The Historical Evolution and Paradigm Shift of Mathematical Statistics Theory - The Theoretical Trajectory from Descriptive Statistics to Data Science

Chengqing Li
Shandong Experimental High School, Jinan, Shandong, 250000, China
2162818770@qq.com

ABSTRACT

In the past 100 years, mathematical statistics has undergone great changes in theory. It has changed from a science that only describes numbers to a science that can predict and provide suggestions, which is very helpful for today's data decision. This study reports that the history of statistical theory is divided into four main stages: first, the classical period of descriptive statistics and probability basis; Followed by Fisher, Neyman and Pearson's statistical testing revolution. Then, Bayesian regression and ideas about decision; Finally, today's data science has an algorithm model, a large number of dimensions of analysis and computer learning. Using the insights of probability theory, philosophy of science, computer science and applied mathematics, we have constructed a story, showing the consistent content and completely changed content in the theory. Our analysis shows that each paradigm shift is driven by the relationship between mathematical innovation, computing power and practical application requirements. At present, the transformation from model-based reasoning to algorithm-based prediction is not only a technical improvement, but also a new way of thinking, which can be learned from data. This study provides a complete map of the history of statistical thinking, shows the tension between frequency and Bayesian epistemology, and provides a framework to better understand how statistical theory and data science change together.

Keywords: Mathematical Statistics; Descriptive Statistics; Inferential Statistics; Bayesian Inference; Data Science; Statistical Learning

1. INTRODUCTION

The discipline of statistics is special: it is between mathematics, science and practical decision-making. Pure mathematics uses rules to seek truth and science to seek explanation, but statistics provides us with a language to think when we are uncertain. This language is very useful in many fields, such as medical testing, physics, economics and artificial intelligence [1].

However, the statistical methods we use today were born not long ago. Things like standard deviation, hypothesis testing, confidence interval and Bayesian dating, which we are studying now, were invented at a specific time to answer scientific questions. Knowing this story is not just for fun; This shows us important insights that still change the way we study data today.

This study shows how mathematical statistics has changed in four main stages: (1) To describe the classical period of descriptive statistics, we mainly aim to summarize and display data; (2) frequent reasoning revolution, which creates a formal framework for estimation and testing; (3) Bayesian regression, questioning other methods with frequent ideas about probability and reasoning; (4) The paradigm of data science, today, through algorithmic methods, high-dimensional models and powerful calculations, changes the discipline.

By observing how statistical theory responds to the changes in science, data availability and computer functions, we hope to better understand the epistemological basis of data analysis today. We say that the current shift from model-based reasoning to algorithm-based prediction is a huge paradigm shift-not just a tool change, but a new way to think about the relationship between theory, data and knowledge.

2. RESULTS AND DISCUSSIONS

2.1 The Emergence of Descriptive Statistics

Before statistics became a formal mathematical discipline, the collection and summary of data was primarily a statecraft enterprise. The word "statistics" comes from the Latin statistic which means "country". This shows that statistics begin with census, taxation and military registration. Early statistical work, such as The Book of Sunday and Grante's observation of the Mortality Act (1662), was mainly descriptive: they counted people, measured resources and found the regularity of birth and death.

The development of descriptive statistics really began in the 19th century, Adolphe Quetelet and others. He used normal distribution to study social data and invented the idea of "average person". Quetelet shows that social things can be as regular as physical science. This opens the door for statistical thinking about human affairs. At that time, charts were also improved to display data, central trend and scattering measurement were invented, and Francis Galton formally determined the correlation.

But descriptive statistics are mainly tools to show and summarize things. Its theoretical basis is very weak: there is no clear stochastic model, no formal framework to make assumptions beyond the data, and no established estimation theory. Data is just something to describe, not the basis of generalization [2].

2.2 The Development of Probability Theory as a Foundation

The development of probability theory laid the theoretical foundation of inferential statistics. In the early days, the history of probability came from questions about gambling and risk, such as cardano, Pascal, Fermat and Bernoulli who found the basic rules for calculating probability. The law of large numbers, proved by Jacob Bernoulli (1713), gave us a very important idea: when you take more samples, this idea enables us to use the data we see to guess what we can't see directly.

