Empirical Analysis of Consumer Price Index of Jinan Based on ARIMA Model

Xingxiang Qi[1,a]

[1]*School of Economics, Shandong University of Finance and Economics*

[a]*Email: qixingxiang@163.com*

ABSTRACT

This paper selects the CPI data of Jinan City from 2011 to 2020 as the sample, and establishes ARIMA (11, 1, 12) consumer price index model for empirical analysis. The results show that the model is accurate. The CPI data in 2020 is predicted, and the predicted value is very close to the real value. Finally, the CPI in the first half of 2021 is predicted by using the model.

Keywords: *CPI, ARIMA, Economy Index Forecasting, Jinan*

1. INTRODUCTION

The consumer price index(CPI) is a very important economic statistical index, which reflects the price level and the living standard of residents in a certain period. Jinan is always the provincial capital of Shandong Province which is the coastal economic big province of East China. Jinan is one of sub-provincial cities around the county, central city of Circum-Bohai Sea Region South Wing. Jinan's economic problems, such as CPI, are worth studying.

Many articles use ARIMA model to study the empirical analysis of CPI. ZHANG Ben-li and Zhang Xiao-qing(2010) analysis consumer price index of Shandong province based on ARIMA(3,2,1) model[1]. Based on the monthly data of Hubei Province from September 2008 to September 2011 (32 months), Zhou Mei-ying(2011) establishes ARIMA (10,1,3) model and analyses the CPI time series of Hubei Province[2]. GUO Xiao-feng(2012) selects the monthly CPI data of China from January 2001 to October 2011 as samples, and establishes ARIMA (12,1,20) model[3]. Based on the monthly CPI data from January 2000 to December 2012, Li Xuan and Huang Dongdong build ARIMA model to analyse and forecast the CPI in the first half of 2013. The empirical results show that ARIMA (3,1,3) model can well describe CPI and provide better prediction[4]. CUI Wen-yan and XU Feng-hua(2016) establish the ARIMA(0,1,2) model on Shandong Province's CPI annual data during 1995-2011[5]. CHU Rui(2018) selects the monthly data of the CPI in January 2005 to June 2017 in Beijing as the sample, and the model ARIMA (0,1,12)

was constructed[6]. Zhang Yangyang, Deng Wei and Song Changyu(2020) use ARIMA(2,1,2) model to fit the CPI data of Shandong Province from January 1995 to May 2020 and predict the CPI of Shandong Province from June to October 2020[7]. LI Shu, WANG Fei and WANG Feng-xiao(2020) take the CPI monthly data of Qingdao from January 2010 to June 2019 as the sample, and establish the ARIMA (11,1,12) model[8]. In reference [1]-[8], although the model parameters are different, they all use ARIMA to fit and predict CPI data in different periods. This paper selects the CPI data of Jinan City from 2011 to 2020 as the sample, and uses ARIMA model to fit and predict.

2. ARIMA MODEL

ARIMA (p, d, q) is the abbreviation of auto regressive integrated moving average model, which is used for the analysis of stationary series or series stabilized by difference. If the d-th difference w_t of a time series $\{Y_t\}$ is a stationary ARMA process, it is called autoregressive moving average summation model. If w_t obeys ARMA (p, q) model, $\{Y_t\}$ is called ARIMA (p, d, q) process.

$$w_t = c + \phi_1 w_{t-1} + \phi_2 w_{t-2} + \cdots + \phi_p w_{t-p}$$
$$+ e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \cdots + \theta_q e_{t-q} \qquad (1)$$

Where c is a constant, φ is the coefficient of the autoregressive model, θ is the coefficient of the moving average model, and e_t is a constant white noise sequence.

When d = 0, ARIMA model becomes ARIM (p, q) model; p represents the maximum order of autoregressive process equation; when d = 0 and P = 0, ARIMA model

becomes MA (q); when d = 0 and q = 0, ARIMA model becomes AR (p) model. Therefore, AR model, MA model and ARMA model are all special ARIMA models.

3. EMPIRICAL ANALYSIS

3.1. Data collection and arrangement

This paper collects the monthly CPI data of Jinan from 2011 to 2020 from the website of Jinan Municipal Bureau of statistics. The results are shown in Table 1.

