



Applications Research of Statistical Analysis in Business Management in Big Data Era

Yikun Zhang^{1,*}

¹College of Letters and Science, University of California, Los Angeles (UCLA) , Los Angeles, California, 90095, the United States

*Corresponding author. Email: yikun1874@g.ucla.edu

Abstract. With the continuous progress and development of science and technology, the life has been closely linked with information. Under the background of the big data era, today's mathematical statistics methods have also changed significantly, which has brought a very important impact on enterprise management. This paper analyses the values of statistical analysis in the era of big data in enterprise management, discusses the application of statistical analysis in the era of big data in the prediction of A product sales volume in S company in 2022 by using the regression tool in mathematical statistics to provide references for relevant researchers.

Keywords: Statistical Analysis, Business Management, Big Data Era1.

1 Introduction

Statistics is a method based on probability theory to study the basic laws of the quantity change of a large number of random phenomena in society and nature [1]. Statistics takes the data obtained from the observation and experiment of machine phenomena as the starting point, and takes probability theory as the theoretical basis to study random phenomena. We choose mathematical models according to random phenomena, and use mathematical data to verify whether the mathematical models are appropriate. On the appropriate basis, we will study its characteristics, properties and laws. This is the basic method of statistics. In recent years, countries around the world have entered the era of big data [2]. Data has become one of the most important resources in enterprise management. Statistical analysis based on big data plays an important role in guiding enterprise management. Its scientific and accurate operation mode effectively improves the work efficiency and achieves the effect of economic management on the basis of promoting the stable development of enterprises. The application of statistical analysis in the era of big data in enterprise management is of great significance. In the process of production and operation of the enterprise, we should fully comply with relevant policies and regulations, and make appropriate adjustments to our products and businesses. The statistical analysis in the age of big data provides high-quality technical of

© The Author(s) 2024

P. Dou and K. Zhang (eds.), *Proceedings of the 2023 International Conference on Economic Management, Financial Innovation and Public Service (EMFIPS 2023)*, Advances in Economics, Business and Management Research 287,

https://doi.org/10.2991/978-94-6463-441-9_54

long-term development strategies of enterprises. This paper will focus on the characteristics of statistical analysis in the age of big data and its application value in different fields of enterprise management, providing reference for relevant researchers.

2 Statistical Analysis in Big Data Era

2.1 Impact of Big Data Era on Sample Size

The traditional statistical principle is to use a small number of samples to spy on the population, which is always clear in advance [3]. However, in the age of big data, people began to give up the fanatical pursuit of the overall, focusing on analyzing individual characteristics, and predicting the future through individual observation. In the era when the ability to obtain and store data was not very developed, people could not observe the population, and could only take samples through sampling technology to achieve an estimation of the overall situation. It was an indirect sample analysis method. With the advent of the big data era, it is no longer a problem to acquire and store data. People have the ability to acquire massive data that covers almost the whole, including the details that were missed in the original sample. The amount of information is sufficient and the error is greatly reduced. In the era of big data, although there are many types and quantities of information involved, there is often a lack of samples of big data. Therefore, in the process of collecting samples, sampling is not allowed. The data collected by traditional mathematical statistics methods show structural characteristics. In the age of big data, there are very obvious differences between sample data and big data. The sample data has a very clear research purpose. Some data can be obtained by sampling, but the data obtained is limited. Big data not only involves a lot of information content, but also will not be limited by various factors when it is applied. In addition, the use of big data can also accurately determine the law of things' development, so that society can enter the era of big data in a relatively fast time.

