

# Research on Classification of the Cities in Shandong Province based on the Population

Si-Lian Shen

School of Mathematics and Statistics  
Henan University of Science and Technology  
Luoyang, 471003, China  
shensilian@126.com

Chun-Wei Wang

School of Mathematics and Statistics  
Henan University of Science and Technology  
Luoyang, 471003, China  
wangpersuit@126.com

**Abstract**—Shandong is one of the populous provinces in china. Although it has advantage in population amount, its demographic structure is clearly imbalanced among the cities, which seriously affects the development of economy and society. Based on the data which is from the statistical yearbook of Shandong province, this paper focuses on studying the classification problem of the cities in Shandong province. Specifically, the SAS software is first used to study the spatial characteristics of the total population in Shandong province and partial results are visualized. Secondly, the regional variations among cities are examined from the aspects of household size and population quality (physical and scientific and cultural qualities included), respectively. Through above analysis, some useful information is extracted to evaluate the population situation and to make the population policy.

**Keywords**—Household size, Cluster analysis, Visualization

## I. INTRODUCTION

In recent years, along with the rapid deterioration of the environment and the severe shortage of resources, the population has become one of the hottest issues which we studied. China is a populous country in the world, and Shandong is a populous province in China (The number of resident population in Shandong province is 95.7931 million in the sixth population census). Although Shandong province has advantage in population amount, its demographic structure is clearly imbalanced among the cities, which seriously affects the development of economy and society. Therefore, it is necessary to explore the regional variations of the population throughout the whole province extensively and deeply so that reasonable population planning can be made.

Research on the population problem has been studied since the nineteenth century. But then, most of the literature focuses on the summarization and accumulation of the data which is associated with population while related induction and analysis is rarely carried out. Until the year 1953, professor G.T.Trevartha makes a speech in the annual meeting of United States scholar Association, which marks the beginning of population research. In the following 60's and 70's, people come to realize the importance of the population study, and some statistical methods are regularly applied to analyze this problem (see [1-3] for details). Recently, with the rapid development of science and the universal application of

computer networks, some new models and technologies such as stochastic differential equation, regressive analysis, grey model and artificial neural network are employed to study the population problem of our country (see [4-9] for examples).

Based on the data which is from the statistical yearbook of Shandong province, we focus on in this paper exploring the variations of the population throughout the cities. Specifically, the SAS software is first used to study the spatial characteristics of the total population in Shandong province and partial results are visualized. Then, the spatial patterns of the house-hold size population in the whole province are explored. Finally, the spatial non-stationarity of the physical quality and scientific and cultural quality of the population among the cities are evaluated.

## II. CLASSIFICATION BASED ON THE POPULATION QUANTITY

The sixth population census data shows that the resident population of Shandong province is 95.7931 million, which ranks second, and is next only to Guangdong province. Compared to 90.7931 million in the fifth population census, there increases 5 million people during the past ten years. The total increase rate is 5.5%, and the annual increase rate is 0.55%. This finding indicates that although the total population still tends to be increasing, the rate is under control and relatively slow, which may be caused by the development of the society and economy as well as the improvement of the people's standard of living.

The total population among the cities displays the regional features and the development process of the population in one country is mainly embodied by the regional variations. Furthermore, as time went on, the regional variations may also change. Therefore, it is necessary to study the classification of the cities based on the total population in order to make reasonable population policy.

### A. Data sources

In order to unveil the regional variations of the total population in Shandong province, we choose related population data of 17 cities and these cities are shown in Fig. 1. Also, the population data both in 2000 and 2010 years are taken so that some comparison analysis can be conducted.

---

This research is supported by the National Natural Foundations of China (No. 11326181, No. 11201123), International Cooperative Project in Henan Province (No. 134300510034) and the start fund of doctoral scientific research (No. 09001624).



Fig. 1. Spatial distribution of the cities in Shandong province.

