

# Reflection on the Talent Training Mode Transformation of Statistics Specialty against the Background of Big Data

MU Xin<sup>1, a</sup>, LI Zhigang<sup>2, b</sup>

<sup>1</sup>College of Mathematics and Statistics, Baicheng Normal University, Baicheng, 137000, China

<sup>2</sup>The Fifth Middle School of Changling County of Jilin Province, Changling, 131500, China

<sup>a</sup>email: keaidemx@163.com, <sup>b</sup>email: 1455115227@qq.com

**Keywords:** Big Data, Statistics, Talent Training Mode, Transformation

**Abstract.** In the information explosion society, big data has entered all aspects of the social economy. Statistics, as a supporting discipline of big data, has received more and more attention. This paper expounds the basic concepts of big data and statistics specialty, analyzes on the demand of statistics major talents and the disadvantages of traditional training mode, and gives some thoughts on the training mode and transformation of statistics major talents under the background of big data. At last, the conclusion is summarized and prospected.

## Basic Concept Elaboration

**Big Data.** Big data's concept was first put forward by Victor Mayer Scheinberg and Kenneth Cook. Big data also called a huge amount of data, that is not by sampling survey shortcut, but to analyze all the data. Big data needs a new processing model to have more decision-making power, insight, process optimization ability, and a variety of information assets. The core of big data's technology lies not in the huge amount of data, but in the specialized processing of data, that is, the ability of data analysis[1].

1.Data acquisition. Due to big data's diversity and abundance, the sources and types of data are extremely complex. It is necessary to use search engine and ETL engine to extract and integrate data, then to clean the data, eliminate useless information and improve the quality of data.

2.Data analyzed. Big data's analysis, using the distributed computing technology, that is, connecting the major computers in the system, forms a large-scale super virtual computer to complete the data analysis and processing in a large number and in a very complicated form.

3.Data presentation. The computer will collect and screen, analyze and process the data, with the help of charts and other visual means, ensure accuracy, at the same time, the results will be presented in an intuitive way for people to understand, to provide guidance for decision makers. It is important to note that if the data is not intuitive and difficult to understand, it can mislead decision makers and even draw the exact opposite conclusion.

**Statistics Professional Talents.** Statistics is a branch of applied mathematics, mainly through the use of probability theory to establish mathematical models, collect the observed system data, quantitative analysis, summary, and then infer and predict, to provide a basis and reference for relevant decisions. It is widely used in a wide range of disciplines, from physics and social sciences to the humanities, and even to business and government intelligence decisions[2].

The major of statistics is divided into three major directions: the direction of mathematical statistics, the direction of economic statistics and the direction of applied statistics. The difference between the direction of mathematical statistics and the direction of economic statistics is not very large. Mathematical statistics mainly studies the basic theories and methods of statistics; economic statistics is to provide scientific surveys, collect economic information, and describe. Analysis of economic data and the socio-economic process of prediction, supervision of a science. The applied statistics is mainly to investigate, collect the data information of the observation object, analyze the characteristics of the observed object, discover the law of things, predict and supervise the object by describing the technique of statistics, so as to realize the benign operation of the society and economy.

## **Current Situation of Statistics Specialty Talent training under the Background of Big Data**

**Analysis of Statistics Specialty Talent Demand.** 1. International demand analysis. Up to now, the major of statistics has been paid more and more attention in the world. Statistics, engineering, life science, environmental science and finance are five hot fields according to the data of the Bureau of Labor Statistics of the United States. Graduates of statistics major have a good chance of being employed in promising positions. Among them, banks and insurance companies, finance companies, pharmaceutical companies have the greatest demand for statistical talent, and the treatment they give is also very good. Some enterprises market research departments and government agencies also have a certain demand for statistical personnel[3].

2. Domestic demand analysis. In our country, the demand for statistics talents is scattered. In the traditional field, government statistics department, enterprise finance department, production and operation management department need a large number of statisticians in each link. With the development of economy, more and more market research and consulting companies appear, which also need a certain number of statisticians, especially in Beijing, Shanghai and other areas with a high level of economic development. In addition, the financial industry, the rapid rise of the insurance industry, but also produced a large gap in statistics professionals.

