The Quantitative Study on the Distribution of Higher Education Resources Based on Information Theory

Miao Dan[(✉)]

School of Economics and Management, Beijing Jiaotong University, Beijing, China
21120744@bjtu.edu.cn

Abstract. The development level of higher education is an important symbol of a country's development level and potential. Whether the distribution of educational resources is relatively balanced and fair is an important issue. China's educational reform and development has made remarkable achievements, but the distribution and allocation of higher education resources still need to be paid attention to. By using the quantitative research methods of information entropy and Theil index in information theory, this paper analyzes the present characteristics of Chinese higher education resources distribution. The results show that the distribution of Chinese higher education resources is unbalanced. Western development lags behind central and eastern regions. The gap among the west, the middle and the east is gradually decreasing, and the distribution of higher education resources is developing in a balanced and coordinated direction. It is suggested to use big data technology and information technology to promote education resource sharing, strengthen the construction of education information technology in less developed areas, boost education decision making, increase resource input, establish the resources integration and sharing system, so as to promote the further balanced development of higher education in China and let higher education benefits more people.

Keywords: distribution of higher educational resources · information theory · the information entropy · Theil index

1 Introduction

China has made significant achievements in education reform and development. It has built the largest higher education system in the world, with the enrollment rate of higher education rising from 42.7% in 2016 to 57.8% in 2021. During the 14th Five-Year Plan period, higher education will enter the stage of high-quality development after it enters the stage of popularization. Universities of all levels and types will build an innovative, coordinated and open higher education system through the integration of science and education and the integration of industry and education [1]. The high-quality development of higher education is a major concern of the Party and the country, which will certainly require and attract more resource input.

On the other hand, "information technology" and "big data" have become buzzwords. Developing big data in education has become a strategic choice to promote the deepening reform and innovative development in our current education field. Good educational data can play an important role in improving the quality of education, promoting educational equity and optimizing the allocation of educational resources. And information technology is having and will still have a profound impact on education. This influence will promote the development of educational equity through shaping the internal and external fair soil of education, spreading the concept of fairness and justice and satisfying students' personalized learning needs and other ways.

Higher education resources are an important foundation for determining the quality and scale of regional higher education and a strong guarantee for our country to build innovative countries and improve international influence. An accurate grasp of the current distribution and evolution process of higher education resources is of great significance to optimize the structure of higher education, promote the coordinated development of regional higher education and social economy, and enhance national status. Thus, the distribution and allocation of higher education resources deserve important attention. From the perspective of quantification, with the help of the data related to the national education resources in recent years, this paper will study and analyze the status quo of the distribution of Chinese higher education resources, draw a conclusion about the time and space dimensions, and put forward corresponding countermeasures and suggestions. In theory, it is helpful to apply relevant empirical methods to the research on the distribution of higher education resources in the whole country, take a wide view of all provinces and cities in the country, and realize the expansion of research objects. In practice, countermeasures and suggestions are put forward to further promote the relatively balanced and fair distribution of higher education resources in China, which can guarantee people's higher education needs and promote the construction of innovative, coordinated and open higher education system.

At first, the research related to higher education resources are not much, especially the research on the resource allocation problems of higher education is more rare. Before 2008, some scholars study the resource allocation of higher education on the theoretical level. And some scholars discuss the system and other aspects of the practical level. But they didn't hit the fundamental problem of China's higher education resources distribution-the actual existence of disequilibrium distribution of regional higher education resources [2]. Later, there were more studies on the status quo, causes and countermeasures of the unbalanced distribution of higher education resources. Relevant scholars from a specific angle, such as the spatial structure change, spatial accessibility, have analyzed the relationship between the distribution of educational resources and another related factor (such as regional economic development, population spatial distribution, innovation ability, etc.), which was from studying the present distribution of China's overall higher education resources to the influence factors and countermeasures, to taking a province or city group as an example [3, 4]. It has experienced a development process from phenomenon to essence and from general to special.