2.2 Impact of Big Data Era on Analytical Methods

The magnanimity of big data has led to the crisis of data storage, processing and analysis. Its diversity has changed the type of data studied. Data in some industries involves hundreds of parameters, and its complexity is not only reflected in the data sample itself, but also in the dynamic interaction between multi-source heterogeneous, multi entity and multi space, which is difficult to describe and measure with traditional methods. The lack of advanced statistical analysis methods is the main obstacle to obtain the maximum value from big data. Therefore, how to develop statistical theories and methods to deal with massive data quickly and effectively is the key technology to solve the problem of mining the value of big data. The ability to obtain knowledge, business opportunities and social services through big data can be applied not only to academic research, but also to ordinary institutions, enterprises and government departments, so that more people can

participate in social management. At present, we still lack the awareness of collecting, sorting out and using data to manage things accurately. The impact of big data can not only change the way we think about things, but also increase the way we explore knowledge. Statistical methods provide convenience for exploring nature and laws, so statistical methods need to be taken as a means. Under the background of big data era, we can use a broader statistical thinking mode to analyse and mine the knowledge hidden in the data and reveal many aspects of things. As far as statistical methods are concerned, it is simply to use all kinds of data to observe things effectively to understand the whole process of things in a timely manner, and to deeply explore the nature and laws of things through data [4].

2.3 Impact of Big Data Era on Statistical Results

In the era of big data, many enterprises have made significant changes in the analysis and application of data, and solved the shortcomings and defects in traditional data analysis [5]. Relevant analysis in the age of big data must meet the requirements of universality and equality, and at the same time, it is necessary to ensure that the results will not be affected by other factors. It can be seen from the actual situation in recent years that the influence of big data has been on the rise. In this context, many experts have proposed the research methods of random correlation coefficient and maximum correlation analysis, mainly because there are still great defects in the traditional analysis methods, which cannot meet the relevant needs of enterprise management in the context of big data. Therefore, effective measures should be taken to strengthen the research on statistical methods so that they can meet the development needs of various industries in the context of big data. Traditional data analysis mainly depends on sample data, which can only solve a single problem. Because data has real-time variability, the value of data functions will disappear with the change of data, and it may also be innovative. Therefore, in the era of big data, it is necessary to constantly add value analysis and develop big data. In the era of big data, statistical data sorting mainly includes statistical data audit, data classification, summary and compilation of statistical data tables and charts, chart preservation and publicity. Because of the large amount and variety of statistical data, there are certain requirements for processing speed, and it will be difficult to compile statistical charts. The audit and preservation of statistical data in the age of big data is different from the traditional statistical methods. It uses modern information tools to audit and preserve data.

3 Values of Statistical Analysis in the Era of Big Data in Enterprise Management

3.1 Promote Enterprises to Gain Advantages in Market Competition

Under the conditions of modern market economy, the market development environment is becoming more and more complex [6]. To gain advantages in market

competition, enterprises should not only focus on reducing costs, but also improve efficiency. Enterprises can find out which links are easy to make mistakes and which links can reduce more costs by making statistics on the cost of each link of production and on the production error rate and error rate. For some mature industries, production and operation have become saturated, and there is not much room for cost reduction [7]. However, through the application of statistical science, we can explore the focus that can be optimized in the production and operation process in detail, which can improve production efficiency, focus on improving efficiency to enhance the competitiveness of enterprises, and provide more opportunities for enterprises to win in the market competition [8]. Enterprises often conduct business forecast by combining static analysis and dynamic analysis. In combination with the business data of enterprises in a certain period and the business data of the same industry and product, enterprises analyse and study the enterprise management mode, production and operation plans and objectives. If the cost space cannot be strictly controlled, statistics should be applied to a large extent, focusing on all aspects during the production and operation period, managing the details of production, continuously improving the business efficiency of the enterprise, and promoting the enterprise to expand the external competition space [9].

