Education Reform and Development, 2025, Volume 7, Issue 1
https://ojs.bbwpublisher.com/index.php/ERD
Online ISSN: 2652-5372
Print ISSN: 2652-5364

Ideological and Political Exploration for Multivariate Statistical Analysis Methods

Jianxi Zhao*

School of Applied Science, Beijing Information Science & Technology University, Beijing 102206, China

***Corresponding author:** Jianxi Zhao, zhaojixiwww@163.com

Abstract: Statistics is an important component of data science, and multivariate statistical analysis methods play a crucial role in analyzing high-dimensional data and extracting effective information from it in the era of big data. Under the wave of domestic curriculum ideological and political education, this paper explores the organic integrations of multivariate statistical analysis methods such as regression analysis, variance analysis, factor analysis, principal component analysis, canonical correlation analysis, k-nearest neighbor algorithm, and hierarchical clustering with ideological and political education. The study makes analogies between the principles, ideas, advantages, and applications of these methods and some ideological and political contents, making abstract multivariate statistical analysis methods more vivid and helping students establish correct worldviews, outlooks on life, and values.

Keywords: Multivariate statistical analysis; Statistical analysis methods; Ideological and political education

Online publication: February 12, 2025

1. Introduction

In the era of big data, statistical analysis methods can extract valuable information and discover hidden patterns from massive, complex, and diverse data, and make reasonable predictions and assist decision-makings. Multivariate statistical analysis is able to analyze high-dimensional data and extract effective information from them. The theories and methods have wide applications in computer vision, pattern recognition, natural language processing, engineering information, biomedicine, management decision-making, economics and finance, and so on fields, and have become important tools to drive scientific innovation and promote social development.

Statistical Analysis Methods is a core professional course of the major Data Science and Big Data Technology of the university. It is an extension and deepening of the Probability Theory and Mathematical Statistics course in dealing with high-dimensional complex practical problems and is the theoretical foundation for students to learn subsequent professional courses. With the development of the Engineering Education Professional Certification, the university has put forward clearer requirements for mathematics courses and clearer quantifiable specific indicator requirements for the Statistical Analysis Methods course. This course is for

educating students to master commonly used multivariate statistical analysis methods and is an important carrier for cultivating their rational thinking. The teaching of this course will strengthen their grasp of basic theories and methods of statistics, cultivate their abstract thinking ability, logical reasoning ability, and ability to solve complex practical problems, and lay a necessary theoretical foundation for their learning of subsequent courses and further expanding their mathematical knowledge.

In recent years, domestic various universities have been actively promoting curriculum ideological and political education, including popular statistics courses such as Probability Theory and Mathematical Statistics [1], Statistics [2], Recommendation Systems [3,4], Regression Analysis [5], Applied Regression Analysis [6], Financial Statistics [7], Time Series Analysis [8], Medical Statistics [9,10], Statistical Calculations [11], Applied Statistical Software [12,13], Statistical Method Selections and SPSS Implementations [14], and Multivariate Statistical Analysis [15]. The main content of the course Statistical Analysis Methods of our university is multivariate statistical analysis, and similar courses have already been offered in many universities. Statistical analysis methods keep up with the trends of the times, have a wide range of applications, and are highly welcomed by students. However, so far, there have been very few studies on the ideological and political aspects of multivariate statistical analysis methods. This paper makes analogies between the principles, ideas, advantages, and applications of these methods and some ideological and political contents, making abstract multivariate statistical analysis methods more vivid, helping students intuitively understand or master the principles, ideas, or advantages of these methods, stimulating their ideological and political sentiments, and helping them establish correct worldviews, life, and moral values.

2. Methods of ideological and political education

This section integrates some ideological and political contents related to worldviews, life, and moral values into main multivariate statistical analysis methods such as regression analysis, variance analysis, factor analysis, principal component analysis, canonical correlation analysis, k-nearest neighbor algorithm that belongs to classification analysis, and hierarchical clustering that belongs to clustering analysis. The principles, ideas, advantages, and applications of these methods are analogized to some ideological and political contents.

Regression analysis establishes a regression model based on given data points, determines an appropriate line (a regression equation) to fit the points, and uses the regression equation to screen factors, predict results, and control adjustments. In fact, a regression model describes a phenomenon, while a regression equation reflects a law. Thus, the process of regression analysis is the process of exploring the inherent law of the phenomenon, and then the law is used to serve humanity. Students can be inspired by regression analysis: To learn to analyze the essence of a phenomenon. In the face of various complex practical problems, it is necessary to have the ability to deeply analyze problems and effectively extract their essences. This ability will help people better understand the world and lay the foundation for future development and innovation.