In view of the distribution status of higher education resource denial, most scholars used descriptive analysis to study the evolution pattern, distribution status, influencing factors and countermeasures. Most of them focused on qualitative discussion. At the same

time, some empirical studies were conducted by building models and using quantitative methods to conduct further research. But more limited in specific provinces and cities or city clusters (such as the Beijing-Tianjin-Hebei region, etc.), which didn't promote in the whole country.

Therefore, this paper uses the quantitative research methods of information entropy and Theil index to conduct an empirical study on the status of higher education resources distribution in the whole country, and draws a relevant conclusion on the status of higher education resources distribution in China.

2   Materials and Methods

2.1   Data Materials

This study takes the whole country as the research area. The data mainly comes from China Statistical Yearbook and statistical yearbooks of various provinces and cities from 2011 to 2020. Due to data collection reasons, the study area of this paper does not include Hong Kong Special Administrative Region, Macao Special Administrative Region and Taiwan Province.

In this paper, the division method of "East, Middle and West three zones", which is frequently used by scholars in previous studies, is used to specifically divide a total of 31 provinces, autonomous regions and municipalities (hereinafter referred to as provinces and cities) except Hong Kong, Macao and Taiwan as follows, as shown in Table 1.

2.2   The Research Methods

2.2.1   The Information Entropy

Information entropy is the basic concept of information theory. By means of entropy, this paper measures the equilibrium of the distribution of higher education resources in China. It reflects the spatial distribution and equilibrium degree of higher education resources in various provinces and cities in China in a certain period of time. The higher the entropy, the lower the order degree and the lower the equilibrium degree. On the contrary, the higher the order degree, the higher the equilibrium degree [5].

Let the total number of colleges and universities in China be S. And let the number of provinces and cities in China except Hong Kong, Macao and Taiwan be n (n = 31).

Table 1. The specific division of China's three major zones

Three Zones	Province and City
East (12)	Beijing, Tianjin, Hebei, Shanghai, Jiangsu, Zhejiang, Shandong, Liaoning, Fujian, Hainan, Guangdong, Guangxi
Middle (9)	Shanxi, Jilin, Heilongjiang, Anhui, Jiangxi, Henan, Hubei, Hunan, Inner Mongolia
West (10)	Chongqing, Sichuan, Guizhou, Yunnan, Tibet, Shaanxi, Gansu, Qinghai, Ningxia, Xinjiang

The total number of colleges and universities is S_i ($i = 1, 2\ldots, n$). The proportion of the total number of colleges and universities in each province in the total number of colleges and universities in China is P_i, that is,

$$P_i = \frac{Si}{S} = \frac{Si}{\sum Si} (i = 1, 2\ldots, n, \quad n = 31) \tag{1}$$

Obviously, $P_i \geq 0$ ($i = 1, 2\ldots n$, $n = 31$) and has normalization, so P_i is equivalent to the probability of events. Thus, the information entropy H of higher education can be defined according to Shannon's entropy formula, which is

$$H = -\sum (Pi \ \ln \ Pi)(i = 1, 2\ldots, n, \quad n = 31) \tag{2}$$

Here, H is information entropy. When the number of colleges and universities in each province is equal, namely $S_1 = S_2 = \ldots = S_n$, the information entropy reaches the maximum, that is, $H_{max} = \ln n$.

Based on the information entropy formula, the equilibrium degree J and dominance degree I of higher education can be defined, that is,

$$J = \frac{H}{Hmax} = \frac{H}{\ln n} = -\frac{\sum (Pi \ \ln \ Pi)}{\ln n}(i = 1, 2\ldots, n, \quad n = 31) \tag{3}$$

$$I = 1 - J \tag{4}$$

where, J is the equilibrium degree. When $J = 1$, it reaches the ideal equilibrium state; when $J = 0$, it is in the least equilibrium state. The larger the value of J, the stronger the equilibrium of higher education in each province. I is the degree of superiority, which reflects the degree of superiority of higher education in each province, and is opposite to the degree of equilibrium.