3.2 Help Enterprises to Cultivate Innovation Awareness

Statistical work and statistical application are new things for enterprises. In the past, statistical work was often used in government work, and enterprises had little contact with it. Although modern science and technology have made rapid progress in recent years, they are still not used in enterprise statistics. In the process of enterprise management, the process of applying statistics is also the process of accepting emerging things, contacting and absorbing innovation consciousness. In today's intense market competition, reducing profits is no longer the main means for enterprises to win in the market competition. Reducing costs and improving management efficiency have become an important breakthrough for enterprise development [10]. Optimizing internal control can improve enterprise management efficiency, reduce internal response time, and maximize the management role brought by limited funds and energy. Through the statistics of the production and operation process of the enterprise, the enterprise management is analysed and controlled to get rid of the loopholes that are formed spontaneously and difficult to find in the management process of the enterprise. Taking the statistical work as an opportunity, the enterprise management is refined and adjusted [11]. Using quantitative means, the best enterprise management mode is formed from the micro level, and the management status quo is corrected to form a standardized and scientific management system. Innovative consciousness and use new methods and ideas to study the production and operation strategy of enterprise. Innovation is the soul of a nation's progress, and also a breakthrough for the long-term development and prosperity of enterprises. Through statistical work, micro details are optimized, which provides new channels for enterprise innovation, and also promotes enterprises to actively explore new development opportunities for sustainable operation [12].

3.3 Impel Enterprises to Pay Attention to Refined Management

The process of statistics is a process of data processing and fine management. The enterprise processes and collates the collected data, draws corresponding conclusions, and then uses statistical conclusions to guide enterprise practice. The statistical results guide the enterprise practice and promote the enterprise to accelerate the process of refined management. The conclusions drawn from data statistics are often micro control in each detail field of operation and management activities, forming the specific process of refined management of the enterprise [13]. Qualitative analysis is often macro. Enterprises cannot ensure the scientific decision-making from a strategic perspective. Once the strategic decisions related to the survival of enterprises are wrong, they will have an irreparable impact. By using statistical tools for statistics and analysis, we can accurately grasp the production quantity, production efficiency, business situation, market share, consumer satisfaction and other data of the enterprise, bring these data into the enterprise development model, and combine with the actual situation of the enterprise, which can provide a reliable basis for the formulation of the enterprise development strategy [14]. In today's increasingly fierce market competition, the progress of fine management has gradually become the decisive factor for enterprises to win in the competition. Modern technology is conducive to improving the scientific and timeliness of statistics, so that the original complicated statistical work that is difficult to implement at the grass-roots level of enterprises can be implemented. Especially for industrial enterprises, statistical data plays a vital role in improving production accuracy [15].

4 Values of Statistical Analysis in Different Fields of Marketing Management

4.1 Statement of Marketing Problems

S Company is an auto parts enterprise. As the sales personnel of S Company did not use mathematical statistics to predict the sales volume of the company's products, they only give the prediction results based on their feelings or experience, and the final results are very inaccurate. Demand forecasting is the basis of marketing management. Now, we need to use mathematical statistics to establish a mathematical model on the basis of big data to forecast the monthly sales volume of Product A of S Company in 2022. The selection of prediction model plays a fundamental and decisive role in the whole prediction. Only by selecting an appropriate forecasting model can we make accurate predictions. We should first analyze the historical data, and select an appropriate prediction model according to the laws of the data. The following table shows the monthly sales volume of Product A for 10 years from 2012 to 2021 collected by the author, as shown in Table 1.

Table 1. A product sale volume from Jan. 2012 to Dec. 2021(table credit: original)

	Ja n.	Fe b.	M ar.	A pr.	M ay.	Ju n.	Ju l.	A ug.	Se pt.	O ct.	N ov.	D ec.
201 2	9	9	7	8	7	7	11	7	11	10	11	14
201 3	13	9	14	13	13	9	7	11	26	16	11	12
201 4	15	15	16	10	14	17	28	17	21	20	25	14
201 5	29	20	24	14	24	21	16	17	26	28	29	31
201 6	36	23	33	35	29	14	25	29	23	29	34	38
201 7	35	27	30	22	35	34	20	36	43	23	30	33
201 8	45	23	50	29	30	21	41	34	47	31	49	52
201 9	58	30	59	41	46	34	32	30	45	32	42	59
202 0	50	33	68	44	51	44	48	32	50	46	50	70
202 1	65	50	65	81	77	50	58	44	55	57	74	74