**B. Classification based on the total population**

Based on the total population of the cities in Shandong province of the years 2000 and 2010, cluster analysis is carried out in this subsection by using the SAS software. Due to the similarity of the analysis and the limited space, we only show and analyze the results of four categories. Related results are displayed in Fig. 2 and Fig. 3. Note that color from light to dark in the figures represents the value from small to large.

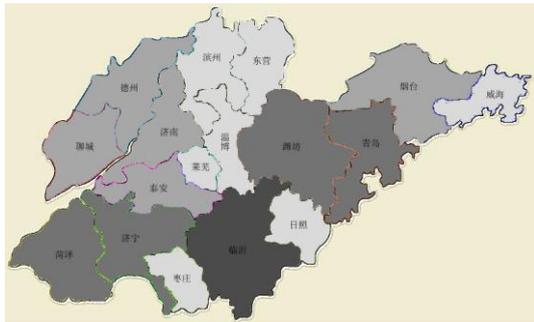


Fig. 2. The clustering results of the cities based on the total population of the year 2000.

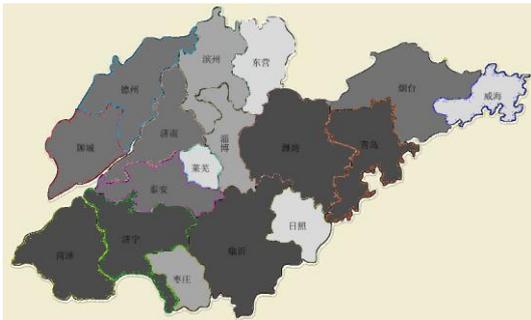


Fig. 3. The clustering results of the cities based on the total population the year 2010.

It can be observed from Fig. 2 that among the seventeen cities in Shandong province, the most populous one in the year 2000 is Linyi, the total population of the cities in Qingdao, Jining, Weifang and Heze ranks second, the total population in Jinan, Taian, Dezhou, Liaocheng and Yantai ranks three and Dongying, Laiwu, Zibo, Zaozhuang, Binzhou, Weihai, Rizhao belong to the fourth populous cities.

Compare to the results of the year 2000 depicted in Fig. 2, we can see from Fig. 3 that, the first and second populous categories are merged, the third populous category remains unchanged. The fourth category is divided into two ones.

In summary, the total population of all cities in Shandong province increases during the ten years, but the increasing rate is slightly different, that is, Binzhou, Zibo and Zaozhuang increases significantly, which may be caused by the rapid development of local society and economics.

**C. Classification based on the house-hold size of the population**

According to the statistical data of the sixth census, there are 8.9855 million household persons and 3.0105 million households. On average, there are 2.98 persons in each family, which is reduced by 0.24 persons compared to 3.22 persons in the fifth census data. This finding indicates that the household size is gradually reducing with the rapid development of the society and economy, which may be caused by the declined fertility levels, increased floating population and dependent living of the young after marriage. In order to unveil the variations of the household sizes among the cities in Shandong province, we compute the household sizes of all cities in the years 2000 and 2010 based on the statistical yearbook. With the software SAS, the cluster analysis is conducted and partial results are visualized in Fig. 4 and Fig. 5.

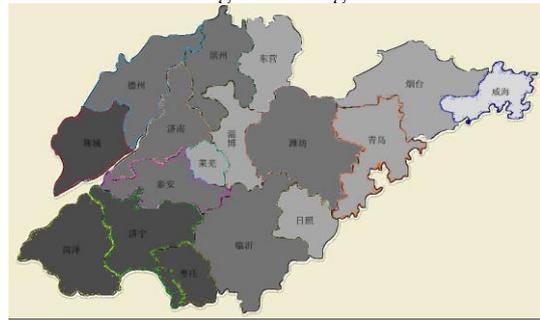


Fig. 4. The clustering results of the cities based on the household size of the year 2000.

