Current Situation of Application and Development Prospects of the Statistical Analysis of Big Data

Zhuoran LI*

Q-Square Business Intelligence, Corp., Tianjin 300000, China

Abstract With the advent of the big data era, modern statistics has enjoyed unprecedented development opportunities and also faced numerous new challenges. Traditional statistical computing methods are often limited by issues such as computer memory capacity and distributed storage of data across different locations, and are unable to directly apply to large-scale data sets. Therefore, in the context of big data, designing efficient and theoretically guaranteed statistical learning and inference algorithms has become a key issue that the current field of statistics urgently needs to address. In this paper, the application status of statistical analysis methods in the big data environment was systematically reviewed, and its future development directions were analyzed to provide reference and support for the further development of theory and methods of the statistical analysis of big data.

Key words Big data; Statistical analysis; Current status; Development prospects

DOI 10.19547/j.issn2152 -3940.2026.01.008

With the rapid development of information technology, the barriers for data collection and storage have significantly decreased, and large-scale data sets are constantly emerging in various industries. In scenarios such as online social media, monitoring systems, online search engines, online shopping platforms, and astronomical observatories, massive amounts of data are generated every moment. The massive data emerging in various fields not only bring unprecedented opportunities for the development of statistics but also pose severe challenges to traditional statistical methods. In this paper, the application of statistical analysis methods of big data was reviewed and summarized, and their future development trends were analyzed to provide theoretical reference for the further development of statistical analysis of big data.

1 Application of statistical analysis of big data

The statistical analysis of big data has wide applicability across various industries, and the application in economic management, urban planning and construction, and enterprise auditing is particularly prominent. It can provide crucial data support for related decision-making and management work.

1.1 Application in the medical field The statistical analysis of big data has been widely applied in clinical medicine and various aspects of medical health. For instance, Cui Lizhen *et al.* [1] have constructed a cutting-edge technology system for multi-source heterogeneous medical big data, which can promote the upgrade of medical health data to achieve multimodality, knowledge graphization and explainability, fully releasing its value and efficiency as a national strategic resource. Other scholars believe that by leveraging big data mining and analysis, the incidence patterns of gynecological diseases can be revealed from multiple dimensions, providing important basis for the prevention and treatment of gynecological diseases in China.

1.2 Application in the field of economic management Many scholars have explored the application value of the statistical analysis of big data in the agricultural and rural sectors from various perspectives. For example, Zhang Lu[2] pointed out that big data can analyze characteristics of historical data, and predict the trends of crop yield and price, providing support for rural production management. Yin Yanfang[3] believes that regression analysis can construct mathematical models and quantify influencing factors, helping to assess agricultural production efficiency. Li Yingshuo[4] proposed that social network analysis can explore consumer relationships and communication patterns, and optimize precise marketing of agricultural products. Lin Jing[5] stated that big data can improve the rural credit scoring system, enhance the accuracy of credit risk assessment and simplify the approval process. Ren Xiaoyu[6] argued that big data can achieve differentiated and refined pricing for rural insurance, and formulate premium strategies based on regional and crop risk characteristics. Overall, big data has significant application prospects in agricultural production, efficiency assessment, marketing, rural credit and insurance.

Some scholars also believe that by using statistical analysis methods of big data, cigarette marketing can be effectively transformed towards precision and digitalization, achieving precise marketing, precise assessment of market dynamics, precise cultivation of cigarette brands, promoting digital terminal construction, and strengthening the improvement of marketing team capability. Through big data technology, it is possible to deeply understand the characteristics of consumer demands and the patterns of consumption behaviors, optimize resource allocation and marketing strategies, and thereby significantly enhance the operational efficiency and economic benefits of tobacco enterprises.

Received: December 15, 2025 Accepted: January 23, 2026

* Corresponding author.

1.3 Application in urban planning and construction Many scholars have conducted diverse research on the prediction of urban spatial development and the application of big data in urban planning. Ma Changfeng *et al.* [7] focused on the traditional prediction system of urban spatial scale, and used various technical methods such as the comprehensive method and natural growth rate method to predict the trend of urban population growth and change, thereby clarifying the future development goals of the city. Li Min[8] pointed out that in the era of big data, the research on urban development needs to rely on big data statistical analysis to accurately judge the main trends of urban development, and also takes into account the internal connections between cities and the core development needs of the city itself. Li Jinhua[9] further refined the research path, and proposed that based on the statistical analysis of big data, multi-dimensional data such as the page structure of professional websites and network domain names can be used to simulate the interaction contact methods between cities, accurately grasp the development orientations of the three main entities – residents, the government, and enterprises, thereby clarifying the future development strategy of urban space.

