Applied Mathematics, 2016, 7, 171-176
Published Online February 2016 in SciRes. http://www.scirp.org/journal/am
http://dx.doi.org/10.4236/am.2016.73016

ARIMA Model in the Application of Shanghai and Shenzhen Stock Index

Shichang Shen[1], Yue Shen[2]

[1]School of Mathematics and Statistics, Qinghai Nationlities University, Xining, China
[2]School of Statistics and Mathematics, Zhongnan University of Economics and Law, Wuhan, China
Email: 13909785766@163.com

Received 13 January 2016; accepted 21 February 2016; published 24 February 2016

Abstract

In the paper, based on the data of Shanghai and Shenzhen 300 stock index in 2011, the *ARIMA* model was established by using Eviews 6, and the historical trend of stock price was found out. The model was used to provide a reference for the investors.

Keywords

Time Series, *ARIMA*, Stock Price Prediction

1. Introduction

Stock in the trading market as a trading object, with the same goods, has their own market and market prices. Because of the stock price to be affected by many factors such as company management, supply and demand, bank interest rate, public psychology and so on, it has a lot of uncertainty.

Shanghai and Shenzhen 300 index is a stock exchange in Shanghai and Shenzhen Stock Exchange in April 8, 2005 to reflect the overall trend of the stock market index A. Shanghai and Shenzhen 300 index sample covering the Shanghai and Shenzhen stock market around 60% of the market value, with good market representation and investment. The goal of Shanghai and Shenzhen 300 index is to reflect the profile and operation status of Chinese stock market stock price changes, and as the criteria for the evaluation of the investment performance, the index of investment and index derivative product innovation to provide basic conditions. So it is necessary for us to find a way to predict the stock price. In recent years, there have been papers investigating the problem (see [1]-[4]). Time series analysis is a very good method [1]. Based on this, this paper in Shanghai and Shenzhen 300 stock index data as the basis, through time series analysis method established the Shanghai and Shenzhen 300 stock index prediction model, and the predictive effect was detected by Eviews 6.0 software [2]. Predictive results provide a reference to the decision-making of investors.

How to cite this paper: Shen, S.C. and Shen, Y. (2016) *ARIMA* Model in the Application of Shanghai and Shenzhen Stock Index. *Applied Mathematics*, **7**, 171-176. http://dx.doi.org/10.4236/am.2016.73016

2. *ARIMA* Model

By 2011, the Shanghai and Shenzhen 300 stock index of 242 data (Due to the holidays, the stock market halted, some months of data is relatively few) as a time series analysis, a prediction model is established which is used in the modeling of 234 data and the prediction of the model is based on the following 8 data. Data comes from the financial research database (RESSETDB) (see **Attached Table**).

2.1. Data Preprocessing

The original data into a line chart was draw. A sequence of *Y* was written, as shown in **Figure 1**. **Figure 1** shows that the data have a downward trend and there are no periodic fluctuations. The initial judgment of the sequence is a non stationary series.

In order to reduce the fluctuation of the sequence, the natural logarithm transformation of the original data is still showing obvious non-stationary, so it is necessary to carry on the differential operation to the data, until after the two order difference, the sequence is obviously smooth. The two order difference is shown in **Figure 2**.

2.2. Model Identification

Autocorrelation function and partial autocorrelation function are the most important tools for the identification of $ARMA$ model. In Eviews 6, the model identification and order determination are usually carried out using a sample of the autocorrelation and partial autocorrelation analysis. Draw autocorrelation and partial correlation diagram of the series as in **Figure 3**.

From **Figure 3**, when $k=1$, the autocorrelation coefficient is beyond the random range, when $k\geq 2$, the autocorrelation coefficients are all fall within the random interval. The autocorrelation function is truncated. In partial autocorrelation analysis, until the lag phase $k=9$, the partial autocorrelation coefficient of the sequence is clearly within the confidence interval. That sequence of partial autocorrelation function is tailing. Therefore, the sequence of *Y* can be established $ARIMA(0,2,m)$ model.

2.3. Parameter Estimation

We can judge the type of time series model more accurately according to the principle of the model [1]. According to the principle can be calculated in **Table 1**, which takes $M=\left[\sqrt{242}\right]=15$, $\frac{1}{\sqrt{242}}=0.0643$.

From the table can be seen as the 1 step truncation, but after 6 steps are not censored, can think the tail, which belongs to $ARIMA(0,2,m)$ model.