The basic idea of variance analysis is to decompose the total change of the dependent variable into the influence of each factor or interaction and the influence of random error, and by comparing the relative size of the two, it is determined whether the influence of each factor or interaction on the dependent variable is significant. One of the key points of variance analysis is to understand the concept of “relative,” which is that each same factor or interaction may have different impacts in different environments or conditions. This understanding of “relative” reflects the importance of dialectical thinking. Dialectical thinking emphasizes that things are constantly evolving and changing, and should be examined in corresponding stages, environments, or conditions.

When solving problems and making decisions, it is important to consider the specific environments and current restrictive conditions and to frequently adjust and optimize methods and strategies to adapt to constantly changing environments and conditions.

Factor analysis refers to a method of using a few factors to describe the correlation among numerous indicators. Its basic idea is to divide observed indicators so that the correlation among indicators in each same group is high, while the correlation among indicators in different groups is low. All indicators in each group reflect a common factor. Each observed indicator can be described by the sum of a linear combination of the least unobservable common factors and a special factor. In the process of ideological and political education, it is necessary to extract core concepts of numerous social phenomena, for example, socialist core values, which is similar to the process of extracting main factors in factor analysis. Ideological and political education helps students understand and grasp deep meanings of these core concepts and cultivates them to become responsible, moral-minded, and innovative social citizens.

Principal component analysis is a linear transformation of a set of numerous correlated original indicators into a set of uncorrelated new comprehensive indicators, and selects the first few comprehensive indicators that contain most information, achieving the goal of significantly reducing spatial dimensions and removing correlations between indicators. Principal component analysis transforms and selects complex high-dimensional data into simplified low-dimensional data, synthesizes and extracts the main information from original data, and captures the main contradictions of problems. In the process of solving problems, it is important to avoid the interference of unimportant factors as much as possible, making decisions and actions timelier and more targeted, thereby breaking through key points of problems and improving efficiency and quality of work.

Canonical correlation analysis focuses on identifying and quantifying the correlation between two sets of random variables, which is an extension of the correlation between two random variables. The main indicators for measuring economic growth include per capita gross domestic product, consumer price index, industrial production index, employment rate, foreign exchange reserves, per capita income, and so forth. The main indicators for measuring environmental pollution include the annual average concentration of inhalable particulate matter, annual average concentration of fine particulate matter, annual average concentration of sulfur dioxide, annual average concentration of nitrogen dioxide, chemical oxygen demand emissions, sulfur dioxide emissions, and so forth. The study uses canonical correlation analysis to explore the correlation between these two sets of indicators, thereby obtaining the correlation between economic growth and environmental pollution, to better understand the balance between economic development and environmental protection, and promote green and sustainable development of society by taking appropriate macroeconomic control measures.

In a feature space, the k (k is a positive integer) points closest to a point in a certain class are searched and determined. Among these k points, if the number of the points belonging to a certain class (e.g. A) is the highest, the point in a certain class to be determined will be classified into class A. This classification method is called the k-nearest neighbor algorithm. This algorithm inspires students that the behaviors and decision-makings of a person are influenced and constrained by their surrounding environments. It reminds people to pay attention to social and cultural factors behind behaviors and decision-makings. Party and government agencies and social organizations need to strengthen guidance in education, culture, and so on, build positive and upward social environments, create harmonious and harmonious social atmospheres, promote social harmony and stability, and comprehensive development of people.

Hierarchical clustering is also known as lineage clustering or systematic clustering. Firstly, each point is considered as a class individually. Secondly, the two closest classes are aggregated into a new class. Thirdly, the

two closest classes are also aggregated into a new class, and so on. As the distance between classes increases, all classes are finally aggregated into one large class, forming a lineage graph aggregated by distances from near to far. The appropriate number of classes can be determined based on specific situations and actual problems. The relationship between points in each same class reflects the similarity of points, while the relationship between points in different classes reflects the diversity of points. Human society is the coexistence and integration of diverse cultures. By promoting communication, tolerance, and integration between different cultural, ethnic, religious, and so on groups, all nations can promote social harmony and stability.