2.3 The Consolidation of Descriptive and Probabilistic Thinking

The integration of descriptive statistics and probability theory began in the works of Francis Galton, karl pearson and William Leksis at the end of 19th century. Pearson invented chi-square test and found a way to adapt distribution to data. Galton used Quetelet's work to create regression and correlation to study how organisms inherit things. This combination provided the basis for formal statistical reasoning, but it was not until the 1920s that a unified reasoning system appeared.

3. THE FREQUENTIST INFERENTIAL REVOLUTION

3.1 The Logics of Estimation, Testing, and Design

The 1920s and 1930s witnessed a transformation of statistics from a descriptive and data-summarizing practice into a formal inferential discipline. The central figure was Ronald A. Fisher, who developed the framework for estimation (maximum likelihood, sufficiency, efficiency), the logic of hypothesis testing (the significance test), and the principles of experimental design (randomization, blocking, factorial designs). Fisher's work gave statistics a coherent logic: data were to be used to estimate parameters, test hypotheses, and design efficient experiments for scientific inquiry.

Jerzy Neyman and Egon Pearson developed another system to test hypotheses. They invented the idea of testing power and confidence interval for type I and type II errors. Fisher's and Niemann-Pearson's systems are different: Fisher's significance test looks at whether the data says anything about the null hypothesis, while Niemann-Pearson's test regards decision-making as a problem of controlling long-term errors. Even though they are different, both methods use frequent visitor probability.

3.2 Fisher, Neyman, and Pearson: Rival Visions

The debate between Fisher, Neyman and Pearson shows that they disagree with the true meaning of statistics. Fisher believes that statistics is the logic of reasoning: meaning test is a tool to look at evidence, and scientists make decisions, not tests [3]. Niemann and Pearson believe that statistics is a method of action: testing is a decision-making process with controllable error rate, and you must make a decision when considering what will happen next.

3.3 Consolidation and Limitations

Frequent visitor paradigm has become the main method of applied statistics and has been widely used in scientific research and government statistics. But he has a problem: people don't understand the concepts of trust and power, they misunderstand the meaning of P value, and they misuse the logic of testing. In addition, the conventional framework needs to choose zero hypothesis and alternative hypothesis, and it is difficult to choose the correct model or use a large number of parameters to solve the problem. These problems leave room for other methods to return [4].

4. THE BAYESIAN RESURGENCE AND THE RISE OF DATA SCIENCE

4.1 Bayes' Theorem and the Bayesian Revival

Bayesian inference, according to Bayesian theorem (1763), is older than many statistical frequencies. However, Bayesian method is a bit shelved because of the frequent consensus in the 1920s. This is because of practical problems (Bayesian calculation is too difficult for computers) and because of differences in ideas.

The modern revival of Bayes began in the 1950s, thanks to the work of L. J. Savage and Bruno de Finetti. They explained that probability can measure what an individual believes. Savage's book "Statistical Basis" (1954) provides a framework for decision-making, indicating that if you want to make a correct decision when you are uncertain, you must use Bayesian dating to change from the previous belief to the later belief.

Bayesian framework is different from frequency in how to understand probability and how to use existing information. By requiring the assignment of priority distribution, Bayesian inference clearly shows that the frequency method sometimes hides ideas, but it adds some subjectivity. The practical problems that hinder the early Bayesian method are becoming more and more important now, thanks to the algorithm for calculating the subsequent distribution.

4.2 The Data Science Paradigm: Algorithms, Computation, and Prediction

The paradigm of modern data science may be the biggest change in statistics since the frequent visitor revolution. With the development of big data, computer power and algorithms, this paradigm has changed important things: from model-based reasoning to algorithm-based prediction, from parameter estimation to function approximation, from hypothesis testing to cross-validation.

The emergence of machine learning has brought new tools-neural network, decision tree, support vector machine, lifting and archiving-which are not easy to use in the old reasoning framework. These methods do not rely on parameter assumptions. They have flexibility and data adaptability, and can model complex nonlinear relationships [5]. They are usually large in scale and require a lot of calculation, which requires new pattern selection, regularization and verification methods.

