Research on Comprehensive Development of Shandong Province Based on Bayesian Quantile Regression

Mengdi Guo¹, Xiaoshan Zhao²*, Yang Ge³
College of Science, Tianjin University of Technology and Education, Tianjin 300222, China
786576956@qq.com

Abstract. Based on the theme of coordinated development in Shandong Province, the article proposes coordinated indicators based on the three aspects of economy, society and ecology, and uses the principal component analysis method and entropy method to integrate the three aspects of economy, society and ecology in Shandong. The indicators were selected. Then the Bayesian quantile model was used to construct an evaluation model and the coordinated development of Shandong Province was evaluated. It shows that Shandong Province needs to vigorously develop ecology and vigorously develop indicators that affect Shandong Province's great changes if it wants to rapidly increase its comprehensive strength.

1. Introduction
Today in modernization, evaluating the development of a region is not just evaluating the development of the economy. Coordinated development is one of China's major international strategies and occupies a major position in the development of China new era. Studying and evaluating the coordinated indicators of economic, social, and ecological development in Shandong Province can see the current status of coordinated development in Shandong Province, which is meaningful for further development.

R Koenker, GS Bassett [1] proposed the basic principle of quantile regression, and Yu K, Moyeed R A [2] proposed a Bayesian quantile model based on it. After that, many scholars used the Bayesian quantile regression model to study the economy. Tian Maoqian and Yu Keming [3] analyzed the gender gap in the income of urban residents in China, and found that narrowing the income gap between men and women needs to be increased. Investing in women's education to increase women work experience; Mei Bo and Tian Mao [4] also used Bayesian quantile model to study the PM2.5 concentration in Beijing, and found that the predicted space is similar to the real space. Help to understand the reasons for the formation of smog; Xiao Beifang [5] used Bayesian quantile model to study the relationship between the urban-rural gap and the financial economy, and found that the relationship between financial development and urban-rural relationship has an inverted "U" relationship; A Kottas, AE Gelfand [6] proposed a Bayesian parameter median regression model; A Kottas, M Krnjajic [7] proposed Bayesian semi-parameter median modeling quantile regression; Mohd Fadzli Mohd Fuzi, Abdul Aziz Jemain [8] introduced the Bayesian quantile regression model of claim count data, and believes that rating factors have a lower risk for higher claim counts. These articles use quantile regression to conduct empirical research and analysis. In this article, three indicators is economy, society, and ecology, and a series of variables such as GDP and science and technology are used to evaluate Shandong the comprehensive development was studied. Principal component
analysis method and entropy method to select variables, and use Bayesian quantile regression to evaluate the coordinated development of Shandong Province.

2. Selection of indicators

The data used comes from the official website of the Shandong Provincial Bureau of Statistics, and the data from the "2019 Shandong Statistical Yearbook" is selected.

In table 1, this article selects three comprehensive indicators based on various factors: economic, social, and ecological.

| Economic            | Social          | Ecological      |
|---------------------|-----------------|-----------------|
| regional GDP        | total population| energy consumption |
| GDP per person      | Urban population| Urban landscaping |
| Primary industry of GDP | Rural population | water consumption |
| Secondary industry of GDP | Number of employees | Ecological water |
| Tertiary industry of GDP | people buy insurance | Chemical oxygen emissions |
| fixed asset investment | Construction land | Waste water |
| imports and exports | Residential area | Ammonia nitrogen emissions |
| General public budget revenue | Highway mileage | Sulfur dioxide emissions |
| Disposable income | students enrolled in schools | Nitrogen oxide emissions |
| Consumption expenditure | Number of civilian cars | Tobacco dust emission |
| Rural residents' disposable income | domestic patent application | general industrial waste |
| Rural household consumption expenditure | Expenditure on colleges and universities | Afforestation area |
| agriculture forestry, animal husbandry fishery output value | traffic accidents | fire accidents |
| industrial enterprises above designated size | beds in hospital | |
| Total sales of wholesale and retail goods above designated size | medical institutions | |
| Total retail sales of consumer goods | |

3. Research methods

This paper selects principal component analysis[9] and entropy method to sort and select the indicators, and then uses Bayesian quantile regression to evaluate and analyze the comprehensive indicators and variables.

