insights, not data itself. In March 2012, the Obama administration announced that it would invest US $200 million to promote the development of big data-related industries and upgrade the "big data strategy" into a national strategy. In 2013, the publication of a book titled "The Age of Big Data" by British scholar Victor Mayer-Schoenberg aroused the attention of many experts, scholars and economic successful people to big data; In May 2013, Ma Yun, Chairman of the Board of Directors of Alibaba Group, said in his speech at the 10th anniversary party of Taobao that the era of big data has come. The future era will not be the era of IT, but the era of DT. In 2015, the State Council officially issued the Action Outline for Promoting the Development of Big Data. In 2016, the State solicited expert opinions on the "13th Five-Year Plan" for Big Data. At present, the world has entered the era of big data. Big data has become a hot word in many industries, forums and economic exchanges. Its development and mathematics, computer, mathematical statistics, management and sales are mutually penetrating, Various teaching reform research and application research based on the background of big data continue to emerge, and the related teaching research on big data and statistics, statistics and probability theory has also received much attention [10].

By analyzing the teaching papers or research reports of these scholars, it is found that the content generally focuses on the single-category curriculum research of probability [11] or statistics [12], while few scholars integrate the two and propose new teaching models; Moreover, there are few discussions on the integration of the theoretical framework of statistics and the application of big data, and there is still a lack of sustainable teaching and interdisciplinary research. On the basis of previous research, this project will make innovation and reform in the discipline foundation, curriculum connotation and talent training direction, make up for the shortcomings of previous research, and achieve two integration, namely, the integration of student-centered probability and statistics teaching, and the integration of output-oriented statistical methods and big data applications, and explore the integration and innovation teaching mode of probability and statistics courses driven by the big data era.

2. Research path

2.1. Preliminary reference of teaching materials

As the first course of statistics courses, probability courses are the compulsory courses for postgraduate entrance examination in many schools. The national planning textbook "Probability Theory and Mathematical Statistics", edited by Sheng Zhouu of Zhejiang University and Wu Chuansheng of Wuhan University of Technology, is the national classic textbook of arts and science; Statistics courses are specialized courses for colleges and universities in economics and management, finance, medicine, biology, education and other majors. The main teaching materials are the national high-quality course textbook Statistics compiled by Chen Zhenzhen of Xiamen University, Business Statistics by David Levin and various professional statistics.

2.2. Research methods

It mainly carries out qualitative and quantitative research through investigation and research method, literature analysis method, analytic hierarchy process, comparative research method, interdisciplinary research method, case analysis method, experience summary method, practical teaching method, chart method, etc., and draws more scientific conclusions.

The investigation and comparative research methods are mainly used for the investigation and comparative analysis of the teaching and research situation of other universities and colleges; The literature retrieval method and the analytic hierarchy process are used to study and analyze the works of experts and scholars at home and abroad, and find out the advantages and disadvantages; The interdisciplinary research method, case analysis method, experience summary method and practical teaching method are mainly used to carry out various teaching activities and summarize the results of activities in practical teaching; The chart method is mainly used for statistical description of survey results and data display in paper writing.

2.3. Research theory and ideas

Based on the era background of big data, using modern education concepts, following the requirements of the new National Standards, and taking the student-centered theory of sustainability with both output-oriented and knowledge learning as the main line; Taking the probability and statistics courses of the financial and economic management students of our school as the main research object, combining with the survey of at least 30 schools (divided into private and public schools) in and outside the province. At the same time, consult various relevant documents at home and abroad, collect the status of teachers, students' learning and teaching materials in various colleges and universities, and grasp the main teaching content and teaching characteristics of probability theory, mathematical statistics and statistics; Analyze the relationship between statistics and big data in the era of big data, find the fusion point of statistics theory and big data application, summarize and sort out, analyze and sublimate, find problems, summarize laws, and give a teaching mode of integration and innovation.

3. Integrating and innovating teaching mode of probability and statistics courses

The research focuses on: Analysis and realization of knowledge fusion, teaching fusion and application fusion of probability and statistics courses; Find the expansion point of the connection between statistics and big data, and analyze the theory and application from point to face; Highlight innovation points and form a relatively complete teaching model.

The difficulties are as follows: the integration of textbook knowledge needs new matching textbooks; The discussion of teachers' integrated teaching needs all departments to jointly mobilize the enthusiasm of teachers; Training and improvement of teachers' teaching ability. The research framework is as follows.

