International Journal of Applied Science and Mathematics
Volume 6, Issue 4, ISSN (Online): 2394-2894

Application of Multivariate Regression of Multifactorial Variables in the Standard Model of the Minimum Living Security - Take Beijing and Shanghai as Examples

Jiayan Niu* and Jingyao Zheng
*Corresponding author email id: 1373929210@qq.com

Date of publication (dd/mm/yyyy): 06/08/2019

***Abstract* – The minimum living security system is a typical way of social relief, providing material assistance to some vulnerable groups of society, which is related to social harmony, stability, fairness and justice. However, the model of accurate subsistence standards has not yet been established, so accurate subsistence standards in each province and city is a problem that needs to be solved. In 2016, this paper collected the minimum subsistence standards, per capita consumption expenditure and per capita disposable income of 30 provinces and cities across the country. The multivariate regression model is used to analyze the standard model and the data is analyzed by using R language and SAS software to provide research value for the establishment of accurate subsistence model.**

***Keywords* – Low Security Standard Cluster Analysis Multiple Regression Model R Software and SAS Software.**

I. INTRODUCTION

By using the common class average method and the longest distance method in the system cluster, the results are consistent, and all provinces (regions) are divided into three categories. The correlation between the same category is stronger and the correlation between different categories is weaker. At the same time, two examples are selected for typical correlation analysis, which verifies our conclusion. Based on the correlation analysis of different regions, we established the multivariate regression model of multivariate variables, and used the gradual regression of multivariate variables and ridge regression respectively to analyze the results of the model fitting. The model is modified by these two methods and the adaptability of the model is tested.

II. ESTABLISHING AND SOLVING THE STANDARD MODEL OF LOW SECURITY

I. *Use Cluster Analysis to Classify 30 Provinces and Cities*

In 2016, we collected the minimum subsistence standards, per capita consumption expenditure, and per capita disposable income of 30 provinces and cities across the country, and conducted a cluster analysis using these three indicators as a sample. Under SAS software, two methods are used for programming. Among the cluster analysis results, the classification results are as follows:

Category	Area
Categ1	Beijing、Shanghai
Categ2	Zhejiang、Tianjing、Guangdong、Jiangsu
Categ3	Liaoning、Neimengu、Hainan、Chongqing、Shanxi、Qinghai、Hebei、Shanxi、Jilin、Heilongjiang、Anhui、Fujian、Jiangxi、Shandong、Henan、Hubei、Hunan、Sichuan、Xizang、Ningxia、Xinjiang、Guangxi、Guizhou、Yunan、Gansu

Therefore, through the cluster analysis, we divide the provinces and cities into three categories in the above table. In order to explain the strong correlation between the same kind and the weak correlation between the different categories, we selected Beijing and Shanghai [1] [2], and carried out a typical correlation analysis of their low insurance standards. To test the reasonableness of the classification results, we collected the minimum insurance standards, per capita disposable income, per capita consumption expenditure in Beijing and Shanghai from 2010 to 2016, and used the R language for typical correlation. It can be seen that the correlation between Beijing and Shanghai is strong, and the correlation coefficients of the three factors are: 0.999, 0.957, and 0.392 [4]. Establishment and solution of multiple regression model.

II. *Establishment and Solution of Multiple Regression Model [3].*

(1) *Multivariate Linear Regression Model and its Regression Diagnosis*

We selected Beijing and Shanghai 2010-2016 data, under SAS software programming analysis. According to the tests of these statistics, we can find that the multivariate linear regression model of multi-factorial variables has a strong multiple collinear relationship, and the value of P is large, and it can not pass the test. The dependent variables and index variables between the same species have strong correlation, Therefore, we consider using multivariate regression and ridge regression to modify the model.

(2) *Multivariate Regression of Multi-Factor Variables [6]*

The idea of multiple gradual regression of multi-factor variables and the gradual regression of a dependent variable are basically similar to the basic steps. The difference is that because the number of dependent variables is greater than one, so in the process of gradual regression, when introducing and removing independent variables, it is necessary to calculate the contribution of these variables to dependent variables. At this time, new statistics (V_i statistics) need to be used to describe them. We selected the relevant data from Beijing and Shanghai for 2010-2016 and analyzed it under SAS software. The results are as follows:

$$\begin{cases} \hat{Y}_1 = -2870.7 + 0.107x_1 \\ \hat{Y}_2 = -2317.28 + 0.2236x_4 - 0.2528x_5 \end{cases}$$

The two regression equations have been tested to be highly significant, and the two regression equations have been tested to be highly significant ($p < 0.05$). Complex correlation coefficient $R_1 = 0.9973$, Decision coefficient ($R_1^2 = 0.9947$). Complex correlation coefficient $R_2 = 0.9999$ ($R_2^2 = 0.9998$). Therefore, the regression equation is tested and the fitting effect is better.

(3) *Ridge Return*

When multivariate linear regression is performed on the equation, we find that there are multiple eigenroots of the least squares matrix $X'X$ that are close to 0. [7] The multiple linear relations between variables are strong. The basic idea of ridge regression is to add a positive matrix to the matrix. The constant matrix $k*I$, the degree to which $X'X + kI$ is close to singularity will be significantly reduced, thus overcoming the collinear relationship between variables. With SAS software programming, the ridge trace map is obtained. When $k \approx 0.01$, the ridge trace has stabilized, thus overcoming the multiple collinear relationships between variables and the fitting effect

is better. The regression equation was obtained and the predictive effect of the SAS software test model was better, as follows:

$$\begin{cases} \hat{Y}_1 = 0.052x_1 + 0.006x_2 + 0.0149x_3 + 0.03x_4 + 0.08x_5 - 1449.62 \\ \hat{Y}_2 = 0.032x_1 + 0.038x_2 + 0.0449x_3 + 0.04x_4 + 0.07x_5 - 2042.97 \end{cases}$$

III. Conclusion

As we can see from this article, [5] the establishment of an accurate standard model for subsistence insurance needs to take into account many factors, which has brought difficulties to the work of precision poverty alleviation. The study of subsistence standards is conducive to the formulation of dynamic adjustment methods for subsistence standards. It is conducive to the improvement and development of the social assistance system and poverty alleviation policies.

References

[1] Beijing Economic and Social Development Statistical Bulletin [J]. Beijing Bureau of Statistics 2010-2016.
[2] Shanghai Municipal Economic and Social Development Statistical Bulletin [J]. Shanghai Municipal Bureau of Statistics 2010-2016.
[3] Gao Huiyi. Undergraduate mathematics basic course teaching materials, applying multiple statistical analysis [M]. Peking University Press, 2005.
[4] Chenchuanbo. Urban and rural residents 'minimum subsistence needs and minimum subsistence standards [J]. Agricultural Economic Issues, 2011(8): 38-44.
[5] Hao Zhichong. The causes of urban poverty and the effect of poverty relief policies [D]. Jilin University, 2007.
[6] Song Yangyang. Regional Economic Research Based on Multiple Statistical Analysis [D]. Value engineering, 2017.
[7] Zhu Shangwei. Two expected constraints of ridge regression parameters [D]. Statistics and Decision-Making, 2015.

Authors Profile'

First A author: **Jiayan** 1996-
Student study in the department of mathemathic in Yanbian University (Yanji,China). email id: 1373929210@qq.com

Second author: **Jingyao Zheng**
Student study in the department of mathemathic in Yanbian University (Yanji,China).