3.1. The Bayesian inference for quantile regression
Let random disturbance terms $\varepsilon_i$ be independently and identically distributed in $ALD(0, \sigma, \tau)$, given sample data $y = (y_1, y_2, \ldots, y_n)$, then the likelihood function of $y$ is

$$L(y \mid \beta, \sigma) = L(y_1, y_2, \ldots, y_n \mid \beta, \sigma) = \frac{\tau^n (1 - \tau)^n}{\sigma^n} \exp \left\{ -\sum_{i=1}^{n} \frac{y_i - x_i' \beta}{\sigma} \right\}$$  \hspace{1cm} (1)

Coefficient vector $\beta$ independent of scale parameter $\sigma$. A posteriori probability of joint priority allocation is $\Pi(\beta, \sigma) = \Pi(\beta) \Pi(\sigma)$ \hspace{1cm} (2), the joint posterior distribution is $\Pi(y \mid \beta, \sigma) \propto L(y \mid \beta, \sigma) \Pi(\beta) \Pi(\sigma)$ \hspace{1cm} (3).

4. Data processing results and regression analysis

4.1. Data processing results

At first, in order to avoid the loss of data indicators, we use the principal component to analyze the data. First standardize the three general indicators of economy, society, and ecology. From Figure 1 we can see that the economy extracted four principal components; society extracted two principal components; ecology extracted six principal components. Let $y_{1,1}, y_{1,2}, y_{1,3}, y_{1,4}$ represent the first, second, third, and fourth principal components of the economy, respectively. The first, second, third, and fourth principal components respectively represent the level of economic development, the level of production, the scale of major output value industries, and asset investment.

Let $y_{2,1}, y_{2,2}$ represent the first and second principal components of society, respectively. The first principal component represents the level of economic development; the second principal component represents the level of production; the third principal component represents the main output value industry scale; the fourth principal component represents asset investment.

Let $y_{3,1}, y_{3,2}, y_{3,3}, y_{3,4}, y_{3,5}, y_{3,6}$ represent the first, second, third, fourth, fifth, and sixth principal components of ecology, respectively. The first, second, third, fourth, fifth, and sixth principal components represent the impact of waste discharge, resource utilization, environmental construction, green area, disaster status, and income on ecology.

The variance contribution rate of each principal component is weighted, and the comprehensive score of Shandong Province's economy, society and ecology is obtained by weighting, that is

$$Z_1 = 0.562y_{1,1} + 0.204y_{1,2} + 0.084y_{1,3} + 0.054y_{1,4}$$

$$Z_2 = 0.744y_{2,1} + 0.174y_{2,2}$$

$$Z_3 = 0.351y_{3,1} + 0.230y_{3,2} + 0.140y_{3,3} + 0.082y_{3,4} + 0.061y_{3,5} + 0.048y_{3,6}$$  \hspace{1cm} (4)

Figure 1: Principal component analysis

Using the comprehensive score, the entropy method is used to calculate the proportions of the three comprehensive indicators of economic, social, and ecological in the comprehensive indicators: 0.3335, 0.3319, and 0.3347. The expression of the comprehensive evaluation index $M$ of Shandong Province is:

$$M = 0.3335Z_1 + 0.3319Z_2 + 0.3347Z_3$$  \hspace{1cm} (5).
4.2. regression analysis

Bayesian quantile regression was performed with principal component M as the dependent variable and $Z_1, Z_2, Z_3$ as the independent variable. The MCMC sampling algorithm generates parameter sampling samples and simulates 4000 iterations. Iterative graph of MCMC parameters at a level of 0.2 is given, as shown in Figure 2-5, MCMC is stable.

Because the MCMC parameter iterative graph converges, a Bayesian quantile regression analysis of the parameters is performed. The following gives estimates of the principal component coefficients at 0.2 quantile, 0.5 quantile and 0.8 quantile.

In the table2, the estimated regression equation for each quantile can be obtained, for example, the regression equation at the 0.5 quantile level is:

$$M = -0.00772 + 0.32771 Z_1 + 0.33807 Z_2 + 0.32853 Z_3 \ (6)$$

| Table.2 Coefficient estimates at each quantile level |
|----------------|-----|-----|-----|
| OLS  | 0.2  | 0.5  | 0.8  |
| Intercept | 0.727 | 0.008 | 677 |
| $Z_1$ | 0.3 | 0.2 | 0.32 |
| $Z_2$ | 335 | 75 | 771 | 411 |
| $Z_3$ | 0.3 | 0.3 | 0.33 |
| 319 | 16 | 807 | 344 |
| 347 | 38 | 853 | 359 |

By bringing the expressions of variables of each principal component into the regression equation of quantiles, the coefficient estimates of each indicator variable can be obtained, and the influence of
In the economic indicators, except for the primary industry's GDP and the secondary industry's GDP, the coefficients of other variables are estimated to be positive numbers, which play a positive role in the comprehensive development of Shandong Province. The economic indicators that significantly affect the comprehensive strength of Shandong Province in the upper middle Big. Therefore, economic variable indicators should be vigorously developed when considering the economic aspects of comprehensive strength development in Shandong Province.