3. Conclusion

This paper explores ideological and political elements in multivariate statistical analysis methods, delves into the organic integrations of main methods and ideological and political education, and makes analogies between the principles, ideas, advantages, and applications of these methods and some ideological and political contents, such as regression analysis (analyzing the essence of a phenomenon), variance analysis (demonstrating the importance of dialectical thinking), factor analysis (extracting core concepts of numerous social phenomena, for example, socialist core values), principal component analysis (capturing main contradictions of problems), canonical correlation analysis (promoting green and sustainable development of society), k-nearest neighbor algorithm (building positive and upward social environments), hierarchical clustering (promoting communication, tolerance, and integration between different cultural, ethnic, religious, and so on groups). The rhetorical device of analogy visualizes these abstract multivariate statistical analysis methods, deepening students' understanding or mastery of the principles, ideas, or advantages of these methods. Integrating ideological and political elements into these methods can help to edify students' ideological and political sentiments, and cultivate their sense of social responsibility and civic consciousness.

Funding

2025 university-level curriculum ideological and political education reform project of Beijing Information Science & Technology University: "Research on the Integration of Science and Education and that of Ideological and Political Education about Statistical Analysis Methods in the Era of Digital Intelligence" (Project No.: 2025JGSZ24); Young Backbone Teacher Support Plan of Beijing Information Science & Technology University (Project No.: YBT 202445)

Disclosure statement

The author declares no conflict of interest.

References

[1] Li Q, 2024, Exploring Ideological and Political Elements in the Course of Probability Theory and Mathematical Statistics. Defence Industry Conversion in China, 2024(2): 155–156.

[2] Xiong W, Pan C, 2024, Exploration and Practice of Statistics Curriculum Ideological and Political Education Reform of Agriculture and Forestry Universities Under the Background of New Agricultural Sciences. South-

Central Agricultural Science and Technology, 45(2): 225–228.

[3] Zhao J, 2023, Ideological and Political Exploration for Teaching Contents of Recommendation System Course. Modern Management Forum, 7(12): 230–232.

[4] Zhao J, 2024, Ideological and Political Exploration for Some Recommendation Methods. Modern Management Forum, 8(4): 122–124.

[5] Wei Q, He L, Ji J, et al., 2024, Teaching Research on Integrating Regression Analysis into Curriculum Ideological and Political Education. Intelligence, 2024(15): 21–24.

[6] Zeng S, Wu H, Yu Y, 2024, Classroom Teaching Exploration of Applied Regression Analysis Based on Curriculum Ideological and Political Education. Industry and Technology Forum, 23(6): 196–199.

[7] Xu Z, Ma D, 2024, Research on Ideological and Political Teaching Strategies of Statistics Courses in Colleges and Universities—Taking Financial Statistics as an Example. Huazhang, 2024(5): 27–29.

[8] Wang Y, Xia W, 2024, Research on Mining and Teaching Implementation of Ideological and Political Elements in the Graduate Course Time Series Analysis. Journal of Higher Education, 10(12): 184–188 + 192.

[9] Li J, Zhu J, Wu S, et al., 2024, Practical Research on Teaching Reform of Curriculum Ideological and Political Education in Medical Statistics. Chinese Journal of Health Statistics, 41(2): 302–305.

[10] Zhou Z, Wu W, Chen H, 2023, Teaching Design Exploration of Curriculum Ideological and Political Education in Graduate Medical Statistics. Medical Education Management, 9(6): 706–710.

[11] Lu J, Zhang Z, 2024, The Challenges and Countermeasures Faced by Curriculum Ideological and Political Education in Undergraduate Statistical Calculations in the Digital Era. Journal of Higher Education, 10(11): 186–189.

[12] Xu L, Zhou Y, 2024, Designs and Applications of Curriculum Ideological and Political Cases in Applied Statistical Software. Computer Knowledge and Technology, 20(5): 159–161.

[13] Xu L, Zhou Y, Zhang H, 2024, Exploration of Curriculum Ideological and Political Education Under the OBE Concept—Taking the Course Applied Statistical Software as an Example. Education Informatization Forum, 2024(3): 117–119.

[14] Fu Y, He J, 2024, Statistical Method Selections and SPSS Implementations for Exploring Connotations of Ideological and Political Education. Education Forum, 2024(5): 165–168.

[15] Xin H, Zhou S, Fu X, et al., 2024, Elementary Analysis for Curriculum Reforms and Ideological and Political Designs of Multivariate Statistical Analysis. Chengcai, 2024(4): 43–45.