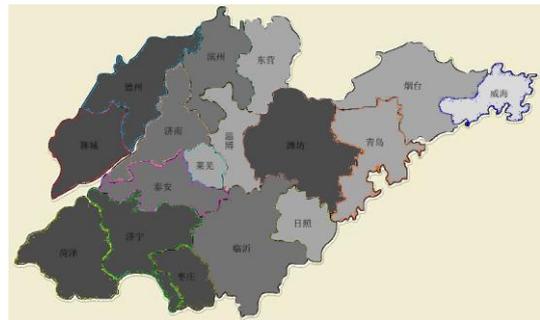


Fig. 5. The clustering results of the cities based on the household size of the year 2010.

From Fig. 4 and Fig. 5, we can observe that, on one hand, the household size of the year 2000 in Liaocheng, Zaozhuang, Jining and Heze ranks first, Linyi, Jinan, Binzhou, Taian, Weifang and Dezhou are the second household size populous cities, Yantai, Rizhao, Laiwu, Qingdao, Zibo and Dongying are the third household size populous ones. The household size in Weihai is relatively small. On the other hand, the clustering results of the year 2010 are similar to those of the year 2000 except that the household sizes in Dezhou and Weifang are increasing significantly. Furthermore, the

household sizes in all cities are decreasing, which may be consistent with the trend of miniaturization in each family.

### III. CLASSIFICATION BASED ON THE POPULATION QUALITY

Generally speaking, population quality consists of the physical quality and Scientific and cultural quality. According to the census data, we define in this section some suitable indices which are used to reflect the population quality and some related data is obtained. Also, the cluster analysis is used to explore the classifications of the cities in Shandong province.

#### A. Classification based on the physical quality

Physical quality generally consists of the structures of ages and sexes. Specifically, the structure of ages is usually measured by some commonly used indices such as children coefficient, labor force coefficient, elder coefficient and so on. While the structure of sexes is quantized by sex ratio (see the literature [10] for details). In this subsection, the values of the children, labor force and elder coefficients as well as sex ratio among all cities are computed with the statistical yearbook data. Also, the cluster analysis is carried out by using the SAS software. Considering the limited space, we take the elder coefficient for example and partial results are shown in Fig. 6 and Fig. 7.



Fig. 6. The clustering results of the cities based on the values of elder coefficient of the year 2000.

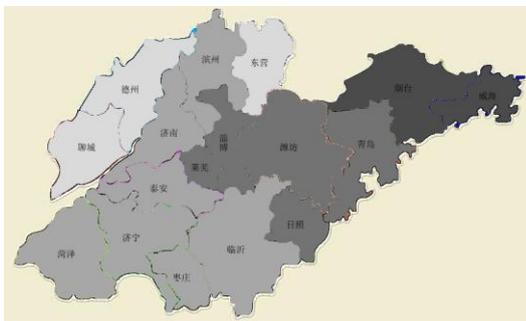


Fig. 7. The clustering results of the cities based on the values of elder coefficient of the year 2010.

It can be seen from Fig. 6 that in the year 2000 the values of the elder coefficient in Qiangdao and Weihai are largest, which implies that the tendency of aging population in the two cities is most serious. The values of the elder coefficient in cities Weifang, Rizhao and Yantai are the second largest, which demonstrates that the tendency of aging population in these cities ranks second. The degree of aging population in

Zaozhuang, Dezhou, Taian, Liaocheng, Jining, Heze, Linyi, Laiwu, Zibo, Jinan and Binzhou ranks third. While the values of the elder coefficient in Dongying are least largest, which shows that the tendency of the aging population in the city is least serious. In summary, the tendency of aging population is becoming more and more serious along the direction from the northeast coast area of Shandong province to southwest area.

Compared to Fig. 6, we find that the variations of the values of the elder coefficient in the year 2010 are similar to those results in the year 2000. Furthermore, the values of the elder coefficient among all cities increase obviously during the ten years, which also indicates that the degree of aging population in Shandong province is becoming more and more serious.

#### B. Classification based on the scientific and cultural quality

According to the sixth census data in Shandong province, we easily find that the numbers of persons with junior middle school education, high school education and college education are all increasing while the number of persons with primary school education is decreasing, which implies that the global schooling throughout Shandong province increases obviously and the population scientific quality improves significantly.

