Analysis on the Per Capita Disposable Income of Sichuan Residents Based on Grey Forecasting Extended Model

Yaning Xiao¹, Dongchen Wu²

¹Faculty of Science, North China University of Science and Technology, Tangshan, China
²School of Mathematics and Statistics, Xi’an University, Xi’an, China

Abstract: In recent years, the development of Sichuan Province has been advancing by leaps and bounds. From 2015 to 2020, the per capita disposable income of residents in Sichuan Province has reached 22,461 yuan, an actual increase of 99.3% year-on-year. In order to predict the future disposable income and provide a basis for formulating policies, GM (1,1) has become an important method. Many existing GM models can only be iterated based on the original data, and cannot add the predicted data to the original data in time. Iterate in. This article predicts the new information GM(1,1) model of metabolism GM(1,1) and compares it with traditional GM(1,1). The experimental results show that the metabolic GM(1,1) has the highest accuracy and is most suitable for predicting the per capita disposable income of Sichuan residents.

Keywords: Sichuan residents, per capita disposable income, new information GM(1,1), metabolism GM(1,1)

1. Introduction

Today in our country with rapid changes, everything is developing rapidly, and the quality of our residents' material life is also growing very fast. The per capita disposable income of residents has a certain gray scale, and the data is more complicated. Common methods are not conducive to analysis and may cause large errors. Therefore, we can choose gray forecasts for forecasting. In recent years, researchers have analyzed and studied the per capita disposable income of residents through VAR model, Granger causality test, data mining, principal component analysis, logistic regression model, entropy method, and multiple linear regression analysis. The method is to study the influencing factors separately or predict the per capita disposable income of residents separately, and the research question is relatively single.

The information of the gray system is partly known and partly unknown, and there are uncertain relationships among various factors in the system. These characteristics are very suitable for the data analysis of this paper, so this paper chooses the gray forecasting method to predict. Gray forecasting method is a forecasting method for predicting gray system. Gray forecasting is through the correlation analysis between various factors of the system, and searching for the law of system changes after processing the data to be analyzed, that is, accumulating the original data to generate (or Generated by other methods) to obtain an approximate exponential law and then model the method. Obtain a data column with relatively strong regularity, and then establish a corresponding mathematical model to predict the development trend of things. The new information model mainly refers to adding predicted data to the original data to predict the next data. Metabolism model means that the predicted data is added to the original data, and the first element of the data is eliminated at the same time.

This paper uses traditional GM (1,1), metabolic GM (1,1), and new information GM (1,1) to predict the per capita disposable income of Sichuan residents, and compare them, and select the model with the smallest error for 2021 - 2025 The per capita disposable income of residents is forecasted to help relevant departments make decisions.
2. Model establishment and testing

2.1. Model establishment

G means grey (gray), M means model (model)

The gray derivative of is defined as 
\[ d(k) = x^{(0)}(k) = x^{(1)}(k) - x^{(1)}(k-1) \]

Let \( z^{(1)}(k) \) Sequence \( x^{(1)} \) Neighbor value generation sequence, which is \( z^{(1)}(k) = \alpha x^{(1)}(k) + (1-\alpha)x^{(1)} \)

Define \( GM(1,1) \) the grey differential equation model is 
\[ d(k) + \alpha z^{(1)}(k) = b \]

\( x^{(0)}(k) \) Grey derivative, \( \alpha \) Development coefficient, \( z^{(1)}(k) \) It is called the whitening background value, and \( b \) is called the ash effect 2,3,...

Introducing matrix vector notation: 
\[
\begin{bmatrix}
  a \\
  b
\end{bmatrix} \quad Y =
\begin{bmatrix}
  x^{(0)}(2) \\
  x^{(0)}(2) \\
  \vdots \\
  x^{(0)}(n)
\end{bmatrix}
\]

\[
B =
\begin{bmatrix}
  -z^{(1)}(2) & 1 \\
  -z^{(1)}(3) & 1 \\
  \vdots & \vdots \\
  -z^{(1)}(n) & 1
\end{bmatrix}
\]

Then the model can be expressed as 
\[ Y = Bu. \]

Then use unary linear regression to find the values of \( a \) and \( b \)
\[
\begin{bmatrix}
  a \\
  b
\end{bmatrix} = (B^T B)^{-1} B^T Y
\]

In order to ensure the feasibility of the modeling method, it is necessary to check the known data.