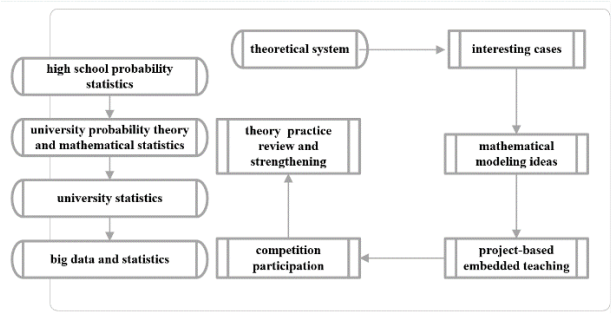


Fig.2 Research framework

Vertically, realize the theoretical content integration of "high school probability statistics → university probability theory and mathematical statistics → university statistics → big data and statistics"; In the horizontal expansion of the teaching process, grasp the context of "theoretical system → interesting cases → mathematical modeling ideas → project-based embedded teaching → competition participation guidance → theory and practice review and strengthening". The specific content is reflected in the following three aspects.

3.1. Student-centered, to achieve the integration of probability courses and statistics courses.

(1) After the new curriculum reform in senior high school, some knowledge points of university probability and statistics have been added to the mathematics content in senior high school, but at present, the university probability and statistics textbooks are basically still using the traditional textbooks, and the content has not been updated in time, resulting in repeated teaching of high school knowledge and university knowledge. This topic will study the knowledge points of senior high school textbooks and university textbooks, and give the new key knowledge points of university probability and statistics teaching.

(2) At present, most university probability textbooks are named Probability Theory and Mathematical Statistics. The content arrangement is usually the first five chapters of probability theory and the last six to nine chapters of mathematical statistics; There are many names of statistical textbooks, most of which are named after Statistics, Business Statistics and other disciplines of statistics. The contents often involve four or more chapters of probability theory, followed

by several chapters of statistics and simple applications. Because of the characteristics of the arrangement of the textbooks of the two courses, the teaching content will be repeated in many cases, In addition, it is impossible to teach students the content that truly reflects the application under the limited class hours. This project will analyze the characteristics of many probability and statistics textbooks in universities, integrate teaching resources, and realize the organic combination of theory and application.

(3) Many probability courses in colleges and universities are taught by math teachers, while statistics courses are taught by economic and management teachers. The lack of communication and handover between teachers results in the inconsistency of knowledge points when teaching the same group of students. And different departments have different requirements for probability theory and statistics. Especially with the arrival of the era of big data and the change of talent cultivation mode in the international perspective, the same teaching mode can no longer meet the needs of different departments. This topic will propose a solution to the problem of teachers' poor communication and different needs of knowledge points under different conditions of college and department training programs, so as to achieve student-centered teacher integration and knowledge point integration.

3.2. Driven by the era of big data, the research on the correlation between probability and statistics courses and big data applications.

(1) Probability courses should have a solid theoretical foundation and smooth application expansion. At present, probability courses in many colleges and universities lack experimental hours. After the knowledge fusion of high school probability statistics and university probability statistics courses is completed, the theoretical class hours of university introduction statistics will be compressed, and teachers will have time to apply case teaching methods to expand the scope of the classroom. Experimental class hours can be appropriately set to enhance students' grasp of knowledge application and achieve a smooth transition between probability theory and statistics.

(2) Integration and innovation of statistical courses and big data-related content. Both statistics and big data aim at solving problems and are based on mathematical models (such as linear regression, logistic regression, clustering, time series, principal component analysis, etc.), but their thinking methods and processes of knowledge discovery and data processing are different. Big data is a "algorithmic" approach, which focuses on computer software implementation, while statistics is a "scientific" approach, which focuses on mathematical methods. Improving statistical methods, expanding the richness of big data processing and practical application problems with statistical methods, integrating the advantages of the two, and learning from each other's strengths and weaknesses, is a key research content of the project.

3.3. Take output as the guidance, and carry out innovation and reform on traditional teaching methods in the teaching process

(1) Consolidate the foundation, integrate the idea of mathematical modeling into classroom teaching, and improve students' ability to analyze and solve problems. Driven by the era of big data, the application of Internet plus is becoming more and more extensive, and the demand for talents is becoming more and more "productive". Cultivating students' ability to solve problems by hand while mastering theoretical knowledge is the focus of talent training in colleges and universities. Mathematical modeling solves practical problems through the main steps of "problem analysis - model hypothesis - model establishment and solution - model evaluation and test", and inserts modeling cases into probability and statistics courses, which is conducive to improving students' ability to analyze and solve problems.