In order to determine the order number of the model, the $ARIMA(0,2,1)$, $ARIMA(0,2,2)$, $ARIMA(0,2,3)$ model was established by using the least square method in Eviews 6. Now look at the different models under

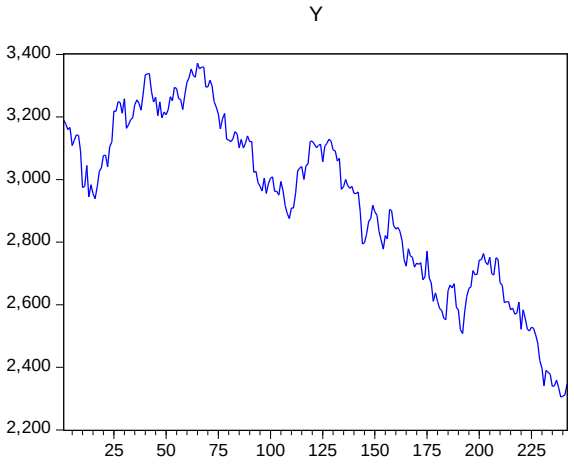


Figure 1. Line chart.

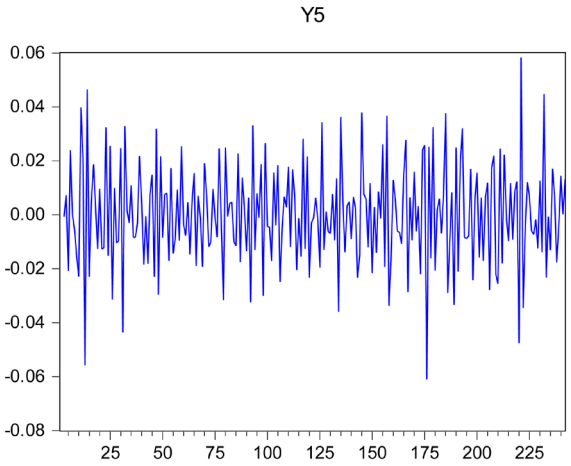


Figure 2. Two order difference chart.

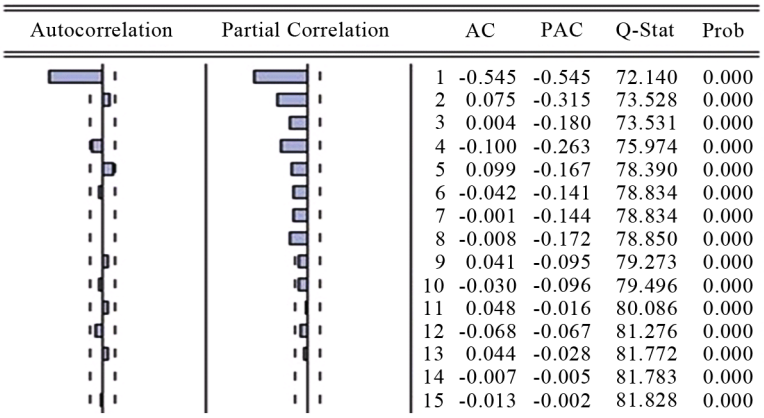


Figure 3. Correlation chart.

Table 1. Autocorrelation-partial correlation analysis value.

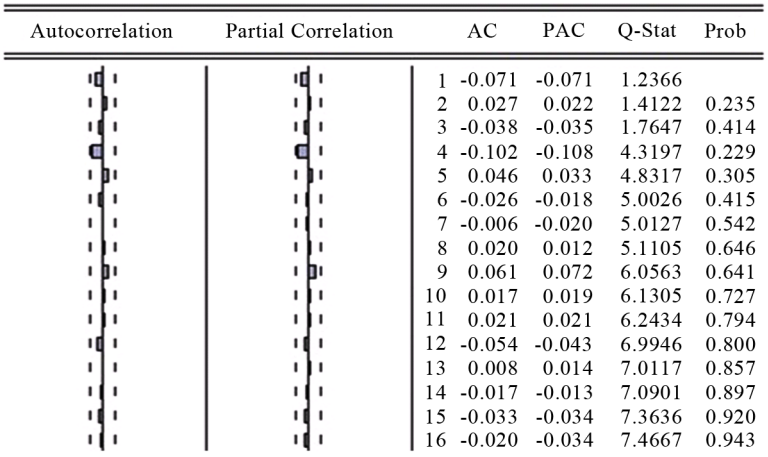
m	1	2	3	…	…
ρ_k	0.084170	0.084467	0.084467	…	…
To meet the conditions of proportion p	13/15 = 0.87	13/15 = 0.87	14/15 = 0.93	…	…
n	1	2	3	…	6
To meet the conditions of proportion p	10/15 = 0.67	9/15 = 0.60	10/15 = 0.67	…	6/15 = 0.446

Adjusted R^2, *AIC*, *SC*, such as **Table 2**.