Table 1. Monthly CPI data of Jinan from 2011 to 2020

Year	Jan.	Feb.	Mar.	Apr.	May.	Jun.	Jul.	Aug.	Sept.	Oct.	Nov.	Dec.
2011	103.4	103.9	104.5	105.3	105.4	106.9	106.9	106.6	106.9	106.4	105.2	103.7
2012	104.1	103.5	103.4	103.1	102.8	102.6	102.5	102.4	102.4	102.3	102.3	102.4
2013	103.9	104.7	103.2	103.8	103.1	103.1	101.9	101.8	102.0	102.6	101.8	101.4
2014	101.1	101.4	102.5	101.9	102.5	103.2	102.8	102.5	102.4	102.0	102.2	102.3
2015	101.2	101.9	101.3	101.3	101.9	101.7	102.4	102.9	102.6	102.0	102.0	101.8
2016	102.1	102.3	102.6	102.6	102.0	102.6	102.8	102.5	103.3	103.2	103.4	103.6
2017	103.7	101.9	101.6	102.1	102.3	101.7	101.4	101.8	101.4	102.0	101.8	102.0
2018	101.4	103.1	102.5	102.4	102.6	102.6	102.4	102.9	103.2	103.2	102.4	102.2
2019	102.3	101.9	102.8	102.5	102.5	102.7	103.2	102.8	103.1	104.4	105.7	105.5
2020	106.1	104.7	103.9	103.3	103	102.6	102.8	103.0	101.9	100.0	98.4	99.3

Data source: Website of Jinan Municipal Bureau of Statistics; Preceding Last Year=100.

3.2. Tests of stationarity

The Augmented Dickey Fuller (ADF) test was used to test the unit root of CPI data and its first-order difference data. The results are shown in Table 2.

Table 2. ADF test results of CPI data and its first-order difference data

Variable	ADF value	Prob.*	1% critical value	5% critical value	10% critical value	Stationarity
CPI	-2.0525	0.2644	-3.4861	-2.8859	-2.5798	Unstable
D(CPI)	-10.2430	0.0000	-3.4866	-2.8861	-2.5799	Stable

According to table 2, the p value of the original series CPI is much larger than 0.05, so we should accept the null hypothesis that the series has unit root and is a non-stationary time series. Therefore, we need to do differential processing. The p value of the first-order difference sequence D(CPI) is less than 0.05, so the null hypothesis is not accepted, that is, the sequence is not considered to have unit root and is a stationary time series. Thus, the first order difference sequence is a stationary time series.

From the above analysis, we can see that the time series after the first-order difference is a stationary non-white noise series, which can fit the ARIMA (p, d, q) model. In order to obtain the relatively optimal fitting model, the values of p and q are determined according to the autocorrelation coefficient and partial autocorrelation coefficient, and different ARIMA (p, d, q) models are tried to be established.

3.3. Establishment of ARIMA model

The values of p and q in ARMA (p, q) are determined by autocorrelation function and partial autocorrelation function. The partial autocorrelation function (PACF) determines the value of p and the autocorrelation function (ACF) determines the value of q. The analysis results of autocorrelation function and partial autocorrelation function are shown in Figure 1.
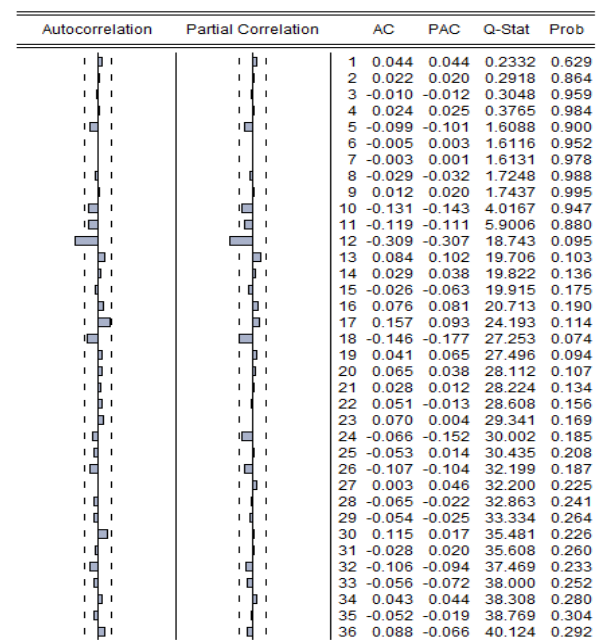

Figure 1 Autocorrelation and partial autocorrelation analysis of D (CPI).

According to the statistical characteristics of autocorrelation function and partial autocorrelation function in Figure 1, the values of P and Q in ARIMA model are preliminarily determined. In order to obtain a better fitting model, different ARIMA (P, 1, q) models are tried to be established for parameter estimation. The regression results are shown in table 3.

