ISSN: 2688-9323

Research on the Construction Level of Urban Infrastructure

Qinghe Guo

School of Hebei GEO University, Shijiazhuang 050000, China

Abstract

Using the above-mentioned multivariate statistical methods and using R language software, this paper analyzes the data related to urban infrastructure construction in 31 provinces and cities in China in 2019, and obtains the classification and model of major Chinese cities based on the level of urban infrastructure construction. Firstly, through the method of cluster analysis, the urban infrastructure construction of 31 provinces and cities is classified and judged. Using discriminant analysis, using the classification results of cluster analysis, randomly select 4 provinces and cities from 31 provinces and cities as the samples to be judged, and use 27 provinces and cities as training samples to judge whether the classification of the samples to be judged is consistent with the cluster analysis. same. Finally, through principal component analysis, the eight original indicators are used to find comprehensive variables that can reflect the original information from various linear combinations.

Keywords

Infrastructure Construction; Principal Component Analysis; Cluster Analysis; Discriminant Analysis.

1. Introduction

Water supply, drainage, roads, traffic, gas, heat, telephone, network and other facilities in urban infrastructure are all indispensable factors in economic production. Urban infrastructure is an important constraint on the quality of life of urban residents. Although the quality of life of urban residents mainly depends on the development of the national economy, comprehensive national strength and the level of per capita GNP, the improvement of urban infrastructure also has an important impact on the quality of life of urban residents. A good urban infrastructure can give impetus to the sustainable development of the urban economy, and its role cannot be ignored. Perfect and good urban infrastructure can make various social and economic units in the city better divide labor and strengthen connections. All aspects of urban infrastructure can quickly transmit the flow of people, logistics and information, and bring together various social and economic elements in the urban area. , greatly improving the urban agglomeration benefits of the organic whole of the economic benefits, urban social benefits and urban ecological environmental benefits of all urban sectors.

The level of urban infrastructure construction is an important aspect to measure urban development, so it needs to be evaluated objectively.

2. Research Method

2.1. Cluster Analysis

Cluster analysis, also known as group analysis, is a multivariate statistical method that studies the classification of samples or indicators. The so-called "class", colloquially, is a collection of similar elements. Cluster analysis is to classify the observed samples (or variables) according to the similarity of the values of the observed samples (or variables), so that the observed samples (or variables) in the same class are similar, and the differences between different "classes" are similar. Observations (or variables) will not be similar.

2.2. Discriminant Analysis

Discriminant analysis is a multivariate statistical method with strong applicability, which has penetrated into various fields. But regardless of the field, the discriminant analysis problem can be described as follows:

There are k m-dimensional populations $G_1, G_2, \ldots, G_k$, whose distribution characteristics are known (for example, the known distribution functions are , $F_1(x), F_2(x), \ldots, F_k(x)$, or it is known from each population training samples). Given a new sample X, we want to determine which population it comes from.

2.3. Principal Component Analysis

Use Z_1 to replace the p variables $X_1, \ldots, X_p$, and use the variance of Z_1 to express the degree to which Z_1 reflects the information of the original p variables. The larger the $Var(Z_1)$, the more information Z_1 contains. To avoid $Var(Z_1) \to \infty$, $a_1{}' a_1 = 1$ should be restricted. If there is a_1 that satisfies the above constraints, so that $Var(Z_1)$ is maximized, Z_1 is called the first principal component (or principal component).

3. Empirical Analysis

3.1. Indicator Selection and Data Description

The data comes from the China Urban Statistical Yearbook 2019. Urban infrastructure includes urban energy and power supply facilities, urban water supply, drainage, sewage treatment facilities, urban road traffic facilities, urban ecological and environmental protection facilities, and urban post and telecommunications facilities. Due to the availability of data, 8 infrastructure construction indicators of 31 major cities in China in 2019 were selected to conduct principal component analysis on the infrastructure construction level of major cities.

3.2. Cluster Analysis

First, using Ward's method, 31 cities were classified. The data were standardized and transformed, and the distance between samples was defined as the Euclidean distance.

According to the comprehensive analysis of the above analysis, it is more appropriate to divide the 31 regions into three categories.

$$G_1^{(3)} = \{Beijing, Shanghai, Zhejiang, Fujian, Guangdong, Jiangsu, Shandong\}$$

$$G_2^{(3)} = \begin{cases} Tianjin, Liaoning\ Shanxi, Xinjiang, Hebei, Inner\ Mongolia, Shanxi, \\ Anui, Jiangxi, HubeiGuangxi, Hunan, Sichuan, Hainan, Yunnan \\ Ningxia, Jilin, HeilongjiangGansu, Qinghai, Henan, Guizhou, Chongqing \end{cases}$$

$$G_3^{(3)} = \{Tibet\}$$

3.3. Discriminant Analysis

4 out of 31 provinces and cities are randomly selected as samples to be judged, and the remaining 27 are training samples. According to the results of cluster analysis, the 27 provinces and cities are divided into three categories, which are regarded as three different populations, and three populations are assumed. The covariance matrix is the same.