In regression analysis, the requirement for the level of parameter t test is not so strict as the regression equation, and more is considered the whole fitting effect of the model. Adjusted R^2, *AIC*, *SC* are important criteria for the selection of models. And the three roots are within the unit circle, to meet the requirements. According to the standard function method, *AIC*, *SC* value reaches minimum is the best model order, so we choose $ARIMA(0,2,1)$ model.

2.4. Model Test

We should further verify the suitability of the model, that is, the residual sequence of the model is tested by white noise. The use of Eviews software for the chi square test, the test results are shown in **Figure 4**.



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.071	-0.071	1.2366	
		2	0.027	0.022	1.4122	0.235
		3	-0.038	-0.035	1.7647	0.414
		4	-0.102	-0.108	4.3197	0.229
		5	0.046	0.033	4.8317	0.305
		6	-0.026	-0.018	5.0026	0.415
		7	-0.006	-0.020	5.0127	0.542
		8	0.020	0.012	5.1105	0.646
		9	0.061	0.072	6.0563	0.641
		10	0.017	0.019	6.1305	0.727
		11	0.021	0.021	6.2434	0.794
		12	-0.054	-0.043	6.9946	0.800
		13	0.008	0.014	7.0117	0.857
		14	-0.017	-0.013	7.0901	0.897
		15	-0.033	-0.034	7.3636	0.920
		16	-0.020	-0.034	7.4667	0.943

Figure 4. Autocorrelation-partial autocorrelation analysis of residual sequence.

Table 2. Adjusted R^2, *AIC*, *SC* value.

$ARIMA(0,2,m)$	1	2	3
Adjusted R^2	0.524603	0.524656	0.523013
AIC	−5.826493	−5.82246	−5.81489
SC	−5.811990	−5.79346	−5.77138

In **Figure 4**, the last two columns for the chi square test, including with *Q* statistics and Adjoint probability of test. The sample size of the residual sequence is 240, and we take the maximum lag period is 15. From the figure, we can see the *Q* value 7.3636, and the probability that the first class error committed by the prob. column is 0.920. This shows that the residual sequence is independent of each other, test pass. That is, $ARIMA(0,2,1)$ model through test.

3. Model Prediction

The model is suitable for the test, can be used for short-term prediction. In order to test the predictive effect of the model, we set aside the last 8 observations in December as the reference object. After the operation of the software, the result of the equation is obtained. The main contents are shown in **Table 3**.

The inverted root of the polynomial in **Table 3** is in the unit circle, shows that the process is stable, and it is also reversible. We use the software to predict the last 8 values of the model. Use $ARIMA(0,2,1)$ model in December of 2011 last eight group in Shanghai and Shenzhen 300 stock index data to predict, the predictive value and the real value, error, error ratio is shown in **Table 4**.

Experimental results show the absolute error of the model and the percentage of absolute error are controlled within a certain range. So the fitting effect of the model is good, and the predictive value is close to the actual value.

4. Conclusion

According to $ARIMA(0,2,1)$ model, we got the last eight values in December. The error is controlled in less than 3% predicted value of stock buying and selling to make a short-term rational decision, for the risk of investing in stocks decrease will have a certain role. The research of this paper is limited to the establishment of the model of the stationary processing of the finite- and non-stationary data. Through the historical data of Shanghai and Shenzhen 300 stock index, it reveals the law of its change with time. Extending this law to the future, so as to predict the future of the Shanghai and Shenzhen 300 stock price index, the fitting effect is not perfect. However, the time series prediction model described by *ARMA* model in the financial, stock and other fields has its theoretical and practical significance. The stock price through the fitting and prediction, time series model has certain reference in the aspect of price volatility of the stock market. The result of fitting prediction

Table 3. Model parameter estimation and correlation test results.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
$ARIMA(0,2,1)$	−0.988651	0.007213	−137.0736	0.0000
R-squared	0.524603	Mean dependent var.		7.97E−05
Adjusted R-squared	0.524603	S.D. dependent var.		0.019016
S.E. of regression	0.013112	Akaike info criterion		−5.826493
Sum squared residual	0.041087	Schwarz criterion		−5.811990
Log likelihood	700.1792	Hannan-Quinn criter.		−5.820650
Durbin-Watson statistic	2.127480			
Inverted MA roots	0.99			

Table 4. Model prediction analysis table.