Table 3. Regression results of different parameter combinations in ARIMA model

	ARIMA(11,1,11)	ARIMA(11,1,12)	ARIMA(12,1,11)	ARIMA(12,1,12)
AR(11)	-0.1050(0.2211)	-0.2196(0.0976)**		
AR(12)			-0.4136(0.0942)***	0.0056(0.0988)
MA(11)	-0.0563(0.2421)		-0.2190(0.1068)**	
MA(12)		-0.8853(0.0230)***		-0.8748(0.0285)***
Adjusted R-squared	0.0086	0.4224	0.1638	0.4011
Akaike info criterion	1.9644	1.4242	1.7993	1.4654
Schwarz criterion	2.0140	1.4738	1.8492	1.5154

Standard errors in parentheses; * $p<0.05$, ** $p<0.01$, *** $p<0.001$

ARIMA (11,1,12) model has smaller values of Akaike info criterion and Schwarz criterion, and the adjusted R-squared is the largest, and the model is simple and feasible. Therefore, this model is selected.

3.4. Test of ARIMA model

The comparison of actual value, fitting value and residual value of ARIMA (11,1,12) model is shown in Figure 2. Autocorrelation and partial autocorrelation analysis of residual value is shown in Figure 3.
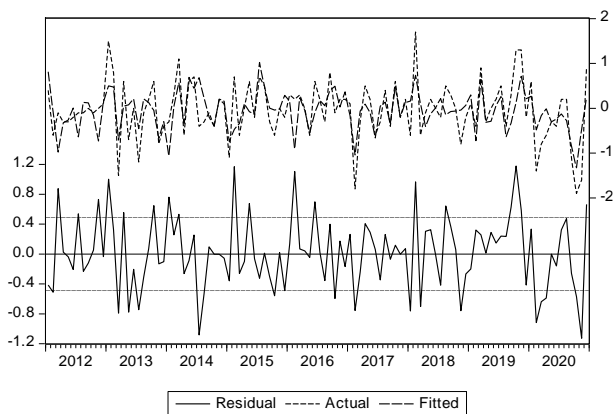

Figure 2   Actual value, fitted value and residual value of ARIMA (11,1,12) model.

In Figure 2, the change of fitting value and actual value of the model is consistent, and the residual of the model fluctuates around the zero mean value. As shown in Figure 3, the autocorrelation coefficient basically falls into the confidence band with significance level of 0.05, that is, the residual sequence of the model is a random white noise sequence, and the ARIMA (11, 1, 12) model established is more reasonable. LM test results also show that there is no correlation between the residual series.
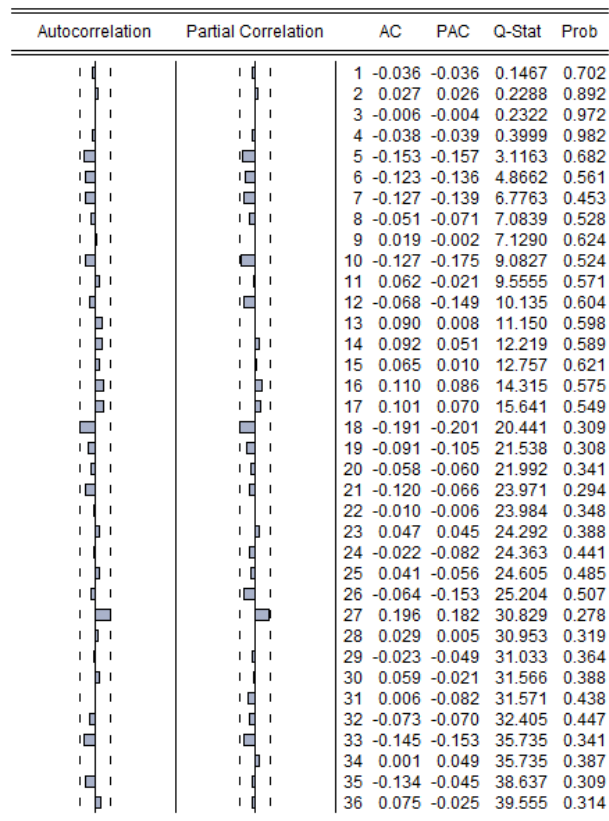

Figure 3   Autocorrelation and partial autocorrelation analysis of residual value.