Let the original data be listed as 
\[ x^{(0)} = (x^0(1), x^0(2),...x^0(n)) \]

Calculate the grade ratio of the sequence \( \lambda(k) = \frac{x^{(0)}(k-1)}{x^{(0)}(k)} \rightarrow k = 2,3,...,n \)

To build a model, all grade ratios need to fall within the acceptable coverage range 
\[ X = \left( e_{n+1}, e_{n+1} \right). \]

Only in this way can grey forecasts be made. Otherwise, the data needs to be processed, such as translation transformation:

\[ y^{(0)}(k) = x(0)(k) + c, k = 1,2,...,n \]

Take \( c \) so that the rank ratio of the data column falls within the acceptable coverage.

2.2. Establish \( GM(1,1) \) model

Assume \( x^{(0)} = (x^{(0)}(1), x^{(0)}(2),...x^{(0)}(n)) \) to meet the above requirements, use it as a data column to build a model: 
\[ x^{(0)}(k) + \alpha z^{(1)}(k) = b \]

The estimated value obtained by regression analysis, so the corresponding whitening model is
\[
\frac{dx^{(1)}(t)}{dt} + ax^{(1)}(t) = b
\]

Solution to
\[
x^{(1)}(t) = \left(x^{(0)}(1) - \frac{a}{b}\right)e^{-at} + \frac{a}{b}
\]

So get the predicted value
\[
x^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{a}{b}\right)e^{-ak} + \frac{b}{a}, k = 1, 2, ..., n - 1
\]

So as to get the predicted value accordingly
\[
x^{(0)}(k + 1) = \hat{x}^{(1)}(k + 1) - \hat{x}^{(1)}(k), k = 1, 2, ..., n - 1
\]

2.3. Test predicted value

(1) Residual error test: calculate relative residual error
\[
\varepsilon(k) = \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)}, k = 1, 2, ..., n
\]

If for all \(|\varepsilon(k)| < 0.1\), it is considered that the higher requirements are met; otherwise, if all the \(|\varepsilon(k)| < 0.2\), it is considered to meet the general requirements.

(2) Level ratio deviation test

Calculate
\[
\rho(k) = 1 - \frac{1 - 0.5a}{1 + 0.5a}\hat{x}(k)
\]

If for all \(|\varepsilon(k)| < 0.1\), it is considered that higher requirements are met; otherwise, if for all \(|\varepsilon(k)| < 0.2\), it is considered to meet the general requirement

3. Empirical analysis

1) Collect data

![2015-2020 Statistics of Per Capita Disposable Income of Urban and Rural Residents in Sichuan Province](image)

Figure 1: 2015-2020 Statistics of Per Capita Disposable Income of Urban and Rural Residents in Sichuan Province
2) Perform a quasi-exponential law test on the data

Index 1: Data with smooth ratio less than 0.5 accounted for 60%

Index 2: Except for the first two periods, the proportion of data with a smoothness ratio less than 0.5 is 100%

The above two indicators meet the quasi-index test law.

3) Using three GM(1,1) models for training

![Figure 3: Comparison of three different GM (1,1) models](image)

<table>
<thead>
<tr>
<th>Model type</th>
<th>Sum of squared errors for the experimental group</th>
</tr>
</thead>
<tbody>
<tr>
<td>GM(1,1)</td>
<td>460852.5496</td>
</tr>
<tr>
<td>New information GM(1,1)</td>
<td>413592.184</td>
</tr>
<tr>
<td>Metabolism GM(1,1)</td>
<td>369927.6214</td>
</tr>
</tbody>
</table>

Because the metabolic GM(1,1) model has the smallest sum of squares, we should choose it for the next five years of forecasting.

4) Using metabolism GM(1,1) to make predictions
The results are calculated using the model, and the following prediction data are obtained:

Figure 4: Forecast of per capita disposable income from 2021 to 2025

4. Conclusion

Aiming at the application of different GM(1,1) models to solve the prediction problem, this article will GM(1,1), new information GM(1,1), metabolism GM(1,1) three models for Sichuan residents per capita Disposable income analysis, predicting the per capita disposable income of Sichuan residents in 2021-2025, and drawing the following conclusions

(1) To analyze the per capita disposable income of Sichuan residents, establish three GM(1,1) models to predict. By comparing the errors of different models, select the metabolic GM(1,1) model to predict. The following results:

(2) The factor of per capita disposable income of Sichuan residents is restricted by many complicated factors, such as the policy arrangement of the year, the situation of the epidemic situation, and the consumption status of residents, etc., which have a relatively large degree of gray. This article uses the metabolic GM(1,1) model to predict per capita disposable income from 2021 to 2025. The results show that there will be an upward trend afterwards, and the government can make corresponding policies according to the size of the increase.

Although this article compares and analyzes the three GM(1,1) models, although the analysis problem is more sufficient, the accuracy of the result is less improved. It should be combined with other machine algorithms to get a detailed relational expression and carry out a detailed analysis of the economic situation analyze.

References

2004(02): 8-12