**The Defects of the Talent Training Traditional Mode in Statistics Specialty.** With the emergence of big data, statistics is becoming interdisciplinary, applied, integrated and market-oriented. The traditional statistical talent training model, whose defects become increasingly prominent, has been difficult to meet the unpredictable talent demand market. The main performance is: first, the goal of talent training is too narrow[4]. The traditional training mode is usually oriented to the government statistics department, the enterprise statistics post, the university training goal is single, the teaching content is too specialized, and the adaptability to the future work post is not strong. Second, the curriculum is unreasonable, the subject cross is few, the synthesis is poor, the level of the teaching plan of some schools is not clear, the basic course is repeated a lot, the teaching content is too old and old. Third, theory is out of touch with practice. Paying too much attention to the training of theoretical knowledge deviates from the practicality and application of statistics specialty, which results in the students' learning content can not be fully displayed in the work post, and it is difficult to be qualified for the post demand of the employing unit.

**The Training Model of Statistics Specialty Under the Background of Big Data (Taking Tsinghua University as An Example).** Tsinghua University announced the establishment of the Institute of data Science on April 26, 2014, and launched a number of multidisciplinary and interdisciplinary related master's degree courses, such as data Science and Engineering, Internet Finance, big data State Governance, Business Analysis, etc. These subjects are co-founded by six faculties. Its personnel training model is as follows:

Talent training Program: big data Master Program, data Science and Engineering Master Program, big data ability Promotion Project, etc.

Personnel training objective: to cultivate practical talents with data analysis and management ability, big data thinking ability, innovation ability, good professional accomplishment, can serve the relevant work in various fields.

Curriculum: basic modules include big data system foundation, big data analysis SAS text data mining and so on. Data ethics, big data governance and policy, etc.; practice modules include big data related expert lectures and big data practical courses[5].

We can find that under the background of big data, the training model of statistics talents has the following characteristics: first, focusing on the combination of theory and practice, setting up a special practice module; second, the subject intersection is obvious. But they are all based on statistics and computer science. Thirdly, the curriculum design adopts modularization, which highlights the practicality and comprehensiveness. However, there are still some problems to be discussed in this new model. For example, the subject of talent training is still too single. In addition to universities and scientific research institutes, the government and enterprises also need to participate; for example, The characteristics and guidance of talent training also need to be more clear. It is obviously difficult to have feasibility to cultivate big data all-around talents. The training

units must pay attention to the changes of market demand and highlight their own advantages and specialties.

## **Transformation of Talent training Mode of Statistics Specialty under the Background of Big Data**

**Pay attention to the Cultivation of Applied Talents with Cross-Disciplinary Background.** With the emergence of big data and the rapid development of computer network technology, the connotation and extension of statistics has been greatly expanded. Therefore, under the background of big data, the position of statistics professionals should be comprehensive talents with computer, mathematics, statistics and other cross-disciplinary background. Moreover, in the practical application of different fields, but also need to be combined with specific disciplines. For example, when studying psychological problems, we need to combine certain psychological knowledge; when studying government statistics, we need to combine the reality of national economy, and so on[6].

In addition, subject intersection is not a hodgepodge. Under the background of big data, the cultivation of statistical talents does not mean that the more subjects involved, the better the effect, the more miscellaneous but not the better, which is not conducive to achieving the training goal. The cultivation of statistics professionals depends on the coordination and cooperation among different departments. In the process of selecting specific models and specifications, it is necessary to classify and design talents in different directions and levels. For example, the cultivation of undergraduate students should focus on the specific application, emphasize the adaptation to the needs of the society, strengthen the ability, graduate training, need to adopt a higher standard, taking into account the theory and practice, To ensure that trained students can independently undertake big data's collection, integration, use of work planning and implementation.