Calculate the results, first give the squared distance (ie Mahalanobis distance) between the three groups, which is 25.4927 from class 1 to class 2, 270.9564 from class 2 to class 3, and 228.1818 from class 1 to class 3.

ISSN: 2688-9323

The resulting linear discriminant function is:

$$Y_1(X) = -242.67 + 0.003x_1 + 0.006x_2 + 4.59x_3 - 1.16x_4 + 0.84x_5 - 1.60x_6 + 0.47x_7 + 1.41x_8$$
$$Y_2(X) = -216.01 - 0.003x_1 - 0.07x_2 + 5.39x_3 - 0.97x_4 + 1.20x_5 - 2.35x_6 + 0.35x_7 + 0.76x_8$$
$$Y_3(X) = -120.89 + 0.05x_1 + 0.23x_2 - 0.26x_3 + 0.75x_4 - 1.49x_5 + 1.40x_6 + 0.21x_7 + 2.50x_8$$

The results returned after the discrimination show that the samples from the three types of cities are all correctly judged; the discriminant results for the four samples to be judged are: Zhejiang and Guangdong are the first type of cities, while Chongqing and Sichuan are the second type of cities. The results obtained by clustering discrimination are consistent in Table 8, which proves that the results are valid.

3.4. Principal Component Analysis

Next, standardize the 8 economic indicators, and then perform principal component analysis on the standardized 8 economic indicators. The analysis results are shown in the table below:

Table 1. The analysis results

	Eigenvalue	Difference	Proportion	Cumulative
1	3.38166635	1.64832928	0.4227	0.4227
2	1.73333707	0.57598683	0.2167	0.6394
3	1.15735023	0.59046554	0.1447	0.7840
4	0.56688469	0.08772708	0.0709	0.8549
5	0.47915762	0.15348321	0.0599	0.9148
6	0.32567441	0.05884798	0.0407	0.9555
7	0.26682643	0.17772323	0.0334	0.9889
8	0.08910320		0.0111	1.0000

From Table 1, the cumulative contribution rate of the variance of the first four principal components has reached 85.49%, indicating that the first four principal components basically contain all the information of the indicators. The first four large eigenvalues of the correlation matrix R are:

$$\lambda_1 = 3.382, \lambda_2 = 1.733, \lambda_3 = 1.157, \lambda_4 = 0.567$$

Therefore, the eigenvectors $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ corresponding to the four eigenvalues can calculate the position of the factor loading vector at 8 points on the plane, as shown in Table 2.

It can be seen from the above table that the 8 indicators can be divided into four categories:

The first category is: urban domestic sewage centralized treatment rate (X_5), per capita urban road area (X_6), per capita public green space area (X_7).

The second category is: electricity consumption (X_1) urban gas penetration rate (X_3) bus ownership per 10,000 people (X_4).

The third category is: domestic water consumption per capita (X_2).

ISSN: 2688-9323

The fourth category is: the harmless treatment rate of domestic waste (X_8).

Table 2. The position of the factor loading vector at 8 points on the plane

	z1	z2	z3	z4
x1	0.263131	0.457161	0.286838	-0.637603
x2	-0.087375	-0.250484	0.822133	0.18911
x3	0.392607	0.276892	-0.31338	0.176284
x4	-0.08659	0.635753	0.279096	0.02505
x5	0.442369	-0.036047	-0.067147	0.118627
x6	0.461806	-0.270716	0.0942	-0.273428
x7	0.448347	-0.335512	0.120684	-0.135699
x8	0.389255	0.24547	0.194354	0.647106

4. Conclusion

Urban infrastructure construction is the foundation and premise of urban development, so the evaluation of urban infrastructure construction is of great significance to urban development. Based on the statistical data of infrastructure construction in 31 major cities in 2019, this paper conducts an empirical analysis on the development level of urban infrastructure.

After analyzing the statistical data, in terms of the level of urban infrastructure construction, the major cities in China can be divided into three categories. The third category belongs to cities with less perfect urban infrastructure. According to the classification of urban infrastructure construction and the degree of economic development of the corresponding cities, it can be seen that Beijing, Shanghai, Zhejiang, Jiangsu and other cities are economically developed, and the corresponding urban infrastructure construction is more complete; the economic development of Liaoning, Shanxi and other cities is in the middle level. , the urban infrastructure construction is relatively complete; Tibet is a relatively backward city in economic development, and the corresponding urban infrastructure construction is not perfect; It shows that there is a close relationship between urban infrastructure construction and economic development.

References

[1] Zhang Zhe, Wang Bo, Yang Zhanmin. Evaluation of Urban Infrastructure Construction Level Based on Principal Component Cluster Analysis [J]. Journal of Xidian University (Social Science Edition), 2008, 18(005):27-32.

[2] Zhou Ling, Wang Zhinan. Research on my country's urban infrastructure level based on cluster analysis [J]. Times Finance, 2017, 000(004):201.

[3] Lu Qing. Comparative analysis of infrastructure construction in 19 sub-provincial cities in my country [J]. Gansu Agriculture, 2006(11):49-50.

[4] Li Wei. Cluster Analysis of Urban Facilities in my country [J]. Market Research, 2012(2):27-27.