1.4 Application in enterprise auditing The continuous advancement of mobile information technology has driven the application of processing technology of big data in the statistical analysis of auditing data, and has become an inevitable trend in the development of auditing field.

For instance, Zhou Xumei *et al.* [10] believe that data statistical analysis has expanded the development space for enterprise auditing work and provided more diverse analytical perspectives for solving problems such as the operation management and financial control of enterprises. At the internal control level, it can effectively restrict authorized access behaviors and ensure system access security. At the business level, its application focuses on the analysis of exceptional matters and rule-based management of business activities. Wei Sunyuan[11] think that the statistical analysis of auditing information in the big data environment helps improve the authenticity of information and the timeliness of auditing business, promotes the professionalization of the auditing team, but at the same time, attention should be paid to the prevention of potential violation risks. Zhang Chunyu *et al.* [12] believe that the application of big data technology can significantly improve the quality and efficiency of auditing work. This study further analyzed the specific application of distributed data processing, data mining, web crawlers, and 3S technology in auditing statistical analysis, and pointed out that the statistical analysis of big data will continue to promote the transformation of auditing work towards informatization and modernization. Under the existing statistical environment and technical conditions, real-time network analysis based on big data analysis is gradually becoming the mainstream method of data analysis and will be widely applied in the field of urban planning and construction.

1.5 Application in the field of environmental protection In the era of big data, environmental decision-making is shifting from being driven by experience to being driven by data. Through the mining of multi-source information and correlation analysis, the inherent laws among environmental elements can be revealed to provide a basis for scientific decision-making in ecological environment protection.

Li Lianshou[13] found that the application of the statistical analysis of big data to the field of environmental supervision and law enforcement can effectively improve the accuracy and efficiency of law enforcement monitoring, strengthen the scientificity and timeliness of environmental supervision, and play an important supporting role in promoting the high-quality development of environmental protection work. Lin Xudong *et al.* [14] believe that the decision-making model in the era of big data is transforming from being driven by experience to being driven by data. Through in-depth mining and correlation analysis of multi-source heterogeneous environmental information, the internal mechanisms and evolution laws of environmental elements can be revealed to provide solid data foundation and technical guarantee for scientific decision-making and scientific governance in the field of ecological environment protection.

2 Problems in the application of the statistical analysis of big data

The statistical analysis of big data has been widely applied in various fields. However, as the digitalization process continues to advance, many problems in its practical application have become increasingly prominent.

2.1 The phenomenon of information silos is widespread Currently, the cross-departmental and cross-industry data sharing and coordination mechanisms still have significant shortcomings. The data barriers and information silos between departments are particularly prominent, and the level of data resource planning, coordinated sharing, and business linkage is insufficient. The openness and sharing level of high-value public data, government information resources, and commercial data are overall low. The key links such as data ownership, sharing exchange, and secure use lack unified norms and institutional guarantees, so that data resources are difficult to achieve efficient circulation, deep integration and optimal allocation, which seriously restricts the market-oriented allocation of data elements and the full release of data value.

2.2 The understanding of the development laws of big data industry is insufficient The society has not yet formed a scientific, objective and unified understanding of the development logic, operation mechanism and value realization path of big data industry. During the process of promoting the construction of big data centers and digital infrastructure, there are obvious issues of path dependence and investment imbalance in some regions. They overly focus on the construction of tangible facilities while neglecting the support system of soft power, resulting in problems such as low integration of data resources, disconnection between technical supply and actual industrial demands, and difficulties in achieving the deep integration and efficient connection of big data

technology, data elements and the real economy, and industry application, which has seriously hindered the high-quality and sustainable development of big data industry to a certain extent.

2.3 The construction and application level of data resources is relatively low The relevant entities have insufficient attention to the construction of data resource systems. Most institutions merely stop at the simple storage stage of data, lacking data preprocessing, integration, and standardization work for practical application. The uneven quality of data resources, the absence of a standard system, and weak management capability have made it difficult to deeply explore and efficiently utilize the value of data.