2.2.2 Theil Index Model

Theil index is an index proposed by Theil (1967) to measure regional development differences, which has been widely used in the study of economic and social development. According to Theil index model, the total regional differences can be divided into two parts: inter-regional differences and intra-regional differences. Intra-regional differences can be further divided into intra-regional differences of different regions, which is of great value to further explore the evolution characteristics and rules of regional development differences. The value of Theil index is between 0 and 1. Generally speaking, the larger the value of Theil index, the greater the regional difference and the higher the degree of imbalance between regions. The smaller the Theil index, the smaller the regional difference. Theil exponent can be calculated as follows:

Total regional difference T_{Total}:

$$T_{Total} = T_{Inter} + T_{Intra} = \sum_{i=1}^{n} Yi \log \frac{Yi}{Pi} + \sum_{i=1}^{n} Yi(\sum_{j} Yij \ \log \frac{Yij}{Pij}) \tag{5}$$

Inter-regional differences T_{Inter}:

$$T_{Inter} = \sum_{i=1}^{n} Yi \ \log \frac{Yi}{Pi} \tag{6}$$

Intra-regional differences T_{Intra}:

$$T_{Intra} = \sum_{i=1}^{n} Yi(\sum_{j} Yij \ \log \frac{Yij}{Pij}) \tag{7}$$

Among them, n is the number of regions (China is divided into the eastern, central and western three zones, n = 3). Y_i represents the proportion of the total amount of colleges and universities in each region in China in the total amount of colleges and universities in the country. P_i represents the proportion of the number of provinces and cities in each region in the total number of provinces and cities in the country (excluding Hong Kong, Macao and Taiwan, a total of 31 provinces and cities). Y_{ij} represents the proportion of the number of institutions of higher learning in the JTH province of region i in the region. And P_{ij} represents the reciprocal of the total number of provinces and cities in each region. T_{Inter}/T_{Total} and T_{Intra}/T_{Total} represent the contribution rate of inter-region difference and intra-region difference respectively.

3   Results & Discussion

3.1   The Information Entropy

The information entropy H, equilibrium degree J and dominance degree I of higher education resources in each province and city from 2011 to 2020 are obtained according to the annual statistical yearbooks and the calculation formulas of information entropy, equilibrium degree and dominance degree. The results are shown in Table 2 and Fig. 1, 2 and 3.

It can be seen that from 2011 to 2020, the equilibrium degree basically increased year by year, from 0.95743 in 2011 to 0.96120 in 2020. And it gradually approached the ideal equilibrium state of 1. The dominance degree basically showed a downward trend year by year, and the changes of the two degrees were obvious before 2017. We know that the distribution of higher education resources in China is still unbalanced, but it is developing in a balanced direction with the passage of time.

Table 2. Results of information entropy analysis

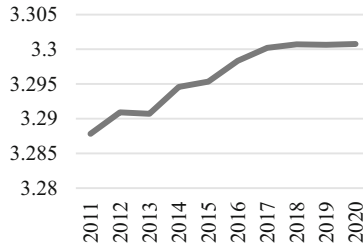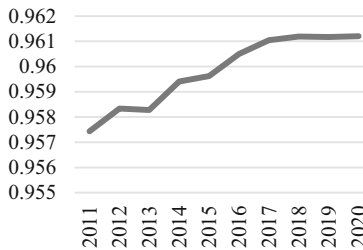| Table Head	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Information Entropy H	3.28781	3.29091	3.2907	3.29459	3.29534	3.29834	3.30021	3.30072	3.30064	3.30076
Equilibrium Degree J	0.95743	0.95834	0.95827	0.95941	0.95962	0.9605	0.96104	0.96119	0.96117	0.9612
Dominance Degree I	0.04257	0.04166	0.04173	0.04059	0.04038	0.0395	0.03896	0.03881	0.03883	0.0388