4.2 Selection of Prediction Model

When we analyse the sales data of each year, we can find that the sales volume in January of each year is high, but it fell back in February, rose in March and April, and reached the first small climax. By May, June, July and August, the sales volume of the products had been declining all the way. From September, the sales volume of the products began to grow continuously until December. This is true from 2012 to 2021. This rule is easy to explain. Product A is an automobile component, and its sales volume is significantly related to the sales volume of automobiles. By observing and analysing the monthly sales of cars, we can know that in January of each year, consumers are psychologically willing to buy new cars before the New Year to show the achievements of this year's work and add to the atmosphere of the New Year. Those consumers who rush to buy new cars have already finished spending in

January, so the sales volume in February will drop a little. In March and April, the weather became warmer, people's desire for consumption increased, and car sales rebounded. Since May, the weather has gradually become hot, and people are often not interested in watching, testing, and buying cars in hot days. The sales volume has gradually declined. Especially in August, when the weather is hottest, few consumers buy new cars, and the sales volume has dropped to the bottom. In September, the temperature slowly dropped. The appropriate temperature and the wealth accumulated in the first half of the year stimulated people's desire for consumption. In addition, with the promotion activities of businesses, the sales of cars began to rise. As November and December approach the end of the year, consumers with a desire to buy are often willing to buy a new car at the end of the year as a reward for their hard work. As a result, the annual sales of auto parts A have a similar trend. This inspired us to study the relationship between monthly sales from 2012 to 2021. Due to space limitations, we only draw a scatter chart of January sales.

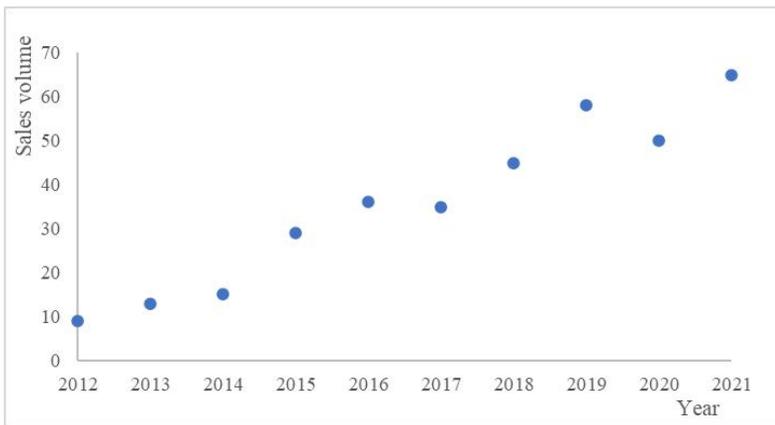


Fig. 1. Sales volumes of Feb. from 2012 to 2021 (figure credit: original)

For a fixed month, with the increase of the year, the sales volume of Product A shows a linear growth trend. We consider using a linear regression model to forecast the monthly sales volume of Product A in 2022.

4.3 Financial Management

We can use SPSS for regression analysis. After selecting “Regression”, we can fill in “Y value input area” and “X value input area” according to the requirements, select “Confidence” of 95%, select “Output area”, and check the linear fitting chart to obtain a series of results such as the estimated value of the regression coefficient. For the sales volume in January, our fitting results are:

$$Y=6.21X+1.33 \quad (1)$$

The F value is 154.56, which is much greater than 5.32. In addition, we can also use the determinable coefficient R^2 of the regression equation to reflect the significance of the regression equation. $R^2 = SSR/SST$, representing the proportion of the sum of regression squares to the total sum of squares, which is the proportion of the sum of squares to the total sum of deviations. R^2 is between 0 and 1. The larger R^2 is, the more significant the regression equation is. Just now, we have obtained the value of the determinable coefficient by using the "data analysis tool" of Excel, so we do not need to calculate it through $R^2 = SSR/SST$. The determinable coefficient of the regression equation for January sales is 0.95, very close to 1, indicating that the regression equation is highly significant. That is, when the significance level is 0.05, the regression equation is significant when the hypothesis is rejected. Therefore, the predicted sales volume of Product A in January 2022 is 69.64. The calculation of regression equation, F value and determinable coefficient in other months are similar, and will not be described one by one. The specific results are as follows:

Table 2. Product volume prediction of every month of year 2022 (table credit: original)

Month	Prediction of 2022	Month	Prediction of 2022
Jan.	70	Jul.	55
Feb.	45	Aug.	46
Mar.	76	Sept.	60
Apr.	63	Oct.	52
May.	69	Nov.	68
Jun.	50	Dec.	81

5 Conclusions

At present, we are in the era of big data. With the continuous reform and innovation of data processing methods, mathematical statistics methods are also in constant progress. Only when it is reasonably applied to the enterprise management can the enterprise's decision-making become more scientific and reliable, promote the enterprise's progress, and bring about more long-term development for the enterprise. This paper expounds the influence of big data era on mathematical statistics, and analyses the application of statistical analysis in marketing management to predict A product sales volume in 2022. The scientific application of big data statistical analysis methods will help promote the development of enterprises, and the development of enterprises cannot be separated from the support of big data information.

References

1. Salazar L R. Exploring the effect of coloring mandalas on students' math anxiety in business statistics courses[J]. *Business, Management and Economics Engineering*, 2019, 17(2): 134-151.
2. Mikhnenko O E, Salin V N. Problems of Modern Transformation of Statistics[J]. *Accounting. Analysis. Auditing*, 2021, 8(4): 18-33.
3. Price C, Walker M. Improving the accessibility of foundation statistics for undergraduate business and management students using a flipped classroom[J]. *Studies in Higher Education*, 2021, 46(2): 245-257.
4. Sharma A M, Srivastav A. Study to Assess Attitudes towards Statistics of Business School Students: An Application of the SATS-36 in India[J]. *International Journal of Instruction*, 2021, 14(3): 207-222.
5. PYLYPENKO O. Integrating Accounting and Logistics Strategy in Business Management[J]. *Scientific Bulletin of the National Academy of Statistics, Accounting and Audit*, 2022 (1): 32-38.
6. Entezami A, Sarmadi H, Behkamal B, et al. Big data analytics and structural health monitoring: a statistical pattern recognition-based approach[J]. *Sensors*, 2020, 20(8): 23-28.
7. Wang R, Wang G. Web text categorization based on statistical merging algorithm in big data environment[J]. *International Journal of Ambient Computing and Intelligence (IJACI)*, 2019, 10(3): 17-32.
8. Sambasivan R, Das S, Sahu S K. A Bayesian perspective of statistical machine learning for big data[J]. *Computational Statistics*, 2020, 35(3): 893-930.
9. He Q P, Wang J. Statistical process monitoring as a big data analytics tool for smart manufacturing[J]. *Journal of Process Control*, 2018, 67: 35-43.
10. Pramana S, Mariyah S. Big data implementation for price statistics in Indonesia: Past, current, and future developments[J]. *Statistical Journal of the IAOS*, 2021, 37(1): 415-427.
11. Gillborn D, Warmington P, Demack S. QuantCrit: education, policy, 'Big Data' and principles for a critical race theory of statistics[J]. *Race Ethnicity and Education*, 2018, 21(2): 158-179.
12. Yongmiao H, Shouyang W. Big data, machine learning and Statistics: challenges and opportunities[J]. *China Journal of Econometrics*, 2021, 1(1): 17-19.
13. Pitcher B W, Kent A J R. Statistics and segmentation: using Big Data to assess Cascades Arc compositional variability[J]. *Geochimica et cosmochimica acta*, 2019, 265: 443-467.
14. Sharples L D. The role of statistics in the era of big data: electronic health records for healthcare research[J]. *Statistics & Probability Letters*, 2018, 136: 105-110.
15. Li X, Chang M, Wang L. Information recognition of pathogenic modules in gene statistics of big data[J]. *Nanomaterials and Energy*, 2021, 10(1): 35-42.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