The final model is ARIMA (11,1,12), which is significantly effective and can be directly used for prediction. The formula is as follows.

$$\Delta CPI_t=-0.2196\Delta CPI_{t-11}+e_t-0.8853e_{t-12} \qquad (2)$$

3.5. Prediction of ARIMA model

The ARIMA (11,1,12) model is applied to forecast the monthly CPI of Jinan in 2020, and the comparison between the predicted value and the real value is shown in Table 4.

Table 4.   The actual value and forecast value of monthly CPI in Jinan in 2020 (Preceding Last Year=100)

	Jan.	Feb.	Mar.	Apr.	May.	Jun.	Jul.	Aug.	Sept.	Oct.	Nov.	Dec.
Actual value	106.1	104.7	103.9	103.3	103	102.6	102.8	103.0	101.9	100.0	98.4	99.3
Forecast value	105.8	105.6	104.5	103.9	103.0	102.8	102.5	102.5	102.2	100.6	99.5	98.6
Relative error	0.28%	0.86%	0.58%	0.58%	0.00%	0.19%	0.29%	0.49%	0.29%	0.60%	1.12%	0.70%

It can be seen from table 4 that the monthly CPI data of Jinan in 2020 fitted by ARIMA (11,1,12) model has very small error with the actual value, and the error in March is zero, indicating that the model fits well and the prediction result is ideal, which can be used for the prediction of future data. Using ARIMA (11,1,12) model to forecast the monthly CPI data of Jinan in the first half of 2021, the results are shown in Table 5.

Table 5.   Forecast results of monthly CPI data of Jinan in the first half of 2021

	Jan.	Feb.	Mar.	Apr.	May.	Jun.
Forecast value	99.6	101.2	102.1	102.7	102.7	102.8

From the forecast results, the CPI data from January to May in 2021 compared with the same period of last year, has decreased every month, but the overall trend is slightly upward. Since November 2020, the trend of CPI declining has changed.

4. CONCLUSION

To a certain extent, ARIMA model can use the analysis of past and present values to predict the future development of time series. ARIMA model is more suitable for short term forecasting. With the extension of forecasting period, the forecasting error will increase accordingly. Using the monthly CPI data of Jinan in recent 10 years, the ARIMA model is used to fit and forecast. The CPI data of the first half of 2021 can be used for reference by economic management departments to make macro-control measures in advance to ensure the stability and long-term of economic development.

REFERENCES

[1] ZHANG Ben-li, ZHANG Xiao-qing, The Analysis of Consumer Price Index of Shandong Province Based on ARIMA Model, Ludong University Journal(Natural Science Edition), vol.26, Yantai, Shandong, 2010, pp.285-288.

[2] ZHOU Mei-ying, Time series analysis and prediction of CPI in Hubei Province Based on ARIMA model, Kunming, Yunnan, Times Finance, 2011, pp. 66,110.

[3] GUO Xiao-feng, Prediction and analysis of China's CPI trend based on ARIMA model, Statistics and Decision, Wuhan, Hubei, 2012, pp . 29-32. DOI:10.13546/j.cnki.tjyjc.2012.11.004

[4] Li Xuan, Huang Dongdong, Analysis and Forecast of CPI based on ARIMA Model, Journal of Shenyang University(Social Science), vol.15, Shenyang, Liaoning, 2013, pp. 306-310.

[5] CUI Wen-yan, XU Feng-hua, The Application of ARIMA Model on the Prediction of CPI in Shandong Province, Journal of Zaozhuang University, vol. 33, Zaozhuang, Shandong, 2016, pp. 24-28.

[6] CHU Rui, Forecast and Analysis of CPI in Beijing Based on ARIMA Model, Value Engineering, Shijiazhuang, Hebei, 2018, pp. 21-22. DOI:10.14018/j.cnki.cn13-1085/n.2018.05.009

[7] Zhang Yangyang, Deng Wei, Song Changyu, Prediction of CPI in Shandong Province Based on Time Series Analysis, Journal of Shandong Normal University(Natural Science), Jinan, Shandong, 2020, pp. 329-336. DOI:10.3969/j.issn.1001-4748. 2020.03.010

[8] LI Shu, WANG Fei, WANG Feng-xiao, Application of ARIMA Model in Consumer Price Index Forecasting of Qingdao, Journal of Lanzhou University of Arts and Science(Natural Sciences), Lanzhou, Gansu, 2020, pp. 11-16. DOI:10.13804/j.cnki.2095-6991.2020.06.003