Fig. 1. Information Entropy H
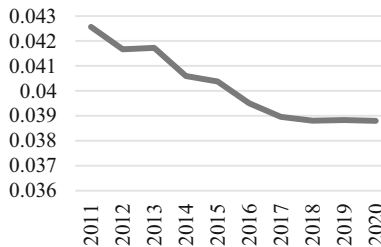

Fig. 2. Equilibrium Degree J

Fig. 3. Dominance Degree I

3.2 Theil Index Model

According to the statistical data in China Statistical Yearbook from 2011 to 2020, the results are shown in Table 3 and Fig. 4 through Theil index model.

As can be seen from the chart information, the National Theil index of the whole country decreased year by year from 2011 to 2020, indicating that the inter-provincial difference in higher education resources in China has been decreasing on the whole. This difference is mainly due to the imbalance within the three regions of the eastern, central and western regions. Specifically, Theil index in the eastern region decreased slightly in 2012, steadily increased in 2013 and 2014, then decreased in 2015, and then rose to the maximum value of 0.0193 in 2016, and steadily declined in the following years. Theil index in western China increased steadily from 2011 to 2013, decreased in 2014, increased significantly in 2015, decreased significantly in 2016, and then kept on an upward trend, reaching the maximum value of 0.211 in 2020. However, Theil

Table 3. Theil Index of Higher Education Resources in China

| Time	Total region	Inter-region		Intra-region		Eastern Regional difference		Central Regional difference		Western Regional difference	
	Theil Index	differences	contribution	differences	contribution	Intra-Regional	Contribution	Intra-regional	Contribution	Intra-Regional	Contribution
2011	0.0635	0.0185	29.09%	0.045	70.91%	0.0186	29.28%	0.007	11.10%	0.0194	30.52%
2012	0.0621	0.0176	28.38%	0.0445	71.62%	0.0181	29.11%	0.0069	11.18%	0.0195	31.33%
2013	0.0622	0.0171	27.49%	0.0451	72.51%	0.0184	29.60%	0.0069	11.02%	0.0198	31.89%
2014	0.0605	0.0162	26.68%	0.0444	73.32%	0.0187	30.96%	0.0068	11.30%	0.0188	31.07%
2015	0.0579	0.0122	21.05%	0.0457	78.95%	0.0184	31.84%	0.0067	11.50%	0.0206	35.61%
2016	0.0589	0.0146	24.84%	0.0443	75.16%	0.0193	32.78%	0.0063	10.73%	0.0186	31.65%
2017	0.0581	0.0141	24.19%	0.044	75.81%	0.0189	32.57%	0.0064	11.10%	0.0187	32.14%
2018	0.0579	0.0129	22.26%	0.045	77.74%	0.0186	32.11%	0.0065	11.24%	0.0199	34.39%
2019	0.0579	0.0122	21.05%	0.0457	78.95%	0.0184	31.84%	0.0067	11.50%	0.0206	35.61%
2020	0.0579	0.0116	19.98%	0.0463	80.02%	0.0183	31.55%	0.007	12.07%	0.0211	36.39%
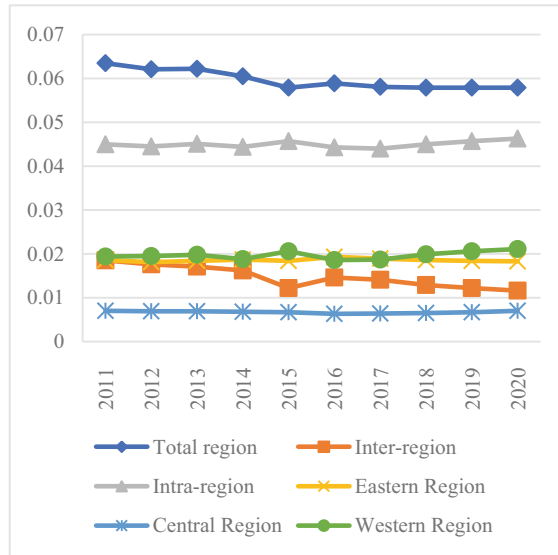
Fig. 4. Theil Index of Higher Education Resources and its Decomposition