**Using Modular Curriculum Settings.** As mentioned above, under the background of big data, the training mode of statistics talents pays attention to the cultivation of applied talents with cross-disciplinary background to meet the needs of the market. Therefore, the corresponding modularization design is also needed in the aspect of curriculum design. Generally speaking, can be roughly divided into two major modules. First of all, it is the professional basic course module, which consists of two parts: statistics course and computer course module. Statistical courses include probability theory, applied mathematical statistics, introduction to statistics, application of stochastic processes, etc. The courses of computing include big data's analytical and statistical basis, big data's analytical computer basis, unstructured big data's analysis, etc. Big data, distributed computing and so on. Second, applied courses, such as monetary and banking, accounting, regional economics, international economics, etc., management fields include management, financial management, human resources management, organizational behavior, etc. Sociology includes introduction to sociology, social psychology, history of social thought, social statistics and computer application, etc. Pedagogy includes educational psychology, teaching theory, educational philosophy, general theory of education, etc.

In a word, the modularized curriculum should be based on statistics and computer courses, combine with different application fields to set curriculum flexibly, combine theory with practice as far as possible, and train the specialized talents of statistics under big data background.

**Adopt Practical Teaching Methods.** Under the background of big data, the cultivation of statistics professionals should pay attention to the application and practice, therefore, it is necessary to match the teaching methods. First, the use of scientific research project-based teaching. Through the case study and discussion of the teaching model, let students more contact with a professional field of practical problems, to cultivate the ability to solve problems, so as to graduate to work, can adapt quickly, learning to use. Second, adopt multi-tutor system. Under the background of big data, the new model of training statistical talents involves the intersection of many disciplines and the synthesis of multiple abilities. It is very difficult for a single tutor to take on all the guiding work, but to adopt a joint training model based on a multi-tutor system can take advantage of its strengths and circumvent its weaknesses. The teachers complement each other to provide high-quality

guidance from both theory and practice. Third, carry on individualized training. Every university, or research institute, has its own field of expertise and not good at the field, as the saying goes across the mountain, even if the comprehensive strength is very difficult to achieve perfection, to cultivate big data is even more difficult to generalist. Therefore, universities and scientific research institutes can choose their own superior disciplines, gather the scientific research and teaching strength of the units, and carry out individualized personnel training on the basis of statistics and computer science.

**The Construction of the Principal-Training Subject of Multi-Party Participation.** As mentioned earlier, under the background of big data, the demand for statistical talents, with a wide range of industries, high professional requirements, and strong market orientation, it is very difficult to achieve good results by relying solely on the single-handedness of universities or scientific research institutes. It is very important to construct a multi-party personnel training subject and realize the precise division of labor and deep docking between the subjects. The government, enterprises, schools, research institutes, big data resources and other main bodies need to unite and cooperate in division of work to realize the docking between the research and application of big data, the supply and reception of statistics talents, and the docking between theory and practice. Construct the training platform of statistics talents together.

## Conclusion

With the development of economy and the innovation of computer network technology, big data has been widely used in the field of statistics. The traditional training model of statistics talents is too narrow, the curriculum is unreasonable and the theory is out of touch with the practice, so it can not meet the needs of the market. Under the background of big data, we need to explore a new model of training statistics professionals. The universities represented by Tsinghua University have explored this new model, which has the following characteristics: first, it has a clear interdisciplinary trend, based on statistics and computer science. Integration of specific applied disciplines; second, the use of modular curriculum to highlight the comprehensiveness and application; third, more emphasis on the combination of theory and practice, the use of scientific research project-based teaching, multi-mentor system, personalized personnel training. We also need to actively construct the main body of personnel training, realize the deep docking between the subjects, and jointly create a modern training platform for statistical talents.

## References

- [1]Manyika J,Chui M,Brown B,et al.Big data: the next frontier forinnovation, competition and productivity [R]. US: Mc Kinsey GlobalInstitute Report,2011.
- [2]LI Jinchang.Big Data and New Statistical Thinking[J].Statistical Research,2014,(1).
- [3]ZHAO Chenting.ZHANG Hao. Big Data Dominates Present and Future[N].First Financial Daily ,2015,11,16(9).
- [4]JIANG Jianhui.The Training Goal of Statistics Specialty in Colleges and Universities in China and the Way to Realize it[J].Statistical Research,2005(8):74-76.
- [5]YANG Tinggan. Research On the Construction of Statistical discipline, Talent training and Teaching Reform[J].Statistical Research,2005(2):10-14.
- [6]XIANG Chengguan,XIONG Shiheng,WANG Dong.Research On big data's Analysis of Talent training Model in Colleges and Universities[J]China Science and Technology Information,2014(9).