

Applied Mathematics and Nonlinear Sciences

<https://www.sciendo.com>

Tourists Forecast Lanzhou Based on the Baolan High-Speed Railway by the Arima Model

Lijun Li^{a†}, Yaru Wang^a, Xiaoyue Li^b*a.* School of Economic & Management, Shijiazhuang Tiedao University, Shijiazhuang 050043, China*b.* Fisher College of Business, Ohio State University, Columbus OH43210, USA

Submission Info

Communicated by Juan Luis Garcia Guirao

Received December 3rd 2019

Accepted January 14th 2020

Available online February 28th 2020

Abstract

According to the analysis from the number of tourists who went to Lanzhou during 2009–2019, the ARIMA model of the number of tourists to Lanzhou was established. The results show that the AR(3) model is used to predict the number of tourists who traveled to Lanzhou during 2009–2019. The average relative error between the predicted value and the actual value is 1.03%, which can be used to predict and analyze the number of tourists in Lanzhou in the future.

Keywords: Baolan high-speed railway; Lanzhou; time series; rational; prediction.**AMS 2010 codes:** 91B74; 62M20; 65G50

1 Introduction

With the expansion of high-speed rail tourism and the improvement of benefits, travel management and decision-making work have also increased many difficulties. To ensure a sustainable and stable development of tourism, scientific methods must be used to manage, analyze, and predict the rules of the tourism market, which has far-reaching significance to the sustainable development of China's tourism industry. Therefore, this paper takes Baolan high-speed railway as an example and uses ARIMA model to compare the passenger flow data and forecast results from 2009 to 2019 to verify the accuracy of the model. By forecasting the passenger flow to Lanzhou in 2020, it can provide reference for the tourism planning and the layout of tourism resources in Lanzhou.

[†]Corresponding author.Email address: 617596186@qq.com

2 The research status of using ARIMA model to predict passenger volume

The opening of high-speed rail has promoted the economic growth of cities along the periphery, but it also brought challenges to developmental opportunities. The surge in the number of people will naturally lead to fullness of the scenic spots and the increase of accommodation and accommodation pressures, especially the problems of hotel accommodation and transportation, which have brought severe challenges to the local tourist reception capacity. How to predict passenger traffic and enhance local hardware and software facilities and increase the carrying capacity according to passenger traffic has become an urgent problem to be solved.

According to the relevant literature, Li Wei [1] believes that for the ARIMA model and the BP neural network model of predictable tourist numbers, the simulation prediction of the ARIMA model is better overall, while the BP neural network model is relatively less. Perhaps, Kong Chaoli [2] believes that with the deterministic factor decomposition model decomposition, the ARIMA model preview effect is better. Liu Ruoyu and Liu Libo [3] believe that compared with regression analysis, which needs basic function model for prediction and estimation, human factors have less influence on time series processing in ARIMA model. Therefore, according to the traditional regression analysis method, the ARIMA model has better precision. It can be seen that in the process of prediction, many researches do not involve the unit root test and white noise residual test, which leads to low reliability [4]. Based on the above research, this paper uses ARIMA model system and Eviews statistical software to process and analyze the number of tourists to Lanzhou during 2009–2019, and make short-term forecast.

3 The data source and description

(1) Data source

This paper uses the ARIMA model to predict the number of tourists to Lanzhou during 2009–2019. The data variable “Lanzhou national economic and social development statistical bulletin,” using Eviews8.0 for comprehensive analysis, sample capacity $n = 11$.

(2) Data description

According to the relevant data, the change in the number of tourists to Lanzhou in the autumn of 2009–2018 is shown in Figure 1. The increase is shown in Figure 2.

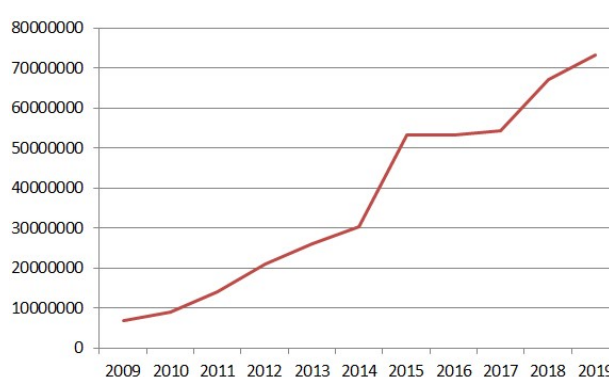


Fig. 1 Number of tourists in Lanzhou City during 2009–2019.