Time	Real value	Predictive value	Error	Error ratio (%)
1	2339.1	2372.733	33.633	1.44%
2	2341.3	2368.373	27.073	1.16%
3	2359.2	2364.022	4.8222	0.20%
4	2335.7	2359.679	23.979	1.03%
5	2305	2355.344	50.344	2.18%
6	2307.9	2351.016	43.116	1.87%
7	2311.4	2346.697	35.297	1.53%
8	2345.7	2342.386	−3.314	−0.14%

can represent the trend of stock price in a certain degree.

Fund

Qinghai Nationalities University Natural Science Foundation Item Number 2015XJZ03.

References

[1] Wang, Z.L. and Hu, Y.H. (2007) Application of Time Series Analysis. Science Press, Beijing.

[2] Yi, D.H. (2008) Data Analysis and Eviews Application. Renmin University of China Press, Beijing.

[3] Fu, S., Peng, B.Y. (2009) *ARMA* Model in the Application of China Stock Market. *Journal of Hengyang Normal University*, **30**, 26-28.

[4] Li, M. (2000) Forecasting the Shanghai and Shenzhen Stock Market with *ARMA* Model. *Journal of Changsha Railway Institute*, **3**, 78-83.

Attached Table. Shanghai and Shenzhen 300 stock index data in 2011.

Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
3189.7	3041.0	3254.9	3272.7	3211.1	3004.2	3049.7	2977.7	2834.5	2557.1	2697.5	2583.6
3175.7	3104.2	3243.3	3311.1	3129.0	2955.7	3122.0	2956.4	2803.9	2552.0	2742.4	2557.3
3159.6	3121.0	3221.7	3324.4	3126.1	2986.4	3122.5	2954.9	2743.8	2644.8	2744.3	2521.4
3166.6	3219.1	3270.7	3353.4	3121.4	3004.3	3113.7	2960.3	2723.3	2662.6	2763.8	2516.3
3108.2	3217.7	3334.5	3333.4	3129.8	3008.7	3101.7	2897.4	2779.1	2653.8	2736.3	2528.2
3124.9	3248.5	3337.5	3326.8	3153.2	2961.9	3109.2	2793.9	2756.1	2667.0	2727.7	2525.0
3142.3	3245.9	3338.9	3372.0	3145.1	2961.9	3113.2	2798.2	2751.1	2592.2	2751.7	2503.5
3141.3	3211.9	3280.3	3353.6	3101.1	2950.4	3056.9	2824.1	2720.3	2583.1	2699.6	2477.7
3091.9	3257.9	3247.4	3358.9	3128.1	2993.6	3106.3	2866.9	2733.1	2520.5	2695.0	2421.9
2974.4	3163.6	3262.9	3359.4	3100.5	2963.1	3115.7	2875.4	2729.1	2507.9	2750.2	2397.5
2977.7	3174.7	3204.0	3295.8	3116.0	2917.6	3128.9	2917.9	2734.0	2576.7	2744.7	2340.8
3044.9	3190.9	3248.2	3295.8	3139.4	2892.2	3122.6	2897.6	2679.3	2625.4	2670.1	2390.1
2944.7	3197.6	3197.1	3317.4	3120.6	2874.9	3095.1	2886.0	2689.8	2651.7	2662.0	2384.4
2983.5	3239.6	3215.7	3299.9	3121.6	2909.1	3091.6	2834.3	2771.0	2657.5	2606.5	2377.1
2954.2		3207.1	3249.6	3023.0	2908.6	3059.1	2807.7	2685.7	2709.0	2609.7	2339.1
2938.7		3223.0	3231.0	3026.2	2957.6	3068.0	2777.8	2669.5	2695.3	2609.5	2341.3
2978.4		3264.9	3209.5	2990.3	3027.5	2968.3	2821.0	2610.9		2584.0	2359.2
3026.5		3251.4	3161.8	2978.4	3036.5	2977.8	2810.0	2637.9		2588.9	2335.7
3036.7		3294.5	3192.7	2963.3	3041.7	3000.1	2903.8	2610.6		2570.0	2305.0
3076.5		3290.6			3000.2	2981.0	2901.2	2588.2		2573.3	2307.9
3077.3		3258.0			3044.1	2972.1	2852.8	2581.4		2608.6	2311.4
		3256.1					2841.7			2521.5	2345.7
		3223.3					2846.8				

Note: the stock market halted due to holidays, some months of data is relatively few.