(2) Realize the sustainable application of knowledge by participating in competition activities. Integrate big data-related examples into the statistics teaching process, combine the "Teddy Cup" big data competition and the "statistical survey competition" in summer, and strengthen students' understanding and grasp of knowledge through project-based grouping work; Guide and encourage students to participate in mathematical modeling competitions, and expand students' ability to apply knowledge through various competitions.

(3) The project-based teaching method is applied to teaching, and the Internet plus platform is used to enhance students' interest in learning. Take the project-based questions as the main points of the assessment at ordinary times, and try to change the traditional assessment method; Timely use cloud ink class for classroom teaching and knowledge consolidation after class to achieve sustainable teaching effect; Integrate hot issues into teaching and improve students' learning interest and learning effect.

4. Results and Discussion

(1) The project meets the requirements of the National Standard, student-centered, and researches the integration of interdisciplinary disciplines, which is exploratory and creative. In January 2018, the Ministry of Education issued the National Standard for the Teaching Quality of Undergraduate and Professional Courses in General Institutions of Higher Learning, or the National Standard for short. The National Standard puts forward three basic requirements for talent training in colleges and universities, namely, focusing on students, output-oriented and continuous improvement. Probability is a subject of mathematics department, while statistics is a professional course of economics and management. According to the recent survey of several colleges and universities and a lot of literature, many college probability courses are taught by mathematics teachers, while statistics courses are taught by economics and management teachers. There is a lack of communication and handover between teachers, resulting in the inconsistency of knowledge points when facing the same group of students teaching; In addition, the teaching materials of probability and statistics courses have a lot of content duplication, and there is also a phenomenon of repeated teaching with the content of high school mathematics. Under the condition that the total class hours of each university are limited, it is a meaningful research topic to study student-centered, integrate teacher resources, integrate teaching content, and innovate teaching models.

(2) The research results can be popularized and used for reference. Through investigating the teaching situation of probability and statistics courses in at least 30 universities and combining the teaching situation and practice of relevant majors in the university, the project found the shortcomings of traditional teaching of probability and statistics courses, explored the integration and innovative teaching mode of

probability and statistics, statistics courses and big data applications under the background of the big data era, summarized and refined the practice and experience of practical teaching, and constructed a teaching model with the characteristics of the university The teaching mode of probability and statistics courses, which can be used for reference and applied by other colleges and universities, can be popularized and used for reference.

References

[1] Steve Lohr，The Age of Big Data［N］. New York Times，2012. 02. 11.

[2] Fan J，Han F，Liu H， Challeges of Big Data Analysis［J］. The Annals of Applied Statistics，2013(7) : 1308-1479.

[3] Zhao Bo, Research on Strategic Foresight Based on Big Data [D] CPC Central Party School, July 2016.

[4] Zhu Huaiqing. The impact of the big data era on undergraduate economics and management statistics teaching and countermeasures [J] Higher Education Research, 2014,31 (3): 35-37.

[5] Geng Zhi, opportunities and challenges faced by statistics in the era of big data [J]. Statistical Research, 2014, 31 (1): 5-9

[6] Chen Shiming, Research on the Impact of Big Data on Financial Statistics [D] Hunan University, April 2015.

[7] Shi Manhong, Probabilistic Theory and Mathematical Statistics Teaching Method Improvement [J] Journal of Chifeng University (Natural Science Edition), 2016.32 (8): 249-251.

[8] Li Xiuzhen, Ma Xiao, Zuo Jiabin, Research on Teaching Reform of Probability Theory and Mathematical Statistics in the Context of Transformation [J], Journal of Jilin Normal University of Engineering and Technology, 2017.33 (3): 84-86.

[9] Jiang Shiquan, Liu Zhongxia, Investigation and experimental research on the teaching reform of management statistics from the perspective of big data [J] Journal of Tongling University, 2017.2:114-117.

[10] Zhao Jianchao, Meier Schoenberg and Kukyer's Research on Big Data Education Thought [J] Contemporary Education Science, July 11, 2017.

[11] Xu Xuhua, Zhao Chunyan, Li Ling, Wang Yan, Zhong Qin, Chen Xiangbing, Li Liang. Teaching reform and practical exploration of probability theory and mathematical statistics based on curriculum ideological and political philosophy [J]. Journal of Xinjiang Normal University (Natural Science Edition), 2022,41 (04): 81-86. DOI: 10.14100/j.cnki. 1008-9659.2022.04.005.

[12] Liu Jinpei, Chen Lijuan, Jin Feifei, Chen Yi, Wu Peng. Exploration and practice of national-level first-class curriculum construction under the background of big data -- taking the "statistics" course of Anhui University as an example [J]. Education and Teaching Forum, 2022 (52): 92-95.