Fig. 2 The increase in the number of tourists in Lanzhou City during 2009–2019.

4 The establishment of the ARIMA model for the number of tourists in Lanzhou

4.1 Selection of the ARIMA model

4.1.1 The reason of forecasting the number of tourists with ARIMA model

As an important industrial base and comprehensive transportation in northwestern China, passengers who choose to travel to Lanzhou will also be affected by traffic, terrain, climate, culture, and so on. Some of these factors are difficult to change, and the relationship between various factors is complicated. It is difficult to predict by regression analysis and factor analysis, and it is difficult to achieve the desired prediction effect. Time series analysis is a statistical method of dynamic data processing. By analyzing some statistics of the data, the dynamic change law of the data is obtained, and the model is established by the boots for prediction [5]. Among those, including a variety of predictive models, the ARIMA model is one of them.

The ARIMA model is a commonly used random time series model, founded by Box and Jenkins. It is a more accurate continuous short-term prediction method. It can better understand the structure and characteristics of time series and the optimal prediction under the minimum variance [6]. So in real life, we often use the ARIMA model to predict and study time series, and get multiple ideal results [7].

4.1.2 The basic theory of ARIMA model

A system, if its response X_t at time t , is related to the previous value of the previous time, and there is a certain dependence on the disturbance of the previous time into the system, then this system is the autoregressive moving average system, corresponding. The model can apply the ARMA to the discontinuous time series containing the short-term trend, and then apply the ARMA to establish the differential autoregressive moving average model. The corresponding model is recorded as the ARIMA (p, d, q) model. Where AR is autoregressive, p is autoregressive, MA is the moving average, q is the moving average, and d is the number of differences when the time series becomes stationary.

4.2 Data preprocessing

The stability and randomness of the sequence are tested before modeling, and a general sequential model is established based on the sequence stability and non-pure randomness. First, test the stationary of time series with ADF. The test results are shown in Figure 3 and can be trimmed at a significance level of 0.05, accepting the presence of a column, $d = 1$, the sequence is first-order difference stationary sequence.

4.3 Model ordering and testing

According to the first-order difference autocorrelation and partial autocorrelation graph (Fig. 4), the model can be preliminarily judged as AR model, and the value of AR is realized by Eviews software. The AIC

Null Hypothesis: D(X) has a unit root		
Exogenous: Constant, Linear Trend		
Lag Length: 0 (Automatic - based on SIC, maxlag=1)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-5.322360	0.0126
Test critical values: 1% level	-5.521860	
5% level	-4.107833	
10% level	-3.515047	

Fig. 3 ADF test with time term and intercept term after first-order difference.**Table 1** AIC equivalence of different AR values

	AIC	SC	HQ
AR(3)	35.06211	35.04297	34.94810
AR(2)	35.25883	35.40352	35.16763
AR(1)	35.45157	35.56009	35.38317

equivalents in several cases are compared (Table 1), finally the AR (3) model is chosen.

Date: 12/31/19 Time: 11:30
Sample: 2009 2019
Included observations: 11



Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.726	0.726	7.5367	0.006
		2 0.452	-0.159	10.782	0.005
		3 0.224	-0.088	11.681	0.009
		4 -0.029	-0.246	11.699	0.020
		5 -0.189	-0.037	12.553	0.028
		6 -0.316	-0.161	15.414	0.017
		7 -0.408	-0.125	21.369	0.003
		8 -0.402	-0.029	29.070	0.000
		9 -0.353	-0.053	37.996	0.000
		10 -0.204	0.129	43.957	0.000

Fig. 4 Autocorrelation and partial autocorrelation after second-order difference.