In social indicators, with the improvement of the median level of spending on medical institutions, and the number of medical institutions showed an increasing trend, and increased significantly at the high end. The degree of increase was greater than the average. Therefore, in medical institutions, the median level of spending on medical institutions, and the number of medical institutions showed a significantly better impact on the comprehensive strength of Shandong Province at the 0.3 and 0.8 quantile levels have a significantly better impact on the comprehensive strength of Shandong Province than the 0.2 quantile levels have a significantly better impact on the comprehensive strength of Shandong Province. The economic indicators variables that significantly affect the comprehensive strength of Shandong Province at the 0.1, 0.2, and 0.3 quantile levels have a significantly better impact on the comprehensive strength of Shandong Province than the 0.4 quantile levels have a significantly better impact on the comprehensive strength of Shandong Province. The economic indicators variables that significantly affect the comprehensive strength of Shandong Province at the 0.5, 0.6, and 0.7 quantile levels have a significantly better impact on the comprehensive strength of Shandong Province than the 0.8 quantile levels have a significantly better impact on the comprehensive strength of Shandong Province.
that of other variable indicators, indicating that the cities located in the middle and lower levels should increase the development of these three areas.

In the ecological indicators, except for the difference between the energy consumption per 10,000 yuan of GDP in each city and 2017, the other variable indicators all decrease with the increase of the quantile level, indicating that except the difference between the energy consumption per 10,000 yuan of GDP in each city and 2017 the development of China has a huge effect on the comprehensive strength of Shandong Province.

5. Conclusion

Through the above analysis, it can be seen that the development of each prefecture and city is uneven and has large differences. In order to comprehensively develop Shandong Province and narrow the development gap between cities, it is necessary to vigorously develop aspects that are faster for one's own ability in light of the actual situation of each city and the existing resource advantages.

ACKNOWLEDGMENTS

The author thanks all anonymous reviewers for their careful reading of this article and their constructive suggestions for the revision of the paper. This work was supported by the Tianjin Natural Science Foundation (No.18JCQNJC70100), the national nature science foundation of China (No.61703307), and the Tianjin nature science foundation of China (No.17JCYBJC15700), and the nature science foundation of Tianjin university of technology and education (No.KJ1814).

REFERENCES

[1] R Koenker, GS Bassett. Regression Quantiles[J]. Econometrica, 1978, 46(1):33-50.
[2] Yu K, Moyeed R A. Bayesian quantile regression[J]. Statistics&Probability Letters, 2001, 14:437-447.
[3] Tian Maoqian, Yu Keming. Gender Difference Analysis of Human Capital Return Rate of Chinese Urban Residents Based on Bayesian Quantile Regression[J]. Journal of Shanxi University of Finance and Economics, 2013, 3509:1-10. (in Chinese)
[4] Mei Bo, Tian Maozai. Bayesian space-time quantile regression model and its research on PM2.5 concentration in Beijing[J]. Statistical Research, 2016, 33 (12): 91-100. (in Chinese)
[5] Xiao Beifang. Empirical Application of Bayesian Quantile Regression Method Based on MCMC Algorithm[D]. Hunan: Hunan Normal University, 2015:1-30. (in Chinese)
[6] A Kottas, AE Gelfand. Bayesian semiparametric median regression modeling[J]. Journal of the American Statistical Association, 2001, 456:1458-1468.
[7] A Kottasa, Alan E Gelfanda. Bayesian semiparametric median modeling inquantile regression[J]. Scandinavian Journal of Statics, 2009, 36:297-319.
[8] Mohd Fadzli Mohd Fuzi, Abdul Aziz Jemain, et al. Bayesian quantile regression model for claim count data[J]. Mathematics and Economics, 2016, 66:124-137.
[9] Long Peng, Guoqing Han. Electric submersible pump broken shaft fault diagnosis based on principal component analysis[J]. Journal of Petroleum Science and Engineering, 2020, 191:1-6.