index in the central part of China has a relatively flat overall change, showing a slight downward trend from 2011 to 2016, reaching the minimum value of 0.0063 in 2016, and then showing a slight upward trend. This shows that in recent years, the distribution of higher education resources in the east has become more balanced, while the intra-regional differences in the central and western region have increased, especially the change in the western region is more significant. According to the Theil index of the whole country and its decomposition, the Theil index of the whole country decreased significantly from 2011 to 2015, increased slightly in 2016, and then continued to decline steadily from 2020. The variation trend of Theil index between regions was similar to that of the whole country, and the variation of Theil index within regions was relatively gentle. There was a slight fluctuation from 2011 to 2016, and a slight upward trend from then to 2020. At the end of the study period, the Theil index between the whole country and regions decreased compared with the beginning of the study period, and the Theil index within regions increased only slightly, which indicates that the differences between provinces are decreasing on the whole, and the distribution of higher education resources in our country is developing toward a balanced direction.

In general, the distribution of higher education resources in China is not balanced. There are still differences among the eastern, central and western regions, but the intra-regional differences are always the main contributors to the inter-provincial differences. However, as time goes by, the Theil index of the whole country basically decreases year by year. The inter-provincial difference of higher education resources in China is shrinking and developing in a balanced direction.

4 Conclusions

On the whole, according to the information entropy and Theil index analysis results, from the spatial dimension, the distribution of higher education resources in China is unbalanced. There is a big gap among the eastern, central and western regions. Higher education resources are more concentrated in the eastern and central regions, while they are relatively scarce in the west.

From the perspective of time, the equilibrium degree of higher education resources in China basically shows an increasing trend year by year. The gap between the agglomeration degree of higher education in the western belt, the eastern and central belt is gradually narrowing. With the passage of time, the distribution of higher education resources is developing in a balanced and coordinated direction.

In future studies, researchers can look across the country, use more comprehensive and specific data, more accurate and effective methods to analyze the distribution of Chinese higher education resources from internal and external aspects, in order to draw more constructive research conclusions.

In order to make China's higher education resources distribution more balanced and benefiting more people, we should make use of big data technology and information technology to promote the sharing of education resources and break through the unbalanced development of education resources between regions, urban and rural areas. Then gradually realize the balanced development of education resources allocation. By big data technology and information technology, we will strengthen the construction of education informatization in less developed areas. Big data technology and information technology can help education decision-making to achieve targeted poverty alleviation in education. And all social entities work together and establish resource integration and sharing system to realize common governance and sharing, which promote the high-quality development of education in China.

References

1. Yang Sa, Jin Haotian. The 14th Five-Year Plan: Higher education based on the present More will lead the future [EB/OL].
2. Guo Guichun. "Research on Higher Education Resources from the Perspective of Multi-disciplines: A review on the Distribution and Coordinated Development of Higher Education Resources in China," J. Journal of Shanxi University (Philosophy and Social Sciences Edition), 2008(04):119–122.
3. Shao Hui, Wen Mengqi. "Matching Characteristics of Urban Public Education Resources and Population spatial Distribution: A Case study of Beijing," J. Macroscopic Quality Research, 2016, 4(2):119–128.
4. Rao Yingxue, Lin Guodong. "The Rationality of spatial Distribution of urban Educational Resources from the Perspective of supply and demand Balance: A Case study of Wuhan City," J. Journal of South-Central University for Nationalities (Humanities and Social Sciences Edition), 2021, 41(5):147–152.
5. WANG Miao, Guo Yanbin, Xie Yanfeng, et al. "Evaluation and analysis of spatial equity of public education resources," J. Science of Surveying and Mapping, 2020(11).