In order to explore the spatial patterns of the scientific and cultural quality of the total population, we quantify in this subsection the scientific and cultural quality by defining the average schooling year. The values of this index in all the cities are computed according to the fifth and sixth census data. Based on the obtained data, we conduct the cluster analysis. Since the clustering results in the year 2000 are exactly the same with those in the year 2010 when the number of category is four, we only show the results of the year 2000 in Fig. 8.

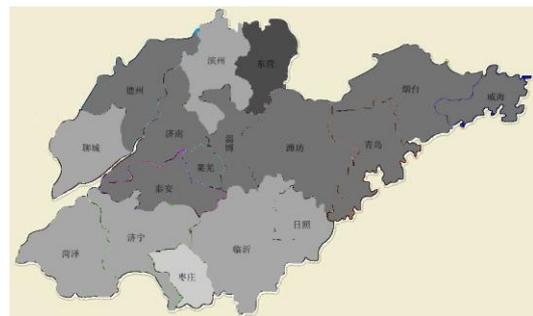


Fig. 8. The clustering results of the cities based on the values of the average schooling year in the year 2000.

We can observe from Fig. 8 that the values of the average schooling year decreases from the northeast coast area to the southwest inland area. Specifically, Dongying, Zibo, Qingdao, yantai, Weihai and Jinan rank first, cities Laiwu, Taian and Weifang rank second, cities Rizhao, Binzhou, Liaocheng, Jining, Zaozhuang and Dezhou rank third, cities Heze and Linyi rank fourth. Furthermore, the values of the average schooling year of the year 2010 are consistently larger than those of the year 2000, which demonstrates that the educational level in Shandong province improves globally and significantly.

## ACKNOWLEDGMENT

Many thanks go to the editor and the reviewers for their valuable comments and suggestions which led to significant improvements on the paper.

## REFERENCES

- [1] S.P. Feng, "On Prediction Model of China Population Growth," *Journal of Anhui Science and Technology University*, vol. 6, pp. 73-76, 2008.
- [2] B.S. Li and L.G. Tong, "The Logistic Model of Population Urbanization in Inner Mongolia and Application," *Journal of Arid Land Resources and Environment*, vol. 2, pp. 32-36., 2007.
- [3] W.Q. Chen, Z. Zhao and D.S. Li, "The Application of Leslie Revised Model in Population Prediction," *World Science-Technology Research and Development*, vol. 2, pp. 219-224, 2008.
- [4] N. Li and H.Q. Hu, "The Prediction of Chinede Stochastic Population. Chinese Journal of Population Scienc," vol. 1, pp. 10-16, 1998.
- [5] N. Li and H.Q. Hu, "The Poulation Forecasting and Project Based on Stochastic Model," *Systems Engineering—Theory Methodology Applications*, vol. 1, pp. 37-41, 1998.
- [6] Y. Fu, "Application of Regression Analysis to Population Prediction," *Liaoning Higher Vocational Technical Institute Journal*, vol. 1, pp. 56-58, 2000.
- [7] X.D. Li, "Study on the Natural Environmental Factors Affecting Population Distribution in the Guizhou Karst Plateau: Multivariate Regression Analysis and Zonality," *Arid Zone Research*, vol. 2 pp. 280-286, 2007.
- [8] A.P. Chen and H.P. An, "A Study on Chinese Population Forecast by Time Series Model," *Population and Economics*, vol. 6, pp. 63-67, 2004.
- [9] S.G. Zhou, "Study on the Grey Predictive Model and Its Application to Chinese Population," *Journal of Mathematical Medicine*, vol. 4, pp. 307-309, 2005.
- [10] Y. L. Li and H. C. Yu, "Spatial dynamic characteristic of the quantity of people in Henan province based on GIS," *Journal of Nanyang Normal University*, vol. 6, pp. 62-65, 2007.