2.4 The information security and data management system is not yet complete There are deficiencies in relevant laws and regulations regarding data ownership, privacy protection, as well as information security standards and norms. The technical security prevention and management capability is relatively weak. A data openness, management and information security guarantee system that takes into account both security and development has not yet been formed, which has restricted the in-depth application and development of the statistical analysis of big data to a certain extent.

3 Future prospects of the application of the statistical analysis of big data

In the context of the rapid development of digital economy and information technology, the statistical analysis of big data has deeply permeated into social production, life services, and various industry practices, becoming an important foundation for promoting industrial transformation and upgrading, optimizing resource allocation efficiency, supporting scientific decision-making, and facilitating technological innovation breakthroughs. With the continuous enrichment of collection methods of multi-source data, the exponential expansion of data scale, and the continuous improvement of data type and structure complexity, the shortcomings of traditional statistical analysis methods in handling massive data, extracting high-dimensional features, conducting dynamic real-time analysis, and verifying the reliability of results have become increasingly prominent, making them unable to meet the practical needs of data-driven decision-making by high-quality development in the new era.

Currently, the theoretical framework, technical framework and application paradigm of the statistical analysis of big data are still in the stage of continuous improvement and dynamic evolution. There is still considerable room for improvement in the construction of basic theories, the development of core algorithms, the realization of efficient computing, and the implementation of multi-scenario application. At the same time, problems such as uneven data quality, weak information security guarantee system, and lack of standard specifications also restrict the standardized, standardized and large-scale development of the statistical analysis of big data.

Therefore, based on the existing methods and technical systems, the development of the statistical analysis of big data should break free from the constraints of traditional research paradigms, strengthen theoretical innovation and the integration of multi-disciplinary technologies, and construct algorithm models and analysis frameworks that are efficient, interpretable and rigorous in theory. By improving the technical standard system, enhancing data governance, and strengthening information security guarantee, it is necessary to continuously promote the development of the statistical analysis of big data towards a more standardized, standardized and intelligent direction, provide strong support for the digital transformation of various industries, the modernization of social governance and scientific research innovation, and promote the deep application and high-quality development of the statistical analysis of big data in a wider range of fields.

References

[1] CUI LZ, SUN XF, LIU N, *et al*. Multi-source heterogeneous medical big data fusion and analysis technology[J]. Journal of Shandong University (Engineering Science), 2026, 56(1): 1-13, 25.

[2] ZHANG L. Application of statistical analysis methods of big data in economic management[J]. Shanghai Business, 2023(12): 36-38.

[3] YIN YF. Discussion on the application of statistical analysis methods of big data in economic management[J]. Business Culture, 2022(7): 32-33.

[4] LI YS. Research on the application of big data statistical analysis methods in the field of economic management[J]. Trade Fair Economy, 2021(20): 125-127.

[5] LIN J. Application of statistical analysis methods of big data in economic management[J]. China Collective Economy, 2020(34): 163-164.

[6] REN XY. Application of statistical analysis methods of big data in the field of economic management[J]. Quality & Market, 2020(22): 100-101.

[7] MA CF, WU WB. Analysis of spatial planning methods for smart cities in the era of big data[J]. Create Living, 2021(3): 32-33.

[8] LI M. Spatial planning methods of smart cities in the era of big data[J]. Theoretical Research in Urban Construction, 2018(10): 181, 140.

[9] LI JH, WU W. Discussion on spatial planning methods for smart cities[J]. Ability and Wisdom, 2017(32): 215.

[10] ZHOU XM, CHU EM. Institutional heterogeneity, supervising distance and efficiency of internal audit: Based on an empirical analysis of audit data of Hunan education institutions[J]. Journal of Xiangtan University: Philosophy And Social Sciences, 2020(1): 93-97.

[11] WEI SY. Research on statistical analysis of enterprise audit data in the context of big data[J]. Market Modernization, 2022(16): 132-134.

[12] ZHANG CY, YANG YH. Review of big data audit literature[J]. Co-Operative Economy & Science, 2021(19): 141-143.

[13] LI LX. Research on precise law enforcement strategy of environmental supervision based on big data analysis[J]. Leather Manufacture and Environmental Technology, 2025, 6(17): 37-39.

[14] LIN XD, ZHOU JF, LIU J. Application of resource correlation big data analysis in agricultural ecological environment protection[J]. Chinese Journal of Agricultural Resources and Regional Planning, 2016, 37(2): 62-65.