4.3 From Parameter Estimation to Function Approximation

Data science has changed the purpose of statistical analysis: not to estimate parameters, but to seek approximate values of functions. Instead of testing hypotheses, we try to predict the results. Instead of measuring uncertainty, we try to make the best decision. This change is helpful to create a theory for the calculation of a large number of data, model selection and regularization. Modern machines need a lot of calculation, so we invented the fast algorithm. This change also raises questions about what statistical evidence is and how statistical theory can help us do our work well.

The difference between Bayes and regulars is not impossible to repair. She gave good ideas, such as robust Bayesian method, prior reference, and the method of checking the model with common ideas. When you mix these two viewpoints together, you will find a new direction of statistical theory and practice [6].

5. CHALLENGES, PERSISTENT DEBATES

5.1 The Role of Statistical Theory in the Age of Data Science

The transition from model-based reasoning to algorithm-based prediction raises important questions about the nature of statistical theory. Traditional theory focuses on estimation, testing and design. She asked: What can we infer from the data? The paradigm of data science raises different questions: What can we predict from data? The transition from inference to prediction requires rethinking the role of statistical theory.

Data science is usually regarded as a "new" discipline, but it is consistent with the oldest statistical tradition. Methods such as cross-validation and guidance come from statistics. The paradigm of data science makes them more important. The continuous increase of big data and computing power has changed statistical practice, but many basic challenges-over-optimization, model selection and causal reasoning-remain the core.

5.2 Reproducibility and Algorithmic Governance

Repeatability in science has prompted changes in methods: pre-registration, transparency and analytical rules. The crisis of repeatability shows that statistical theory is very important to guide practice, and it also shows that there is a gap between statistical methods and practical problems [7].

The great transformation to data science has brought new challenges: algorithm deviation, fairness, interpretability and the relationship between prediction and interpretation. When statistical models become more complex, they become more difficult to understand, which raises questions about trust and accountability. Creating a theory that can help make the model easy to understand is an important direction of future research.

6. CONCLUSION

This study tracks the historical evolution of mathematical statistics in four paradigms, from descriptive statistics to data science. The evolution of statistical thinking shows the path from data generalization to parameter estimation to result prediction. Every paradigm shift is driven by the interaction between mathematical innovation, computing power and actual needs. The transformation from model-based reasoning to algorithm-based prediction is a new way to think about the relationship between theory, data and knowledge. The future of statistics will need a new theoretical framework, which will mix reasoning and prediction, reconcile Bayesian and frequent visitor methods, and solve the problems of repeatability, fairness and interpretability. This will be a very important comprehensive view of development statistics. As a discipline, it provides tools for reasoning and uncertainty in complex situations, which have many dimensions and are constantly changing.

REFERENCE

[1] Qasem, S. N., & Al-Najjar, H. Short-term load forecasting using deep neural networks with fuzzy logic integration for uncertainty handling. Journal of Electrical Engineering & Technology, 2025, 20(3): 1875-1890.
[2] Kumar, A., Singh, R., & Sharma, P. FIRE-CNN-LSTM: A fuzzy rough set-evolved hybrid deep learning model for short-term load forecasting. Neural Computing and Applications, 2024, 36(18): 10845-10865.
[3] Yolcu, O. C., Lam, H. K., & Yolcu, U. Short-term load forecasting: Cascade intuitionistic fuzzy time series—univariate and bivariate models. Neural Computing and Applications, 2024, 36(22): 14189-14208.
[4] Nguyen, T. H., & Tran, Q. V. Stacked hybrid model for load forecasting: Integrating transformers, ANN, and fuzzy logic. Scientific Reports, 2025, 15: 19688.
[5] Hossain, M. A., & Lee, S. J. Design of a hybrid learning model for establishing consistency in smart grid environment. Scientific Reports, 2025, 15: 44903.
[6] Zhao, Y., & Chen, W. Uncertainty quantification in smart grid load forecasting: A fuzzy logic-based framework. IEEE Transactions on Power Systems, 2024, 39(4): 6123-6135.
[7] Park, J. H., & Kim, D. H. Online fuzzy rule-based learning for short-term load forecasting with adaptive model selection. Applied Energy, 2025, 358: 122612.