Arch test is a method to observe the variance change of time series. Its idea is that in time series data, we can consider the existence of heteroscedasticity as arch process, and judge whether there is heteroscedasticity in time series by testing whether this process is tenable. Perform an ARCH test on the model to determine the lag order, as shown in Table 2. According to the Akaike info standard (AIC), the lag order can be determined to be 1. In the ARCH(1) test, the Ps value corresponding to the $\text{Obs} * R^2$ statistic is at the significance level of 5%, and the null hypothesis is not rejected, that is, the test result indicates that there is no ARCH effect, that is, there is no heteroscedasticity of the ARCH form. The statistics of the model mean and the autocorrelation coefficient are all tested by the significance test. The model fully extracts the effective information through the residual autocorrelation test and can be used for prediction [8].

4.4 Model prediction

We use time series analysis to establish an autoregressive prediction model for the data of tourists to Lanzhou during 2009–2019. The predicted and actual values are shown in Table 3.

It can be seen from the table that the model is accurate in the short term, and the average error is 1.03%. However, because the modeling uses the smoothing data, the prediction results are underestimated, and the model error is expanding. At the same time, it is predicted that the number of tourists to Lanzhou in the autumn of 2020 will reach 81045529.

Table 2 ARCH test

	AIC	SC	HQ
Lag order 3	59.53666	59.50575	59.15463
Lag order 2	59.06760	59.09739	58.86668
Lag order 1	58.98383	59.02766	58.88925

Table 3 Comparison of actual and predicted passenger flow to Lanzhou during 2009–2019

Years	Actual value	Predictive value	Relative error (%)	Average error (%)
2009	70,01000	69,213,96	-1.14	1.03
2010	88,75000	86,820,37	-2.17	
2011	14,036000	13,726,730	-2.20	
2012	21,015100	20,804,175	-1.00	
2013	26,025200	25,844,980	-0.69	
2014	30,307700	29,978,300	-1.09	
2015	37,001800	37,101,749	0.27	
2016	53,375700	52,987,680	-0.73	
2017	54,314000	54,283,760	-0.05	
2018	67,185600	66,904,703	-0.42	
2019	75,240032	74,826,047	-0.55	

5 Conclusion

In this paper, using the relevant theory of time series, the AR(3) model is used to predict the number of tourists in Lanzhou City during 2009–2019. The experimental results show that using AR(3) model to better predict the short-term forecast of Lanzhou tourism population, this is a feasible way for Lanzhou tourism managers to predict the development of tourism market. Only considering the impact of the opening of the Baolan high-speed rail on the tourist population in Lanzhou, there is a one-sidedness. On the other hand, the ARIMA model is used to make short-term predictions of the number of tourists. The applicability of long-term prediction has not been studied in depth, and there are deficiencies.

Acknowledgments: This work is supported by science project of national railway administration of PRC: Research on High-Speed Railway Tourist Economy (KF2019-011-A).

References

- [1] Li Wei, (2019), Forecast of the number of tourists based on the ARIMA model, Jiangxi University of Finance and Economics.
- [2] Kong Chaoli, (2019), Hainan monthly tourist population forecast based on ARIMA model, China Science and Technology Information, 05, 76-78.
- [3] Liu Ruoyu, Liu Libo, (2019), Analysis and prediction of the number of visitors based on the ARIMA model, Computer and Telecom, Z1, 1-4.
- [4] Liu Yanyu, Long Jinping, (2016), An empirical analysis of the number of tourists in Jiangkou County based on the ARIMA model, Statistics and Management, 06, 36-37.
- [5] Chen Peng, Wu Ling, Song Hui, (2012), Forecast of the number of inbound tourists in Anhui Province based on ARIMA model, Journal of Anhui Agricultural University (Social Science Edition), 21, 32-35+116.
- [6] Cheng Lingqiang, (2019), High-speed rail travel train to create a win-win situation, Qingdao Daily, 006.
- [7] Yan Xue, Zhang Hong, (2019), The Impact of High-speed Rail on the Agglomeration and Diffusion of Tourism Industry—Taking Beijing-Tianjin-Hebei as an Example, Resources Development and Market, 35, 986-992.
- [8] Zhang Lizhong, Zhang Tianwen, Zhang Letong, Zhao Linlong, (2019), Study on the Influence of “High-speed Railway

Era” on Guilin Tourism Economic Development——Taking Guiguang High-speed Railway as an Example, Hubei Agricultural Sciences, 58, 159-161.